International Journal Evolving Sustainable and Renewable Energy Solutions ISSN: 3068-9325

PRIME OPEN ACCESS

Volume 2, Issue 3

Research Article

Date of Submission: 01 Oct, 2026
Date of Acceptance: 19 Oct, 2026
Date of Publication: 31 Oct, 2026

Performance Analysis for Leveraging AI to Reduce Energy Consumption of a Data Center: A Safe Deep Reinforcement Learning and Workload-Aware Framework

Mohammad Ali*

School of Science and Engineering, Atlantic International University, Honolulu, HI 96813, USA

***Corresponding Author:** Mohammad Ali School of Science and Engineering, Atlantic International University, Honolulu, HI 96813, USA.

Citation: Ali, M. (2026). Performance Analysis for Leveraging AI to Reduce Energy Consumption of a Data Center: A Safe Deep Reinforcement Learning and Workload-Aware Framework. *Int J Evol Sus Renew Energy Sol, 2*(3), 01-14.

Abstract

Electricity demand from data centers is rising quickly with the deployment of artificial intelligence (AI), and cooling remains the largest non-IT load, particularly in hot-arid climates. This paper presents a performance analysis of leveraging AI to reduce the energy consumption of a 10 MW AI/cloud data center located in Abu Dhabi. A component-level energy model is developed that couples server and accelerator power, uninterruptible power supply (UPS) losses, airflow and hot-air recirculation, server-fan penalties, air-cooled chiller efficiency and dry-cooler economization. On this digital twin, a five-layer framework integrates a long short-term memory (LSTM) load forecaster, a double deep Q-network (DDQN) that sets supply-air temperature, chilled-water temperature and airflow under a digital-twin safety shield, forecast-driven server consolidation and thermal-aware shifting of flexible workloads. Trained on one synthetic year and tested on an independent year against a static baseline, an ASHRAE-optimized rule-based strategy and model-based oracle bounds, the LSTM gives the lowest day-ahead error (MAPE 3.04%). The complete framework reduces annual facility energy by 15.3% (13.85 GWh), cooling energy by 40.8% and peak demand by 14.1%, and improves power usage effectiveness (PUE) from 1.421 to 1.291, while holding the 99th-percentile rack inlet temperature below 27 °C. It captures 99.9% of the oracle-attainable saving and pays back in 1.7 years under stated assumptions. The results also show that PUE alone can misrepresent AI-driven savings.

Index Terms: Artificial Intelligence, Data Center, Deep Reinforcement Learning, Energy Efficiency, Long Short-Term Memory, Power Usage Effectiveness, Safe Reinforcement Learning, Thermal Management, Workload Consolidation

Introduction

Data centers underpin cloud services, digital commerce and, increasingly, generative artificial intelligence (AI). The International Energy Agency (IEA) estimates that data centers consumed about 415 TWh in 2024, roughly 1.5% of global electricity, and projects about 945 TWh by 2030 in its base case, with accelerated AI servers as the main driver. Recent national assessments confirm a renewed acceleration of data center demand after a decade in which efficiency gains offset growth, and large-scale AI training introduces new power-quality and grid interaction challenges [1-4].

The power usage effectiveness (PUE), the ratio of total facility energy to IT energy, has stagnated at about 1.54 on a global average. Whereas leading hyperscale operators report values close to 1.1. Most of the gap is cooling energy. The gap is widest in hot-arid regions such as the United Arab Emirates, where high ambient temperatures raise chiller lift, limit economizer hours and discourage evaporative heat rejection because of water scarcity [5-10].

AI can close part of this gap because data centers are densely instrumented, non-linear and strongly coupled: the best chilled-water temperature depends on weather, IT load and airflow, which in turn affect rack inlet temperatures, server-fan power and hot-spot risk [11,12]. Neural-network recommendations reduced cooling energy by 40% in a hyperscale facility [13]. Deep reinforcement learning (DRL) has since been applied in simulation and on live plants [14-16]. Recently with safety-aware and offline formulations; and forecasting and carbon-aware scheduling [17-21]. Have been used on the IT side. A 2025 systematic review of 65 RL studies, however, identifies the lack of joint IT–cooling optimization, of standardized multi-scale metrics and of rigorous validation as open problems [22].

This paper addresses these gaps for an AI-dominated facility in a hot-arid climate.

Its Contributions are as Follows:

- A transparent hourly energy model of a 10 MW AI/cloud data center, summarized as a system architecture in Fig. 1, that captures the trade-offs among supply-air temperature, chilled-water temperature, airflow provisioning, server-fan power and economizer availability.
- A five-layer AI framework that couples an LSTM forecaster, a DDQN cooling agent whose actions are masked by a digital-twin safety shield during training and deployment, forecast-driven consolidation and thermal-aware temporal workload shifting.
- A full-year, out-of-sample performance analysis against static, rule-based and oracle references, reporting energy, PUE, peak demand, thermal compliance, service impact, optimality gap, climate sensitivity and economics.
- Evidence that PUE can hide IT-side savings and reward the transfer of energy into server fans, with a recommendation to report absolute facility energy alongside PUE.
- Section II reviews related work. Section III presents the system model and Section IV the AI framework. Section V describes the case study, and Section VI analyzes the results. Section VII discusses implications and limitations, and Section VIII concludes the paper.

Related Work

AI-Based Cooling Control

Early data-driven work predicted PUE with neural networks for offline what-if analysis [23] and then issued cooling recommendations in a live facility [13]. Lazic et al. learned a linear floor model and applied model-predictive control (MPC) [16-24]. Li et al. formulated cooling as an off-policy actor–critic problem , and Luo et al. reported 9% and 13% energy savings from RL on two commercial plants [14,15]. More recent studies address safety and data efficiency: Wang et al. combine imitation learning with physics-guided post-hoc action rectification and report 18–26.6% power savings with far fewer thermal violations than reward shaping [17]. Chen et al. use a DQN with an XG Boost thermal model for multi-set-point control [25]. And Zhan et al. deploy offline RL with a physics-informed graph model in production data centers, reporting 14–21% cooling-energy savings without constraint violations [18]. Shielding, in which a monitor restricts an agent to verified-safe actions [26,27]. Has received less attention in data center cooling.

Workload Forecasting and IT-Side Optimization

Server power is approximately linear in utilization with a large idle floor [28]. Which motivates consolidation of lightly loaded servers [29]. Recurrent networks, particularly the LSTM [30]. And hybrid decomposition–LSTM models [19]. Achieve accurate workload forecasts, while classical ARIMA models [31]. And gradient-boosted trees [32]. Remain strong baselines. Temperature-aware placement [33]. And carbon-aware temporal shifting [20,21]. Exploit workload flexibility, and Deep EE jointly optimized job scheduling and airflow with a parameterized DQN, saving up to 15% energy against siloed baselines [34].

Research Gap

Table I summarizes representative studies. Three gaps remain. First, cooling and IT optimization are usually studied separately, so their interaction and relative contribution are rarely quantified. Second, results are seldom benchmarked against an explicit optimality bound, so the share of attainable savings captured by a learning controller is unknown. Third, evidence for GPU-dominated facilities in hot-arid climates, with scarce economization and high chiller lift, is limited [10-22].

Ref.	Year	Method	Controlled variables	Setting	Reported outcome
[13]	2016	Ensemble DNN recommendations	Cooling plant set-points	Live hyperscale DC	40% cooling energy; 15% PUE overhead
[34]	2019	Parameterized DQN (DeepEE)	Job scheduling + airflow	Simulation	Up to 15% vs. siloed baseline
[15]	2022	RL with constraint handling	Commercial chiller plant	Two live sites	9% and 13% energy
[20]	2023	Forecast-based scheduling	Flexible compute timing	Google fleet	Carbon-aware load shifting
[17]	2024	Physics-guided safe DRL	Cooling set-points	Simulation, traces	18–26.6% DC power; 94.5–99% fewer violations
[25]	2024	DQN + XGBoost thermal model	Multiple set-points	Air-cooled DC model	2.4–5.7% vs. other DRL controls
[18]	2025	Offline RL, physics-informed	Air-cooling units	Production DCs	14–21% cooling energy
This work	2026	LSTM + shielded DDQN + workload optimizer	T_{sa}, T_{chw}, α, active servers, batch timing	10 MW AI DC, hot-arid	Section VI; oracle-bounded

Table 1: Comparison of Representative Studies on Ai-Based Data Center Energy Optimization

System Modeling

System Architecture

Figure 1 shows the architecture of the model. Utility power supplies a 10 MW critical-IT load (a 3 MW CPU cloud cluster, 4 MW GPU training cluster, 2 MW GPU inference cluster and 0.5 MW of network and storage) through double-conversion UPS units and power distribution units (PDUs). Heat is removed by computer-room air handlers (CRAHs) with electronically commutated fans in a contained cold-aisle layout, supplied by air-cooled chillers and a dry-cooler water-side economizer, because water scarcity discourages evaporative heat rejection. The AI supervisory controller acts on the blocks tagged in Figure 1, and the model runs at a one-hour step Δt.

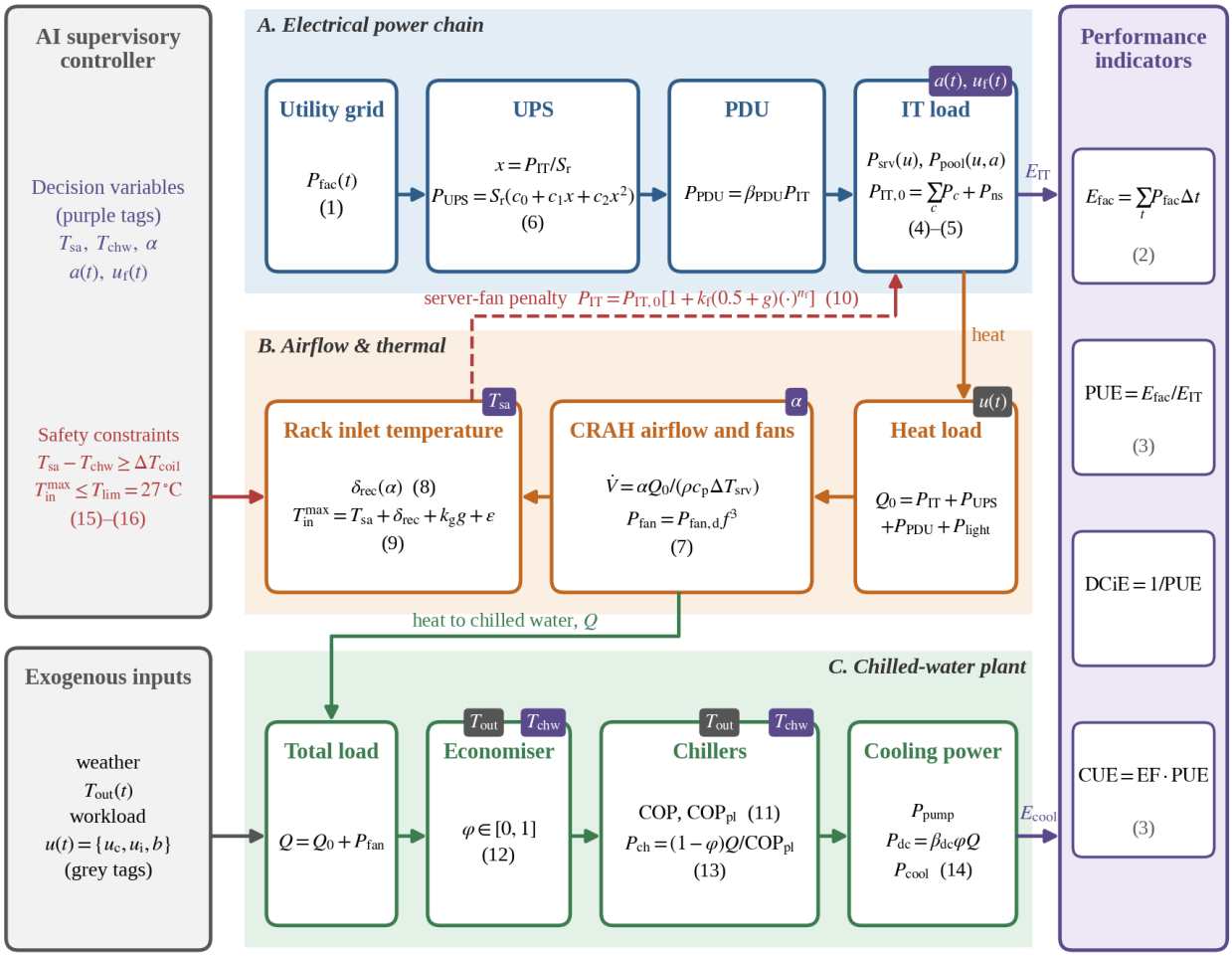


Figure 1: System Architecture of the Data Center Energy Model

Electrical chain (A), airflow and thermal subsystem (B) and chilled-water plant (C), annotated with the governing equations (1)–(16), the AI decision variables (purple tags), the exogenous inputs (grey tags) and the performance indicators.

Energy Balance and Indicators

The facility power and annual energy are

$$P_{fac} = P_{IT} + P_{UPS} + P_{PDU} + P_{cool} + P_{light} \quad (1)$$

$$E_{fac} = \sum_{t=1}^{T} P_{fac}(t)\, \Delta t \quad (2)$$

and the PUE and carbon usage effectiveness (CUE) for a grid emission factor EF are

$$PUE = \frac{E_{fac}}{E_{IT}}, \qquad CUE = EF \cdot PUE \quad (3)$$

C. IT and Electrical Model

A server with rated power P_{max}, idle ratio κ and utilization u draws [28]

$$P_{srv}(u) = P_{max}[\kappa + (1 - \kappa)\, u] \quad (4)$$

with κ = 0.50 for CPU servers and κ = 0.25–0.30 for GPU nodes. If a console datable pool of rated power P_r keeps a fraction a of nodes active and puts the rest in a sleep state of power ratio s, its power is

$$P_{pool} = P_r[a\kappa + (1 - \kappa)u + (1 - a)s], \quad u \leq a \quad (5)$$

so consolidation saves $P_r(1 - a)(\kappa - s)$. UPS losses follow the quadratic curve

$$P_{\mathrm{UPS}} = S_r(c_0 + c_1 x + c_2 x^2), \quad x = P_{\mathrm{IT}}/S_r \tag{6}$$

with S_r = 12 MW, and PDU and transformer losses are $P_{\mathrm{PDU}} = \beta_{\mathrm{PDU}} P_{\mathrm{IT}}$.

D. Airflow and Thermal Model

The heat to be removed is $Q_0 = P_{\mathrm{IT}} + P_{\mathrm{UPS}} + P_{\mathrm{PDU}} + P_{\mathrm{light}}$. With airflow provisioning ratio $\alpha \geq 1$, the normalized CRAH flow

$$P_{\mathrm{fan}} = P_{\mathrm{fan,d}}\, f^3, \qquad f = \max\left(f_{\min}, \frac{\alpha Q_0}{Q_{\mathrm{d}}}\right) \tag{7}$$

Lower α saves fan power cubically but increases hot-air recirculation, modeled as

$$\delta_{\mathrm{rec}} = \delta_0 + k_r[\max(0, \alpha_{\mathrm{ref}} - \alpha)]^{n_r} \frac{L}{L_{\mathrm{ref}}} \tag{8}$$

where $L = P_{\mathrm{IT},0}/Q_{\mathrm{d}}$ is the load factor. The maximum rack inlet temperature is

$$T_{\mathrm{in}}^{\max} = T_{\mathrm{sa}} + \delta_{\mathrm{rec}} + k_g\, g + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_T^2) \tag{9}$$

where g is the GPU share of IT power, which captures hot-spot severity in GPU rows, and ε represents sensor noise. With the mean inlet temperature $T_{\mathrm{in}}^{\mathrm{avg}} = T_{\mathrm{sa}} + \delta_{\mathrm{rec}}/2$, server fans and leakage raise IT power as

$$P_{\mathrm{IT}} = P_{\mathrm{IT},0}\left[1 + k_f(0.5 + g)\, \Delta_T^{\eta_f}\right] \tag{10}$$

where $\Delta_T = \max(0, T^{\mathrm{in}}_{\mathrm{avg}} - T_0)$. This term penalizes set-points that are too warm and creates an interior optimum.

Chilled-Water Plant Model

The chiller coefficient of performance (COP) is modeled as a fraction ηc of the Carnot value between the chilled-water temperature and the condensing temperature:

$$\mathrm{COP} = \eta_c \frac{T_{\mathrm{chw}} + 273.15}{T_{\mathrm{out}} + \Delta T_{\mathrm{cond}} - T_{\mathrm{chw}}} \tag{11}$$

A part-load correction $\mathrm{COP}_{\mathrm{pl}} = \mathrm{COP}[1 - 0.35(\mathrm{PLR} - 0.75)^2]$ is applied per staged chiller. With the chilled-water return temperature $T_{\mathrm{ret}} = T_{\mathrm{chw}} + \Delta T_{\mathrm{chw}}$, the dry cooler provides the free-cooling fraction

$$\varphi = \min\left[1, \max\left(0, \frac{T_{\mathrm{ret}} - T_{\mathrm{out}} - \Delta T_{\mathrm{ec}}}{\Delta T_{\mathrm{chw}}}\right)\right] \tag{12}$$

so that, with $Q = Q_0 + P_{\mathrm{fan}}$, the chiller power and the total cooling power are

$$P_{\mathrm{ch}} = \frac{(1 - \varphi)\, Q}{\mathrm{COP}_{\mathrm{pl}}} \tag{13}$$

$$P_{\mathrm{cool}} = P_{\mathrm{ch}} + P_{\mathrm{fan}} + P_{\mathrm{pump}} + P_{\mathrm{dc}} \tag{14}$$

where P_{pump} scales with the square of the load and $P_{\mathrm{dc}} = \beta_{\mathrm{dc}}\, \varphi Q$.

F. Operating Constraints

The CRAH coil requires a minimum approach, and the inlet temperature must stay within the ASHRAE recommended envelope [11]:

$$T_{\mathrm{sa}} - T_{\mathrm{chw}} \geq \Delta T_{\mathrm{coil}} \tag{15}$$

$$18\ ^\circ\mathrm{C} \leq T_{\mathrm{in}}^{\max} \leq T_{\mathrm{lim}} = 27\ ^\circ\mathrm{C} \tag{16}$$

Table II lists the model parameters.

Parameter	Symbol	Value
CPU / GPU-train / GPU-inference rating	P_r	3.0 / 4.0 / 2.0 MW
Idle ratio CPU / GPU-train / GPU-inf.	κ	0.50 / 0.25 / 0.30
Consolidatable CPU share; sleep ratios	s	60%; 0.05 / 0.08
Network and storage power	P_{ns}	0.5 MW
UPS rating; loss coefficients	S_r; c_0, c_1, c_2	12 MW; 0.008, 0.012, 0.010
PDU loss fraction; lighting	β_{PDU}; P_{light}	1.5%; 100 kW
Design heat load; fan power; min. flow	Q_d; $P_{fan,d}$; f_{min}	10.8 MW; 450 kW; 0.25
Recirculation δ_0, k_r, α_{ref}, n_r	—	0.8 K, 12 K, 1.30, 1.2
Hot-spot coeff.; sensor noise	k_g; σ_T	1.2 K; 0.3 K
Fan penalty k_f, T_0, n_f	—	0.005, 23 °C, 1.3
Carnot fraction; cond. approach	η_c; ΔT_{cond}	0.45; 12 K
Chiller unit capacity; pump power	Q_{ch}; $P_{pump,d}$	3.0 MW; 150 kW
Economizer approach; CHW ΔT	ΔT_{ec}; ΔT_{chw}	4 K; 6 K
Coil approach; inlet limit	ΔT_{coil}; T_{lim}	6 K; 27 °C

Table 2: Model Parameters

Proposed Ai Framework

Figure 2 shows the five-layer framework. Layer 1 acquires building management system (BMS), data center infrastructure management (DCIM), rack sensor, scheduler and weather data. Layer 2 cleans and stores the data and hosts the digital twin of Section III. Layer 3 forecasts load and utilization with an estimate of uncertainty. Layer 4 runs three coordinated agents: the shielded DDQN cooling controller, the consolidation module and the temporal-shifting module. Layer 5 writes set-points to the BMS, issues sleep/wake commands and reports indicators. The supervisory problem is

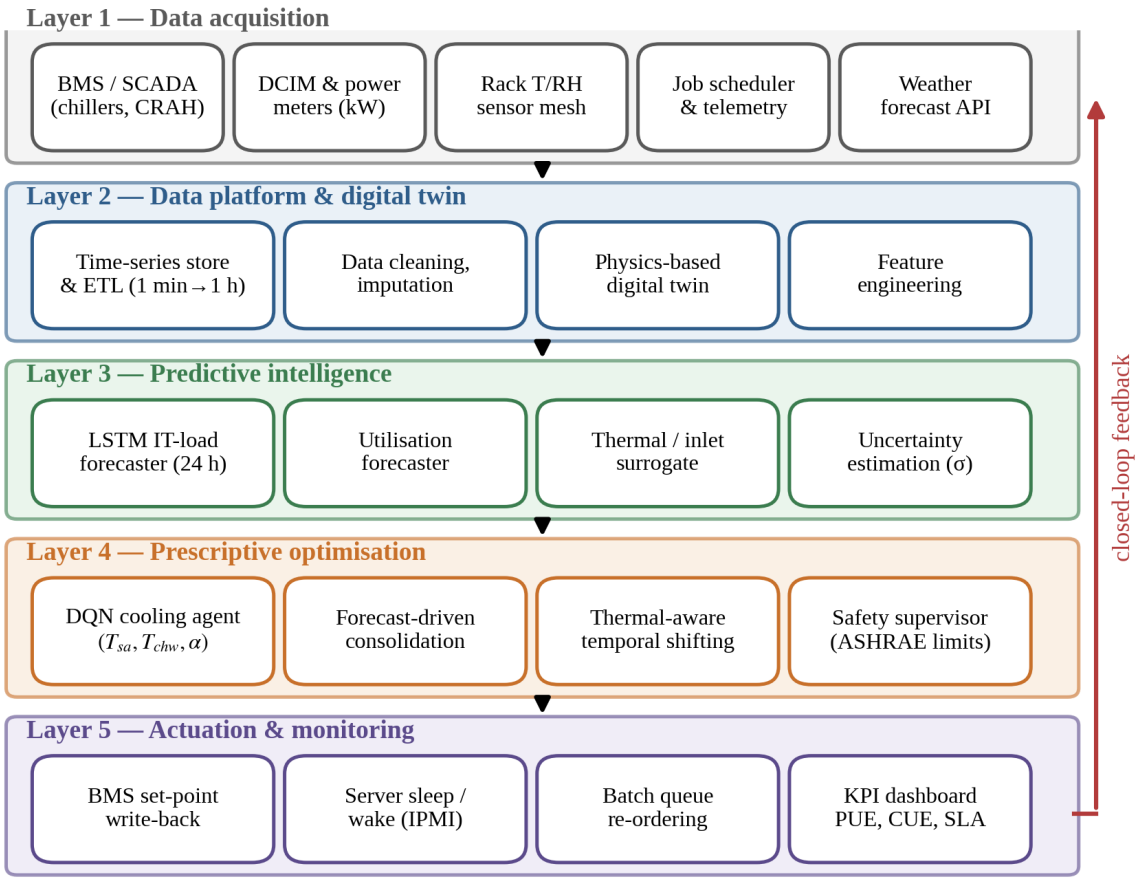


Figure 2: Five-Layer Architecture of the Proposed ai Framework for data Center Energy Optimization, from data Acquisition to Closed-Loop Actuation.

$$\min_{T_{sa},T_{chw},\alpha,a,u_f} \sum_{t=1}^{T} P_{fac}(t) \qquad (17)$$

subject to (1)–(16), capacity sufficiency and a same-day completion bound on flexible work. Because the problem is non-convex and stochastic, it is decomposed: the IT optimizer sets the active fraction a and flexible load u_f , and the DRL agent learns the cooling set-points and observes the resulting IT power.

LSTM Load Forecasting

A two-layer LSTM [30] with 64 units reads a 48-h window of nine features: IT power, CPU and inference utilization relative to their latest values, the latest values themselves, hour-of-day encodings and a weekend flag. With the gate vector $[\mathbf{f}_t; \mathbf{i}_t; \mathbf{o}_t] = \sigma(\mathbf{W}x_t + \mathbf{U}h_{t-1} + \mathbf{b})$, the cell updates

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (18)$$
$$h_t = o_t \odot \tanh(c_t) \quad (19)$$

A dense layer maps h_t to a direct 24-step forecast of three targets. The model predicts changes relative to the last observation, which makes it robust to workload growth, and it is trained with Adam [35] on the mean-squared error over N windows and H = 24 horizons, with early stopping:

$$\mathcal{L}(\Theta) = \frac{1}{NH}\sum_{n=1}^{N}\sum_{k=1}^{H}\|y_{n,k} - \hat{y}_{n,k}\|^2 \quad (20)$$

Shielded Double DQN for Cooling Control

1) MDP formulation: Cooling control is cast as a Markov decision process [36]. The 11-dimensional state contains the current and next-hour outdoor temperature, IT power, GPU share, cyclic encodings of hour and day, and the previous action. The action is a discrete triple

$$a_t = (T_{\text{sa}}, T_{\text{chw}}, \alpha) \in \mathcal{A} \quad (21)$$

with $T_{\text{sa}} \in \{18, 20, \ldots, 26\}$ °C, $T_{\text{chw}} \in \{7, 9, \ldots, 20\}$ °C and $\alpha \in \{1.00, \ldots, 1.40\}$. Combinations that violate (15) are removed, which leaves $|\mathcal{A}| = 190$ actions.

2) Reward: With the controllable overhead $P_{\text{oh}}(s, a) = P_{\text{fac}} - P_{\text{IT},0}$, the overhead relative to a reference action and the reward are

$$\Delta P_{\text{oh},t} = P_{\text{oh}}(s_t, a_t) - P_{\text{oh}}(s_t, a^{\text{ref}}) \quad (22)$$
$$r_t = -\frac{\Delta P_{\text{oh},t}}{P_{\text{n}}} - \lambda_1 v_t - \lambda_2 v_t^2 - \lambda_3 \| \Delta a_t \|_1 \quad (23)$$

where $v_t = \max(0, T_{\text{in}}^{\max} - T_{\text{lim}})$, P_{n} = 500 kW and Δa_t is the normalized set-point change. Because weather and load are exogenous, subtracting the overhead of the reference action a^{ref} = (22 °C, 12 °C, 1.20) acts as a state-dependent baseline: it removes the seasonal swing from the return without changing the ranking of actions. The server-fan term (10) is part of the overhead, so warmer set-points are not rewarded unconditionally.

3) Double DQN with shielding: A Q-network with two 256-unit layers [37] is trained with the double-DQN target [38], restricted to the safe action set:

$$y_t = r_t + \gamma\, Q\left(s_{t+1}, \arg\max_{a \in \mathcal{A}_{\text{safe}}} Q(s_{t+1}, a; \theta); \theta^-\right) \quad (24)$$

Following the shielding paradigm [26], the digital twin predicts the inlet temperature of every candidate action, and only actions with a margin m are admissible:

$$\mathcal{A}_{\text{safe}}(s) = \{a \in \mathcal{A} : \hat{T}_{\text{in}}^{\max}(s, a) \le T_{\text{lim}} - m\} \quad (25)$$
$$\pi(s) = \arg\max_{a \in \mathcal{A}_{\text{safe}}(s)} Q(s, a; \theta) \quad (26)$$

The margin m = 0.7 K keeps the exceedance probability below 1% at the boundary for σ_T = 0.3 K. The shield is active during exploration, target computation and deployment, so unsafe actions never reach the plant (Figure 3). Algorithm 1 summarizes training.

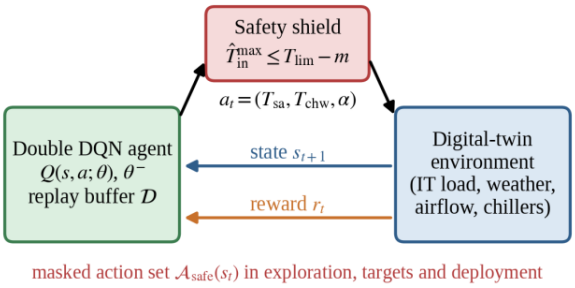


Figure 3: Interaction Between the Ddqn Agent, the Digital-Twin Safety Shield and the Environment.

Algorithm 1: Shielded double DQN training

1: Initialize θ, $\theta^{-} \leftarrow \theta$, replay buffer □; select six validation weeks
2: for episode = 1, …, 500 do
  3: Sample a 168-h window; alternate original/consolidated IT load
  4: for t = 1, …, 168 do
    5: Compute $□_{safe}(s_t)$ with the digital twin, Eq. (25)
    6: Choose ε-greedy a_t within $□_{safe}(s_t)$; apply a_t
    7: Observe r_t, s_{t+1}, $□_{safe}(s_{t+1})$; store in □
    8: Update θ on a 256-sample batch with target (24) and Huber loss
    9: $\theta^{-} \leftarrow \tau\theta + (1-\tau)\theta^{-}$; decay ε
  10: Every 20 episodes (after 100): validate; keep best θ
11: return shielded greedy policy (26)

Workload-Aware IT Optimization

Using the 1-h utilization forecast $\hat{u}(t)$ and the residual standard deviation σ_u, the active fraction of each consolidatable pool is

$$a(t) = \min\left[1, \max\left(a_{\min}, \frac{\hat{u}(t) + z\,\sigma_u}{u_{\text{tgt}}}\right)\right] \qquad (27)$$

with safety factor z and target utilization u_{tgt} = 0.8. A shortfall occurs if $u(t) > 0.98\ a(t)$. A share ψ = 20% of CPU demand is delay-tolerant within the day, and its daily total F_d is redistributed toward cooler forecast hours:

$$u'_f(h) = F_d \frac{w_h}{\sum_j w_j}, \qquad w_h = e^{-\beta(\hat{T}_{\text{out}}(h) - \bar{T}_d)} \qquad (28)$$

subject to a utilization cap of 0.90. The service impact is measured by the mean deferral, i.e., the earth-mover distance between the original and shifted flexible profiles, in hours.

Case Study And Simulation Setup

Climate, Workload and Data

The facility is located in Abu Dhabi (hot desert climate, Köppen BWh). Hourly temperatures are generated from monthly means (18.8 °C in January to 35.4 °C in August), diurnal amplitudes, a first-order autoregressive (AR(1)) synoptic component and noise. The test year has a mean of 27.7 °C, a peak of 44.6 °C and 4,496 h above 27 °C (Fig. 4(a)). CPU and inference utilization follow diurnal and weekly patterns with 5% annual growth and AR(1) noise, and the training cluster runs at about 82% occupancy with 14 random job gaps and maintenance windows per year. The mean IT power is 7.26 MW with a GPU share of 64% (Figure 4(b)). Two independent synthetic years are generated: Year A for training and validation and Year B for testing. All scenarios use the same Year B weather, workload and noise, so differences are attributable to control alone.

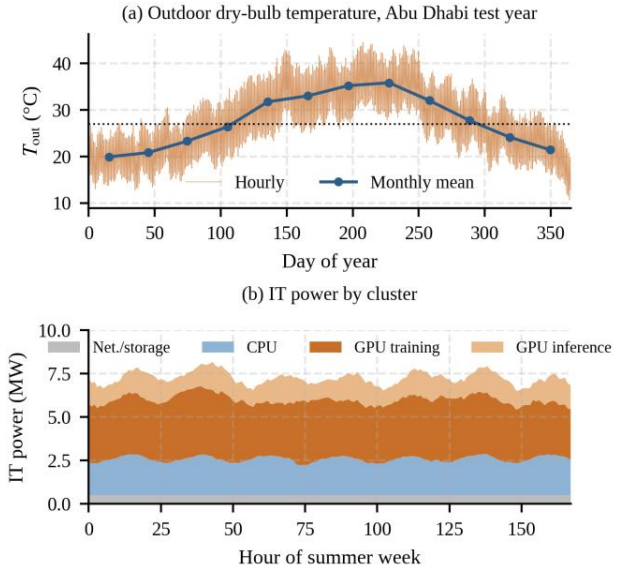


Figure 4: Case-Study inputs: (a) Hourly and Monthly-Mean Outdoor Temperature of the Abu Dhabi Test Year; (b) IT Power by Cluster for a Summer Week.

Scenarios and Implementation

Table III defines the scenarios. S0 reflects common legacy practice, S1 a competent static retrofit within the ASHRAE envelope, S2 AI cooling alone and S3 the full framework. The oracles OC and O search the 190-action set exhaustively every hour, with perfect knowledge of the nominal model and the same margin, and bound S2 and S3 respectively. The LSTM is benchmarked against seasonal naive, seasonal ARIMA with exogenous Fourier regressors (SARIMAX) [31]. Multilayer perceptron (MLP) and gradient-boosting [32] models. The DDQN is trained for 500 episodes (84,000

interactions) with the hyper-parameters of Table IV. The pipeline was implemented in Python with PyTorch, scikit-learn and statsmodels on a two-core CPU; training the DDQN took about 7 min.

ID	Cooling (T_{sa}, T_{chw}, α)	IT workload
S0	Static 18 °C, 7 °C, 1.35	All on; no shifting
S1	Static 22 °C, 12 °C, 1.20	All on; no shifting
S2	Shielded DDQN, hourly	All on; no shifting
S3	Shielded DDQN, hourly	Consolidation (z = 2) + shifting (ψ = 20%)
O_C	Exhaustive search	As S2
O	Exhaustive search	As S3

Table 4: Evaluated Scenarios

Model	Hyper-parameter	Value
LSTM	Window / horizon	48 h / 24 h
LSTM	Layers; dropout; batch	2 × 64; 0.1; 128
LSTM	Adam rate; weight decay	10^{-3} (×0.94/epoch); 10^{-5}
DDQN	State; actions; hidden	11; 190; 2 × 256
DDQN	γ; τ; batch; buffer	0.5; 0.005; 256; 1.2 × 10^{5}
DDQN	Adam rate; ε schedule	5×10^{-4}; 1 → 0.05 (60%)
DDQN	λ_1, λ_2, λ_3; P_n	4, 0.5, 0.03; 500 kW
Shield	Margin m	0.7 K
Workload	z; u_{tgt}; a_{min}; β	2; 0.8; 0.15; 0.25 K^{-1}

Table 5: Hyper-Parameters Of The Ai Models

Performance Metrics

Forecast accuracy is measured by the mean absolute error (MAE), root-mean-square error (RMSE), mean absolute percentage error (MAPE) and R^2. Energy performance is measured by the relative saving and the capture ratio of the attainable saving:

$$\Delta E_k = \frac{E_{fac}^{S0} - E_{fac}^{k}}{E_{fac}^{S0}}, \qquad \eta_{cap} = \frac{E_{fac}^{S0} - E_{fac}^{AI}}{E_{fac}^{S0} - E_{fac}^{O}} \tag{29}$$

Thermal compliance is measured by the hours with $T_{in}^{max} > 27$ °C and the 99th percentile of T_{in}^{max}. Economics use the simple payback $T_{pb} = C_{cap}/(\Delta C - C_{om})$ and the net present value (NPV) at 8% over five years, with an assumed tariff of USD 0.08/kWh, emission factor of 0.40 kg CO2/kWh, capital cost of USD 1.5 million and operating cost of USD 0.2 million per year.

Results And Performance Analysis

Load Forecasting

Table V compares the forecasters on the test year. At the 1-h horizon, SARIMAX (MAPE 1.16%) and the LSTM (1.18%) perform best, because strong autocorrelation makes the latest observation highly informative. At the day-ahead horizon, which temporal shifting requires, the LSTM is best (MAPE 3.04%), reducing RMSE by 6.5% relative to gradient boosting and by 24.3% relative to the seasonal naive model. The moderate day-ahead R^2 of all models reflects unpredictable job gaps in the training cluster. Fig. 5 shows the forecasts for a test week and the learning curves; the validation loss reaches its minimum at epoch 8.

Model	MAE 1 h	RMSE 1 h	MAPE 1 h (%)	R^2 1 h	MAE 24 h	RMSE 24 h	MAPE 24 h (%)	R^2 24 h
Seasonal naive	283	375	3.94	0.131	283	376	3.94	0.125
SARIMAX	83	119	1.16	0.912	—	—	—	—
MLP	89	127	1.24	0.900	237	307	3.30	0.415
Gradient boosting	90	126	1.25	0.902	234	304	3.26	0.427
LSTM (proposed)	85	121	1.18	0.909	218	284	3.04	0.498

Table 6: Forecast Accuracy for Total It Power on the Test Year (Mae And Rmse In Kw)

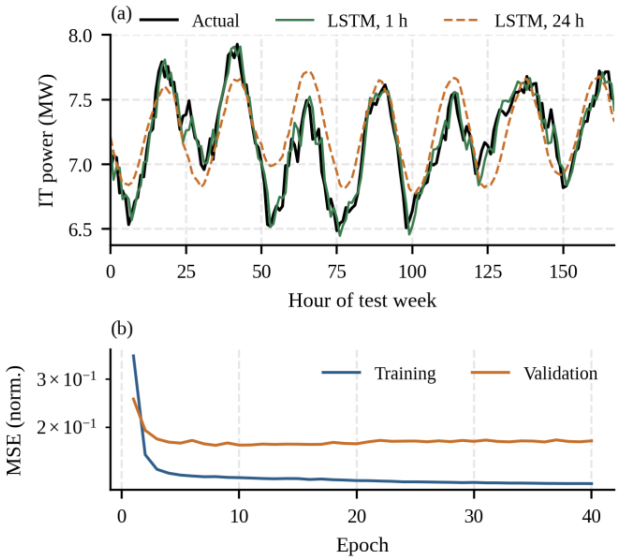


Figure 5: (a) Actual IT Power and LSTM Forecasts at 1-h and 24-h Horizons for a Test Week; (b) LSTM Training and Validation Losses.

DDQN Training Convergence

Because episodes come from different seasons, the training curve in Fig. 6 normalizes the agent's controllable overhead by that of the oracle on the same window. The gap falls from 21.3% in the first 20 episodes to 1.1% in the last 50, compared with about 42% for the static baseline. The validation energy on six held-out weeks varied by less than 0.05% after episode 120. With the shield active, only 125 of 84,000 training hours exceeded 27 °C, all through measurement noise; a preliminary unshielded agent produced several hundred violation hours per test year.

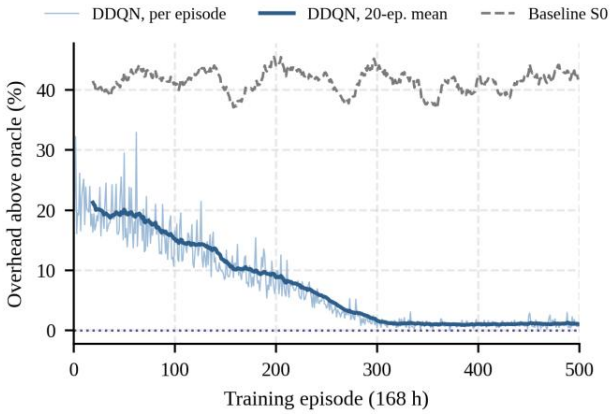


Figure 6: DDQN Training Convergence: Controllable Overhead Above the Oracle on Identical 168-h Windows.

Learned Control Behavior

Figure 7 shows the set-points chosen in S3. In summer the agent raises T_{sa} from 18 °C to about 22 °C and T_{chw} from 7 °C to about 16 °C, and it lowers α to about 1.01. In winter it lifts T_{chw} to 16.7 °C on average to use the economizer at night, which gives partial free cooling in 41% of winter-week hours. Over the year the economizer operates for 915 h (none in S0), and the mean COP rises from 4.04 to 5.88. Fig. 8 explains this behavior. At 38 °C the overhead falls steeply with T_{chw}, but the optimum lies on the boundary set by (15) and the inlet limit, not at the warmest T_{sa}, because of the recirculation and fan penalties. At 20 °C the optimum moves toward higher Tchw, where economization is possible, and a static set-point cannot follow this shift.

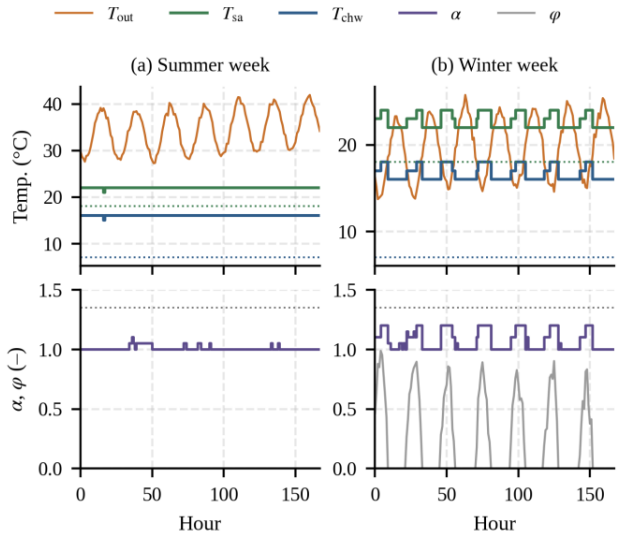


Figure 7: Set-Points Selected by the Shielded DDQN (S3) in (a) a Summer Week and (b) a Winter Week; Dotted Lines Mark the S0 set-Points.

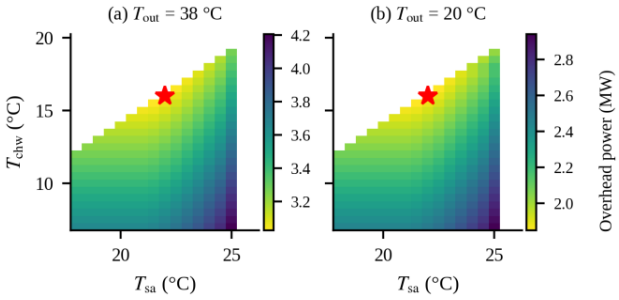


Figure 8: Minimum Overhead Power (over α) Versus T_{sa} and T_{chw} at an IT load of 7.8 MW for (a) 38 °C and (b) 20 °C Outdoor Temperature; Blank Regions are Infeasible and the Star Marks the Optimum.

Annual Energy Performance

Table VI and Fig. 9 report the annual energy by component. S0 consumes 90.40 GWh (PUE 1.421), of which cooling accounts for 25.4%. S1 saves 4.87%, S2 saves 8.57% and S3 saves 15.32% (13.85 GWh per year), with 40.8% less cooling energy, 6.8% less IT energy and a 14.1% lower peak. The waterfall in Fig. 10 attributes 32% of the total saving to correcting static over-cooling, 24% to adaptive DRL cooling and 44% to workload optimization. In this GPU-dominated facility, IT-side optimization therefore contributes the largest single share.

Scenario	IT	Elec. losses	Chillers	CRAH fans	Pumps + DC	Cooling	Facility	PUE	Peak (MW)	Saving (%)
S0 Baseline	63.63	2.95	18.51	3.55	0.89	22.95	90.40	1.421	12.67	—
S1 Rule-based	63.63	2.95	15.19	2.49	0.87	18.55	86.00	1.352	12.03	4.87
S2 AI cooling	63.87	2.95	12.43	1.61	0.91	14.95	82.65	1.294	11.58	8.57
S3 Full AI	59.31	2.78	11.47	1.28	0.84	13.58	76.55	1.291	10.88	15.32
O_C Cooling oracle	63.87	2.95	12.43	1.60	0.91	14.95	82.64	1.294	11.58	8.58
O Full oracle	59.32	2.78	11.46	1.26	0.84	13.56	76.54	1.290	10.88	15.33

Table 7: Annual Energy by Component (Gwh) and Performance Indicators on the Test Year

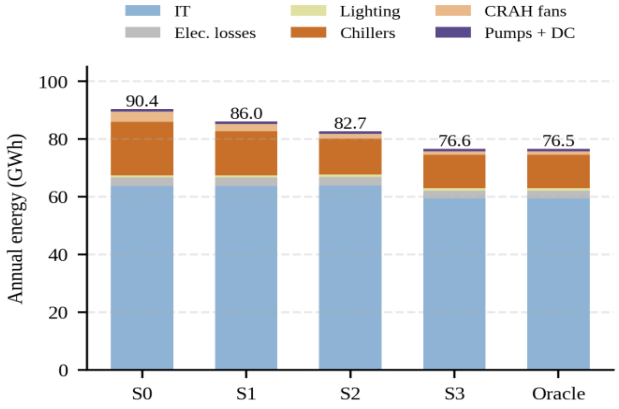


Figure 9: Annual Facility Energy by Component for Each Scenario.

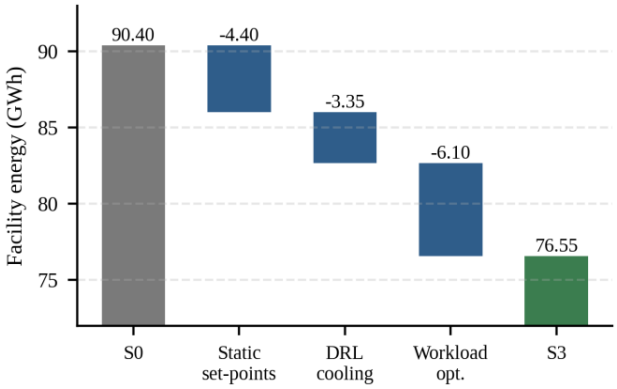


Figure 10: Decomposition of Annual Facility Energy Savings from the Baseline S0 To The Full Framework S3.

PUE Analysis

Figure 11 shows monthly PUE. S0 ranges from 1.351 in winter to 1.494 in August, and the AI scenarios lower PUE in every month. However, S3 uses 7.4% less energy than S2 while its PUE is almost unchanged (1.291 versus 1.294). Consolidation removes IT energy from the denominator of (3), while fixed overheads do not scale down, and warmer set-points shift 0.24 GWh into server fans in S2, where it counts as IT energy. PUE thus under-reports IT-side savings

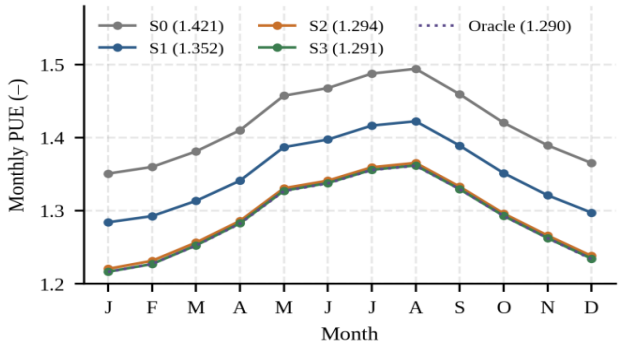


Figure 11: Monthly PUE for all Scenarios (Annual Values in the Legend).

Thermal Compliance and Service Quality

Figure 12 shows that S0 over-cools, with a mean T_{in}^{max} of 19.6 °C. The AI scenarios operate closer to the limit, while the 99th percentile remains 26.81 °C in S3. Table VII shows 10 exceedance hours in S3 (0.11% of the year), all small and noise-driven, with a maximum of 27.21 °C, far below the A1 allowable limit of 32 °C. On the IT side, consolidation at z = 2 kept 54.6% of the CPU pool and 65.9% of the inference pool active on average, with 6 shortfall hours, and flexible jobs were deferred by 5.3 h on average. Fig. 13 shows the safety-factor trade-off: z = 0 saves 8.3% of IT energy with 282 shortfall hours, whereas z = 2.5 eliminates shortfalls at a saving of 6.8%.

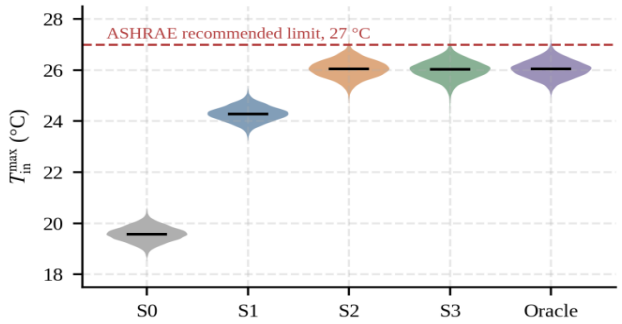


Figure 12: Distribution of Hourly Maximum Rack Inlet Temperature by Scenario, with the Ashrae Recommended Limit.

Indicator	S0	S1	S2	S3
Mean T_{in}^{max} (°C)	19.57	24.28	26.04	26.01
99th pct. T_{in}^{max} (°C)	20.25	24.98	26.82	26.81
Hours > 27 °C	0	0	16	10
Economizer hours	0	45	865	915
Fan penalty (MWh)	0	0	238	223
Shortfall hours	0	0	0	6
Mean deferral (h)	0	0	0	5.34

Table 8: Thermal and Service Indicators

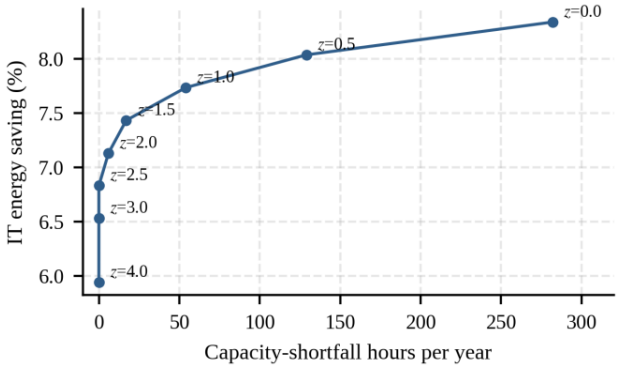


Figure 13: Trade-off Between it Energy Saving and Capacity-Shortfall Hours for Consolidation Safety Factors Z.

Optimality Gap and Climate Sensitivity

With (29), the cooling oracle O_C saves 8.58% and S2 saves 8.57% (capture ratio 99.9%). The full oracle saves 15.33% and S3 saves 15.32% (99.9%). The gap is small because the plant is quasi-steady at hourly resolution and the reference-subtracted reward gives a low-variance signal over a shield-reduced action set; model mismatch in a real facility would widen it. To test transferability, the temperature series was shifted by −12 K to +4 K (Fig. 14). The baseline PUE rises from 1.310 to 1.456, and the attainable saving falls from 16.2% in cooler climates, where higher T_{chw} unlocks more economizer hours, to 15.2% in hotter ones.

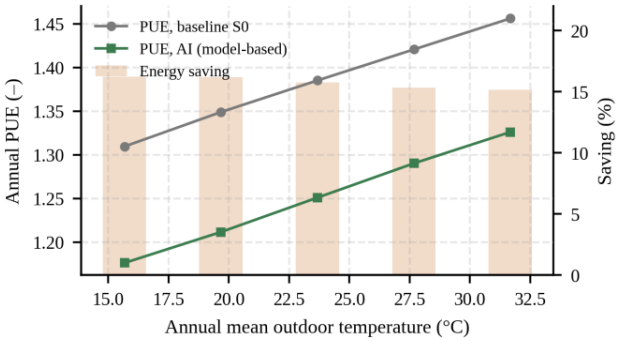


Figure 14: Sensitivity of Annual Pue and Attainable Facility Energy Saving to Site Mean Outdoor Temperature (Abu Dhabi: 27.7 °C).

Environmental, Economic and Comparative Assessment

Table VIII summarizes the environmental and economic results. S3 avoids 13,845 MWh and 5,538 t CO_2 per year and saves USD 1.11 million annually, which gives a payback of 1.65 years and a five-year NPV of USD 2.12 million. The 40.8% cooling-energy reduction is of the same order as the 40% reported by DeepMind [13]. And the 14–21% of offline RL in production [18]. And the 8.6% facility saving of AI cooling alone is comparable to the 9–13% of live RL trials [15]. Direct comparison requires care: the conservative S0 baseline inflates relative savings, whereas the hot climate limits them.

Indicator	S1	S2	S3
Energy saving (MWh/yr)	4,398	7,743	13,845
Avoided CO_2 (t/yr)	1,759	3,097	5,538
Cost saving (M USD/yr)	0.35	0.62	1.11
CUE (kg CO_2/kWh$_{IT}$)	0.541	0.518	0.516
Payback (years)	—	—	1.65
NPV, 5 yr at 8% (M USD)	—	—	2.12

Table 9: Environmental And Economic Performance Relative To S0

Discussion

Why AI outperforms static rules: S1 captures a substantial share of the benefit because S0 is severely over-cooled. S1 cannot, however, follow the weather-dependent optimum of T_{chw}, the load-dependent recirculation–fan trade-off or the change in hot-spot severity when consolidation raises the GPU share. It must also be sized for the worst case, whereas the shielded agent re-checks safety every hour and can approach the limit whenever conditions allow.

Deployment: A staged roadmap is recommended: instrument the facility and calibrate the digital twin (chiller curves, CRAH characteristics, rack inlet sensors); train offline and verify the shield on historical replay; run an advisory mode with operator approval, as in early industrial deployments [13], [15]; and then grant bounded, supervised autonomy with a watchdog that reverts to S1 set-points on stale data or model faults. Network segmentation, authenticated and rate-limited set-point write-back, and post-hoc explanations such as SHAP [39]. Support cyber-security and operator trust. The computational cost of the AI itself is negligible: training took minutes on a CPU, far below 0.1 kWh.

Limitations

The study relies on a calibrated simulation with quasi-steady hourly dynamics and a lumped recirculation model rather than computational fluid dynamics; weather and workloads are synthetic; humidity, dehumidification and water use are not modeled; the action space is discrete; and economic results depend on the stated assumptions. The near-perfect capture ratio should therefore be read as an upper estimate for the learning component. Validation on operational data, direct-to-chip liquid cooling of dense GPU racks [10] and the power dynamics of AI training [4] are important next steps.

Conclusion

This paper analyzed the performance of an AI framework for reducing the energy consumption of a 10 MW AI/cloud data center in a hot-arid climate. The framework combines LSTM forecasting, a DDQN shielded by a digital twin, forecast-driven consolidation and thermal-aware workload shifting. Over an independent test year, it reduced facility energy by 15.3%, cooling energy by 40.8% and peak demand by 14.1%, and improved PUE from 1.421 to 1.291, while keeping the 99th-percentile inlet temperature below the ASHRAE limit with 6 shortfall hours. The framework captured 99.9% of the oracle-attainable saving and paid back in 1.7 years under stated assumptions. Correcting over-cooling is the largest single cooling measure, but adaptive control and IT-side optimization together roughly triple the saving of static best practice. Safety can be built into DRL through shielding, and PUE should be complemented by absolute energy when AI-driven savings are reported. Future work will validate the framework on field data, extend it to liquid cooling and water use, and couple it with grid and carbon signals.

References

1. International Energy Agency, "Energy and AI," IEA, Paris, France, World Energy Outlook Special Report, Apr. 2025.
2. Masanet, E., Shehabi, A., Lei, N., Smith, S., & Koomey, J. (2020). Recalibrating global data center energy-use estimates. Science, 367(6481), 984-986.
3. Shehabi, A., Newkirk, A., Smith, S. J., Hubbard, A., Lei, N., Siddik, M. A. B., ... & Sartor, D. (2024). 2024 united states data center energy usage report.
4. Choukse, E., Warrier, B., Heath, S., Belmont, L., Zhao, A., Khan, H. A., ... & Alben, J. (2025). Power stabilization for AI training datacenters. arXiv preprint arXiv:2508.14318.
5. Information Technology—Data Centres—Key Performance Indicators—Part 2: Power Usage Effectiveness (PUE), ISO/IEC 30134-2:2016, 2016.
6. Avelar, V., Azevedo, D., French, A., & Power, E. N. (2012). PUE: a comprehensive examination of the metric. White paper, 49, 52.
7. Uptime Institute, "Uptime Institute global data center survey 2025," Uptime Institute Intelligence, New York, NY, USA, 2025.
8. Google, "Power usage effectiveness," Google Data Centers. [Online].
9. Zhang, Q., Meng, Z., Hong, X., Zhan, Y., Liu, J., Dong, J., ... & Deen, M. J. (2021). A survey on data center cooling systems: Technology, power consumption modeling and control strategy optimization. Journal of Systems Architecture, 119, 102253.
10. Al Kez, D., Foley, A. M., Wong, F. W. B. H., Dolfi, A., & Srinivasan, G. (2025). AI-driven cooling technologies for high-performance data centres: State-of-the-art review and future directions. Sustainable Energy Technologies and Assessments, 82, 104511.
11. ASHRAE Technical Committee 9.9, Thermal Guidelines for Data Processing Environments, 5th ed. Atlanta, GA, USA: ASHRAE, 2021.
12. Patel, C. D., Bash, C. E., Sharma, R., Beitelmal, M., & Friedrich, R. (2003, January). Smart cooling of data centers. In International Electronic Packaging Technical Conference and Exhibition (Vol. 36908, pp. 129-137).
13. R. Evans and J. Gao, "DeepMind AI reduces Google data centre cooling bill by 40%," Google DeepMind Blog, Jul. 2016. [Online].
14. Li, Y., Wen, Y., Tao, D., & Guan, K. (2019). Transforming cooling optimization for green data center via deep reinforcement learning. IEEE transactions on cybernetics, 50(5), 2002-2013.
15. Luo, J., Paduraru, C., Voicu, O., Chervonyi, Y., Munns, S., Li, J., ... & Mankowitz, D. J. (2022). Controlling commercial cooling systems using reinforcement learning. arXiv preprint arXiv:2211.07357.
16. Lazic, N., Boutilier, C., Lu, T., Wong, E., Roy, B., Ryu, M. K., & Imwalle, G. (2018). Data center cooling using model-

predictive control. Advances in Neural Information Processing Systems, 31.
17. Wang, R., Cao, Z., Zhou, X., Wen, Y., & Tan, R. (2024). Green data center cooling control via physics-guided safe reinforcement learning. ACM Transactions on Cyber-Physical Systems, 8(2), 1-26.
18. Zhan, X., Zhu, X., Cheng, P., Hu, X., He, Z., Geng, H., ... & Zhao, F. (2025, May). Data center cooling system optimization using offline reinforcement learning. In International Conference on Learning Representations (Vol. 2025, pp. 47735-47757).
19. Yuan, H., Bi, J., Li, S., Zhang, J., & Zhou, M. (2024). An improved LSTM-based prediction approach for resources and workload in large-scale data centers. IEEE Internet of Things Journal, 11(12), 22816-22829.
20. Radovanović, A., Koningstein, R., Schneider, I., Chen, B., Duarte, A., Roy, B., ... & Cirne, W. (2022). Carbon-aware computing for datacenters. IEEE Transactions on Power Systems, 38(2), 1270-1280.
21. Acun, B., Lee, B., Kazhamiaka, F., Maeng, K., Gupta, U., Chakkaravarthy, M., ... & Wu, C. J. (2023, January). Carbon explorer: A holistic framework for designing carbon aware datacenters. In Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2 (pp. 118-132).
22. Kahil, H., Sharma, S., Välisuo, P., & Elmusrati, M. (2025). Reinforcement learning for data center energy efficiency optimization: A systematic literature review and research roadmap. Applied Energy, 389, 125734.
23. Gao, J., & Jamidar, R. (2014). Machine learning applications for data center optimization. Google White Paper, 21, 1-13.
24. Rawlings, J. B., Mayne, D. Q., & Diehl, M. M. (2020). Model predictive control: theory, computation, and design.
25. Chen, Y., Guo, W., Liu, J., Shen, S., Lin, J., & Cui, D. (2024). A multi-setpoint cooling control approach for air-cooled data centers using the deep Q-network algorithm. Measurement and Control, 57(6), 782-793.
26. Alshiekh, M., Bloem, R., Ehlers, R., Könighofer, B., Niekum, S., & Topcu, U. (2018, April). Safe reinforcement learning via shielding. In Proceedings of the AAAI conference on artificial intelligence (Vol. 32, No. 1).
27. Garcıa, J., & Fernández, F. (2015). A comprehensive survey on safe reinforcement learning. Journal of Machine Learning Research, 16(1), 1437-1480.
28. Fan, X., Weber, W. D., & Barroso, L. A. (2007). Power provisioning for a warehouse-sized computer. ACM SIGARCH computer architecture news, 35(2), 13-23.
29. Beloglazov, A., Abawajy, J., & Buyya, R. (2012). Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future generation computer systems, 28(5), 755-768.
30. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.
31. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
32. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).
33. Moore, J. D., Chase, J. S., Ranganathan, P., & Sharma, R. K. (2005, April). Making Scheduling" Cool": Temperature-Aware Workload Placement in Data Centers. In USENIX ATC, General Track (pp. 61-75).
34. Ran, Y., Hu, H., Zhou, X., & Wen, Y. (2019, July). Deepee: Joint optimization of job scheduling and cooling control for data center energy efficiency using deep reinforcement learning. In 2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS) (pp. 645-655). IEEE.
35. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
36. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
37. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. nature, 518(7540), 529-533.
38. Van Hasselt, H., Guez, A., & Silver, D. (2016, March). Deep reinforcement learning with double q-learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 30, No. 1).
39. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.