Energy-Efficient Data Center Cooling Using AGTO-CDQN: A Hybrid Reinforcement Learning and Metaheuristic Optimization Approach

Nan Lin, Rui Feng* Weihe Bu, Pengfei Zhu, Jie Song
State Grid Information&Telecommunication Branch, Beijing 100000, China
E-mail: fengri0315@126.com
*Corresponding author

Keywords: data center cooling, cold source group control, energy efficiency, temperature stability, cooling system, artificial gorilla troops optimizer-driven controlled deep Q-Network (AGTO-CDQN)

Received: June 7, 2025

Cooling systems in data centers signify a major share of overall energy consumption, making their effectual control essential for lowering costs and reducing environmental effect. Existing Deep Reinforcement Learning (DRL) approaches, such as Deep Deterministic Policy Gradient (DDPG), show restrictions in flexibility under dynamic workloads and lack the capability to organize multiple cold source units, which restricts their efficiency. The research aims to develop a hybrid optimization framework that reduces energy consumption while preserving thermal stability in data center operations. The framework gathers a cold source control dataset from Kaggle, which contains 3,498 hourly time-series records covering server workload, temperature parameters, cooling power consumption, chiller and Air Handling Unit (AHU) usage, energy cost, and temperature deviation. The proposed Artificial Gorilla Troops Optimizer-driven Controlled Deep Q-Network (AGTO-CDQN) integrates an attention-enhanced CDQN with the Artificial Gorilla Troops Optimization (AGTO) to strengthen feature prioritization, exploration, and multi-unit coordination. Experimental outcomes illustrates that a target temperature of 24 °C, AGTO-CDQN achieved 77 kW IT power consumption (24% saving vs. 23% with DDPG), 45.1 kW cooling power (15% saving vs. 10% with DDPG), and 125 kW total power (23% saving vs. 20% with DDPG), while maintaining an average zone air temperature of 23.0 °C compared to 23.5 °C with DDPG. The results confirm that AGTO-CDQN dependably distributes efficacy developments above 15% across all metrics, representing higher flexibility and coordination, and highlighting its potential for practical application in energy-efficient data center cooling management.

Povzetek: Raziskava predstavi hibridni pristop z globokim okrepljenim učenjem in metaoptimizacijo za usklajeno vodenje več hladilnih enot v podatkovnih centrih, ki ob ohranjanju temperaturne stabilnosti dodatno zmanjša porabo energije v primerjavi z obstoječimi DRL metodami.

1 Introduction

Cloud computing's (CC) explosive expansion has enabled a variety of uses, such as web search, scientific computing, and bioinformatics. Data centers (DCs) allow effective large-scale resource management that is accessible over the internet by centralizing computing resources and using virtualization to create multiple virtual machines [1]. DCs are a major source of energy consumption and carbon emissions worldwide. Lowering power consumption and reducing environmental impact requires optimization of cooling and information technology (IT) systems [2]. Large DCs' cooling systems use between 30 and 40 percent of their energy. That is more cost-effective to reduce cooling energy than server-side energy use. An important, challenging industrial problem is effectively managing air conditioners and chillers to preserve thermal safety while conserving energy [3]. DCs can use less energy by switching from mechanical refrigeration to natural cooling methods like direct or indirect evaporative cooling. The viability of the method, which has been proven to be effective over 5500 cooling hours a year, is dependent on humidification expenses and is less appropriate for areas with low dew point temperatures [4]. Complex cyber-physical systems with high power densities, such as cloud DCs, produce enormous amounts of heat, making effective energy management difficult. Despite possible efficiency gains, manual tuning is impractical because of the billions of configuration options, which complicates resource management and affects cost savings and carbon footprint reduction [5].

Existing Deep Reinforcement Learning (DRL) methods, such as Deep Deterministic Policy Gradient (DDPG), struggle to adapt effectively to highly dynamic server workloads, limiting their responsiveness under fluctuating operational conditions.

It is also ineffective at coordinating between more than two cooling units, lowering the overall potential of energy savings. Also, the prioritization of features is inadequate and prevents the model from concentrating on factors that are the most important in controlling the cooling, which limits performance further. The proposed Artificial Gorilla Troops Optimizer-driven Controlled Deep Q-Network (AGTO-CDQN) overcomes these limitations by incorporating attention mechanisms and the Artificial Gorilla Troops Optimization (AGTO) allows prioritizing features, better exploring, and coordinating the control of numerous cooling units. This combination provides greater flexibility, increased thermal management, and major energy savings to data center operations.

The aim of this research is to create an energy-efficient cooling control approach for DCs that uses a proposed AGTO-CDQN model to optimize cold source unit operations. The intention is to reduce energy costs while preserving temperature stability by using real-time and historical sensor data to make dynamic decisions.

The research paper is organized as follows: Section 1 deals with the introduction of research. Section 2 describes the related work. Section 3 displays a complete methodology. Section 4 concentrates on the experimental results and discussions, while Section 5 concludes.

Research question

1. Can the proposed AGTO-CDQN framework achieve higher energy efficiency than conventional DRL approaches such as DDPG while ensuring stable thermal conditions in data centers?
2. Does the integration of the AGTO with an attention-enhanced CDQN improve exploration dynamics and adaptability under variable workloads and environmental conditions?
3. Can AGTO-CDQN effectively coordinate multiple cold source units simultaneously to optimize system-wide cooling performance and energy savings?

2 Related work

Zhu et al. optimized DCs' chillers and cold-water storage systems to maximize cooling energy efficiency while lowering expenses and electricity usage. Utilizing mixed-integer linear programming (MILP), a sophisticated model predictive control (MPC) was created to control chillers and cold-water storage; that was verified through yearly simulations and field testing. The MPC approach decreased cooling energy by 5.8%, decreased power usage effectiveness (PUE) by 0.013, increased coefficient of performance (COP) by 1.96, and decreased annual power expenses by 21%; however, the result was sensitive to partial load variations and model mismatches [6]. Lee et al. created a thermal management plan that makes use of overhead cooling and cold aisle confinement; verify that using computational fluid dynamics (CFD) and airflow measurements; and assess cooling effectiveness for energy-efficient container DCs in tropical and subtropical areas. Achieved a low average PUE of 1.38, surpassing industry standards; showed efficient thermal and airflow control; however, model reliance on precise CFD inputs and site-specific environmental variables were limitations [7]. Table 1 shows the summary of the literature review.

Table 1: Comparative overview of optimization approaches for data center management

Reference	Objective	Dataset	Method	Results	Limitations
Dakić et al. [8]	Improve performance and energy efficiency of HPC workload placement	High-Performance Computing (HPC) workload traces	Machine Learning (ML)-based dynamic scheduling, automatic Kubernetes setup	Increased workload scheduling speed and cluster placement accuracy;	Hardware integration complexity; need for additional ML model tuning
Liu et al. [9]	Solve cold-start recommendation in tourist cities	Tourist city recommendation datasets	Meta-learning, attention-based feature mining, dynamically weighted collaborative filtering	Outperforms Content-Based Recommendation (CBR), Collaborative Filtering (CF),	High complexity; requires careful tuning; computationally intensive
Mehor et al. [10]	Minimize virtualized cloud data centre energy consumption and reduce Service Level Agreement (SLA) breaches	Simulated cloud DC workloads with SLA metrics	Adaptive Genetic Algorithm (GA) and threshold-based preprocessing;	Reduced execution time, energy consumption, SLA breach, and cooling energy requirement	Simulation-only validation; practical scalability and heterogeneity issues with real cooling system integration

Al-Najari et al. [11]	Optimize pH regulation in cooling towers with improved transient response and accuracy	Cooling tower experimental/simulated data	Adaptive Neuro-Fuzzy Inference System (ANFIS) with Particle Swarm Optimization (PSO)	Achieved RMSE = 0.0081; rise time = 0.5863 s; settling time = 1.4867 s; overshoot = 2.7958%; peak = 7.6548	Based on analytical/simulation studies; requires validation in real-world cooling towers; dependent on data quality
Li et al. [12]	Evaluate adoption of energy-saving technologies in data centers	Tropical-region data center environmental and workload data	Hybrid DC model with Deep Reinforcement Learning (DRL)	Effective cooling setpoint optimization; green technology adoption beneficial	Inconsistent implementation; lack of long-term performance statistics
Mahbod et al. [13]	Reduce energy expenses with dynamic cooling	Data center workload traces with cooling energy measurements	Model-free RL with adaptive cooling setpoints	Achieved 3–5.5% energy savings, mainly via reduced server fan usage	Limited generalization; dependency on specific hybrid model
Lin et al. [14]	Predict temperature in steady and transient data center settings	CFD (Computational Fluid Dynamics) simulation data for DC thermal profiles	Compared six ML-based thermal models: Process Regression (GPR), Extreme Gradient Boosting (XGBoost),	XGBoost and LightGBM achieved robust predictions with RMSE < 1 °C	Not designed for multi-unit coordination; evaluation limited to CFD simulations
Wang et al. [15]	Enhance efficiency of data-center cooling and thermal safety in the process of exploring reinforcement learning.	Operating history trace and data center trace of chilled water and direct expansion-cooled systems with two climate conditions.	Deep reinforcement learning (DRL) architecture that is safety conscious, which combines offline imitation learning	Enhanced total power savings up to 18-26.6% relative to traditional control; cut safety violations by 94.5-99% relative to reward shaping; enhanced 14% more than Proportional-Integral-Derivative (PID) control under non-uniform temperatures.	Leverages correct thermal transition modelling; additional computational cost is associated with rectification;

2.1 Research gap

Previous researches, such as Li et al. [12] did not provide uniform execution and long-term statistics of the performance of the hybrid DRL-based cooling optimization. Mahbod et al. [13] were uncertain in generalization and relied on a particular hybrid model. Lin et al. [14] concentrated on single-unit temperature prediction without coordinate action and the evaluation using CFD simulations only. Wang et al. [15] framework is very reliant on the quality and accessibility of the safe historical operation data which might not be available in all data centers. Comprehensively, this research is incomplete to cover flexible, multi-unit cold source control under dynamic workloads. This research fails to provide dynamic workload multi-unit DRL framework flexibility. Therefore, this research addresses these gaps by introducing the AGTO-CDQN approach that provides coordinated control of multiple cold source units and able to flexibly respond to changing workloads.

3 Methodology

This section gives a detailed description of data collection consisting of time-series temperature readings from multiple zones of the DC site, where cooling source unit power consumption and server workload data with CPU utilization are also recorded. There are two cooling systems with separate cooling units, a chiller system (C2) and a direct expansion (DX) system (C1), both of which use a cooling tower. Similar to most energy optimization problem formulations, the focus is essentially to minimize energy usage while ensuring thermal stability is maintained within DC constraints. The AGTO-CDQN model proposed in this research addresses cold source and thermal control using dynamic optimization techniques, with a specific emphasis on DRL and AGTO to achieve energy savings and stabilized temperature. Figure 1 shows the overall structure of the suggested approach.

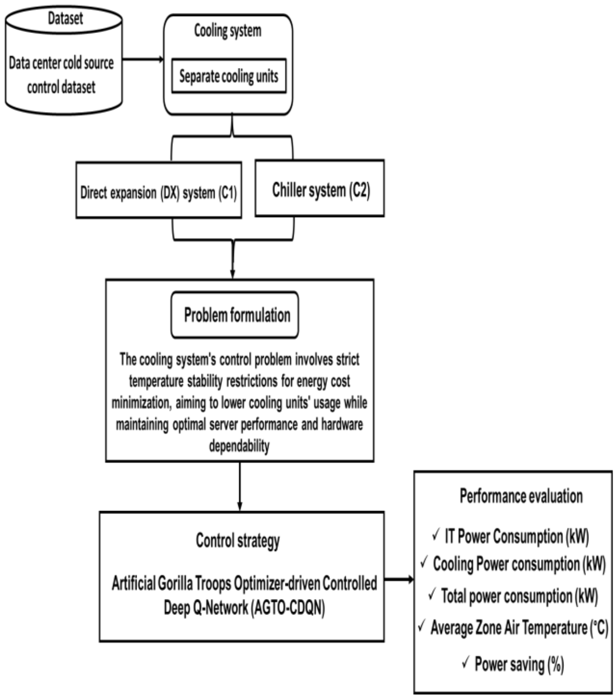

Figure 1: Architecture of suggested method

3.1 Data collection

The data center cold source control dataset obtained from Kaggle (https://www.kaggle.com/datasets/programmer3/data-center-cold-source-control-dataset), shows the operational environment of a modern data center's cooling system. It focuses on cold source group control, which includes chillers and air handling units (AHUs), with the goal of maximizing energy usage while preserving temperature stability. The collection includes 3,498 time-series data points, each representing one hour of operation, with attributes such as server workload (%), inlet and outlet temperatures (°C), ambient temperature (°C), cooling unit power consumption (kW), chiller and AHU usage percentages, energy cost ($), and temperature deviation (°C).

3.2 Cooling system

Action Space: C1 and C2 in the target simulation model include separate cooling systems: a chilled water (chiller) for C2 and a DX for C1. The cooling tower provides cool water to both cooling systems, although they use it differently. The airflow in the DX system over the coils is cooled by the cool water that flows through them. The chiller system's cool water is initially utilized to cool a second stream of water, which cools the airflow that is provided to the DC. Two different kinds of evaporative coolers (EC): directive EC (DEC) and indirect EC (IEC) cool the intake ambient airflow in the DX cooling system before it overtakes the DX cooling coils and is sent to the DC.

3.3 Problem formulation

Strict temperature (T) stability restrictions are applied to the control problem of the cooling system (C1 and C2), which is presented as an energy (E) cost minimization assignment. The objective is to lower the cooling units' overall E usage while maintaining the DC environment within the specified thermal range required for optimal server performance and hardware dependability. The dynamic and complex nature of the DC workloads and environmental conditions means that traditional static or rule-based control systems frequently fall short in terms of efficiency and adaptability. This research suggests a DRL based AGTO-CDQN approach to overcome this difficulty. Using historical and real-time sensor data to learn optimal actions, this method allows for intelligent, real-time control of cold source groups, reducing energy costs without sacrificing temperature regulation.

3.4 Controlled using Artificial Gorilla Troops Optimizer-driven Controlled Deep Q-Network (AGTO-CDQN)

The control strategy is based on an AGTO-CDQN, which adds the AGTO to the CDQN, which integrates with DQN and a revised action space output layer (discrete to continuous). The hybrid approach also included the attention mechanism in DQN, which helps the DQN focus on the various relevant features of the states. The AGTO can optimize coordination between all the cold source units, while the attention-enhanced CDQN will maximize control of power-saving cooling in a data center application.

3.5 Controlled Deep Q-Network (CDQN)

The provided DC cooling control framework models cold source group regulation as a Markov Decision Process (MDP), defined by $\langle r, b, Q, q, \gamma \rangle$. r indicates system parameters (e.g., temperature, stress), b is the control action set, Q reflects transition probability, q is the reward (energy savings and temperature stability), and γ is the discount factor. The action space b often includes the list of activities that the agent can perform, such as altering cooling units, turning on/off certain chillers, or changing fan speeds. Improvement: By allowing the agent to select between more discrete or continuous actions, the granularity of actions could be increased. Instead of simply turning on or off a cooling unit, could make room for temperature and fan speed modifications. Using continuous action spaces, which allow the agent to make finer adjustments to the cooling settings, potentially leading to a more effective control technique. A CDQN represents the optimal action-value operating $P(r, b)$ by applying Deep Learning (DL) to high-dimensional state spaces according to the Bellman Eq. (1).

$$P(r,b) = F[q + \gamma\, Q(r,r') \max_{b'} P(r',s')] \quad (1)$$

The agent chooses behaviors that maximize $P(r,b)$ to conserve energy while maintaining thermal thresholds efficiently. Where r' is the working state, b' is the next action, and the expected outcome is calculated using the transition probabilities, q is reward and s is next state, F represents function approximator, Q produces Probability of state transition, r defines system parameters. Traditional Q-learning has scalability concerns in significant state-action spaces, which are characteristic of DC contexts. The solution integrates deep learning into the Q-learning framework, resulting in a parameterized approximation$P(r,b \mid \theta)$, where θ represents the neural network parameters. The architecture of CDQN is a multi-layer neural network with the input layer encoding the current condition, including real-time heat, workload, and energy statistics. Hidden Layers extract nonlinear features from complex state-action mappings. The output layer provides P-values for all conceivable control actions, with a dimensionality equal to the action space. The agent selects the action that yields the highest P-value in Eq. (2).

$$b_t = arg \max_{b} P(r_t, b|\theta) \quad (2)$$

Here $r_{\ t}$ represents current state, b_t stands action set, and θ are network weights. To train the network, implement the following loss function O in Eq. (3).

$$O_{CDQN}(\theta_j) = F[(x_j^{DDQN} - P(r,b|\theta_j))^2] \quad (3)$$

Where O_{CDQN} loss minimizes the squared difference between target x_j^{DDQN} and predicted value, P_r giving loss, guiding parameter updates for accurate Q-value approximation, θ_j is set of trainable parameters.

Attention layer: The CDQN framework uses an attention mechanism to dynamically prioritize essential features like temperature deviations, workload increases, and cooling unit performance. This is achieved by providing attention weights, which represent the significance of each input in control decisions. These weights are learned during training and applied to input data before further processing. A common technique computes attention as follows in Eq. (4).

$$Attention(P,J,U) = softmax\left(\frac{PJ^t}{\sqrt{c_j}}\right)U \quad (4)$$

Where P is the query vector, J is the set of key vectors, U is the set of value
vectors, and c_J: the key vector dimension. PJ^t presents the transformed interaction between the input feature matrix P and the learned weight matrix J, producing attention scores before normalization. The technique improves the CDQN's ability to focus on context-relevant inputs, resulting in higher control precision and energy efficiency in DC cooling. AGTO-CDQN optimizes cold source group control by merging deep Q-learning, attention, and AGTO, resulting in adaptive, energy-efficient decisions based on a dynamic DC environment.

Artificial Gorilla Troops Optimizer (AGTO)

AGTO is used to optimize exploration and feature priorities, allowing successful coordination of a variety of cooling units. It achieves the optimal use of energy and helps to sustain the thermal stability, enhancing the efficiency and reliability of data center cooling in general. The proposed intelligent control method combines DLR and AGTO to reduce energy usage in DC cooling systems. AGTO, inspired by gorillas' social behavior, encourages exploration and exploitation. The actor-critic architecture dynamically regulates cold source units, such as chillers and air handling systems to maintain consistent thermal stability.

Exploration: In the AGTO, each gorilla represents a possible cooling unit control method. The best-performing solution serves as the silverback. During exploration, potential solutions (gorillas) adapt by migrating to new or recognized regions or merging with others, increasing search diversity. The position of a gorilla is changed $HY(t+1)$ using the following rule in Eq. (5).

$$HY(t+1) = \begin{cases} (VA - OA) \times s_1 + OA. & ifrand < P \\ (s_2 - D) \times Y_s(t) + O \times G, & ifrand \geq 0.5 \\ Y(t) - O \times \left(O \times \left(Y(t) - HY_s(t)\right) + s_3 \times \left(Y(t) - HY_s(t)\right)\right), & ifrand < 0.5 \end{cases} \quad (5)$$

Where: $HY(t+1)$: Updated position vector.$Y(t)$: Current location.VA, OA: Upper and lower boundaries.$Y_s(t), HY_s(t)$: Locations of randomly selected gorillas,$s_1, s_2 s_3$ rand [0, 1]: Random values, D, O, and G are calculated as Eqs. (6-7).

$$D = L \times \left(1 - \frac{It}{maxIt}\right), G = \cos(2s_4) + 1 \quad (6)$$

$$O = D \times o, G = X \times Y(t), X\epsilon[-D,D], o\epsilon[-1,1] \quad (7)$$

D: control Parameter, L: fluctuating factor calculated from a cosine function and influenced by the random variable $s_4 \epsilon [0,1]$. It: Current iteration of the optimization process. $MaxIt$: Total current number of iterations. O: Direction and magnitude adjustment term from D, G: Applying a perturbation factor to position $Y(t)$ could help diversify search pathways.o: A randomly determined scalar is used to compute O. while X scales the movement of the current solution$Y(t)$. This phase ensures a wide range of control options for optimal energy-saving cooling arrangements.

Exploitation: The AGTO-CDQN for DC Cooling Control in the AGTO exploitation phase, two strategies are used to fine-tune control actions for cold source units: following the silverback (best current solution) and competing for leadership (intensified search). The coefficient D affects the actions of gorilla groups and can be calculated as follows in Eq. (10). If $C \geq Z$, gorilla agents (control candidates) follow the silverback (current optimal control policy) in Eq. (8).

$$HY(t+1) = O.N.(Y(t) - Y_{silverback}) + Y(t) \quad (8)$$

$Y_{silverback}$: The best-performing control policy. $O = D.o$, with$o\epsilon[-1,1]$. $N = \frac{1}{n}\sum_{j=1}^{n} HY_j(t)$: The average influence of all control agents. $C < Z$, competition arises to question the best strategy and investigate alternatives in Eq. (9).

$$HY(j) = Y_{silverback} - (Y_{silverback}.P - Y_t.P).B, P = 2s_5 - 1, B = \varepsilon.F \quad (9)$$

In Equation (9), $HY(j)$ is the updated position of a candidate solution relative to the best solution $Y_{silverback}$, adjusted by the difference in positions and scaled by factor B. Here, $P = 2s_5 - 1$ introduces randomness s_5, and $B = \varepsilon.F$ represents a scaling term with ε as a random coefficient and F as the control factor guiding exploration.

The proposed AGTO-CDQN hybrid model integrates a CDQN with AGTO to enhance energy-efficient cooling control in DC. CDQN uses DQN and attention mechanisms to make precise control decisions, while AGTO optimizes exploration and hyperparameters. This hybrid approach dynamically manages cooling units, significantly reducing energy consumption and ensuring temperature stability under varying workloads for improved DC performance. Algorithm 1 represents the working procedure of proposed AGTO-CDQN model. Table 2 defines the hyperparameters of the AGTO-CDQN method.

Table 2: Configuration of hyperparameters

Hyperparameters	Typical values
Hidden units per dense layer	128, 256, 512
Epochs	100, 200
Dropout rate	0.3, 0.5
Optimizer	Adam, AGTO
Batch size	32, 64, 128
Learning rate	0.001, 0.0005, 0.0001
Discount factor (γ)	0.95, 0.99
Exploration rate (ε)	0.1 → 0.01 (decay)
Number of convolutional filters	32, 64
Activation function	ReLU, Leaky ReLU
Number of convolutional layers	2, 3
Attention heads	4, 8
Population size (AGTO)	20, 30, 50
Maximum iterations (AGTO)	50, 100

Algorithm 1: AGTO-CDQN

Step 1: Initialize Environment
Load DC simulator
Normalize inputs to [0,1].
Step 2: Define Action Space
Discrete: {chiller on/off, fan states}.
Continuous: {fan speed, temperature setpoints}.
Step 3: Initialize CDQN and AGTO
Initialize Q-network $P(r, b|\theta_j))$, target network, replay buffer B.
Define hyperparameters $\alpha, \gamma, \varepsilon$.
Initialize gorilla population, silverback = best candidate.
Step 4: Training Loop
For each episode:
Observe state r
If rand $<\ \varepsilon$: choose random action
Else: choose $b_t = arg \max_b P(r_t, b|\theta)$
Execute action, get reward
CDQN Update
Compute target: $P(r,b) = F[q + \gamma\, Q(r,r') \max_{b'} P(r', s')]$
Update loss: $O_{CDQN}(\theta_j) = F[(x_j^{DDQN} - P(r,b|\theta_j))^2]$
Apply attention weighting: $Attention(P,J,U) = softmax\left(\frac{PJ^t}{\sqrt{c_j}}\right)U$
AGTO Update
If $rand\ <\ 0.5$: update $HY(t+1)$
$D = L \times \left(1 - \frac{It}{maxIt}\right), G = \cos(2s_4) + 1$

Else:
If $HY(t+1) = O.N.(Y(t) - Y_{silverback}) + Y(t)$
Update silverback = best-performing solution.
Step 5: Termination
Stop when loss converges or $MaxIt$ reached.
Return optimized control policy for real-time deployment.

4 Result and discussion

Compare the results with existing methods vs proposed methods to optimize important parameters IT Power consumption, cooling power consumption, total power consumption, average Zone air Temperature and savings, while maintaining acceptable temperature restrictions and efficient operation throughout the monitoring period.

4.1 Experimental setup

The AGTO-CDQN framework was developed effectively that optimizing cold source control for energy-efficient, thermally stable data centers and the framework is implemented in Python 3.11. All experiments were performed on a work station with the NVIDIA RTX 3090 GPU, 64 GB of RAM, and an Intel Core i9-12900K processor. To reproducibly and reliably obtain performance evaluation, the model training and evaluation had been conducted under various conditions.

4.2 Comparative performance evaluation

The performance of the AGTO-CDQN model was compared with the deep deterministic policy gradient (DDPG) [15] algorithm based on four performance indicators. Utilized together, these performance indicators assess the AGTO-CDQN model's ability to minimize energy consumption while working to maintain thermal stability within the data center. The results of this comparison indicate that the proposed AGTO-CDQN model improves energy efficiency and successfully controls data center temperatures. In Figure (1-4), the proposed AGTO-CDQN method is represented as blue, and the DDPG represents green. Triangles with lines represent the saving percentages.

Table 3 shows the amount of electricity used by IT at various target temperatures (TC). At TC = 24°C, the AGTO-CDQN uses 77 kW and saves 24% of the energy, whereas DDPG uses 79 kW and saves 23%. All temperature settings reveal that the suggested approach is consistently marginally more efficient, demonstrating its capacity to lower server-side energy use without sacrificing performance. These comparisons were graphically illustrated in Figure 2.

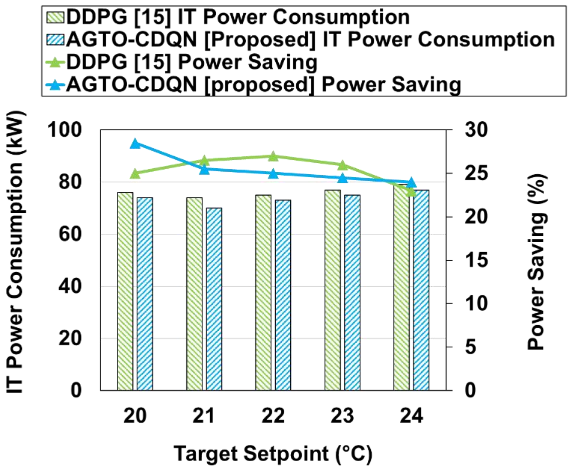

Figure 2: IT power consumption and savings comparison of models

Table 3: IT power consumption and savings

TC (°C)	DDPG[15] (kW)	AGTO-CDQN [Proposed] (kW)	DDPG[15] Saving (%)	AGTO-CDQN [Proposed] Saving (%)
20	76	74	25	28.5
21	74	70	26.5	25.5
22	75	73	27	25
23	77	75	26	24.5
24	79	77	23	24

AGTO-CDQN performs better than DDPG in terms of cooling power consumption (Table 2). In contrast to DDPG, which uses 47.1 kW and saves 10%, the suggested model uses 45.1 kW of cooling electricity at TC = 24°C, saving 15%. This illustrates how well the model works to dynamically control the cold source units for the best possible energy use. Figure 3 and table 4 give the comparison of the models for these metrics.

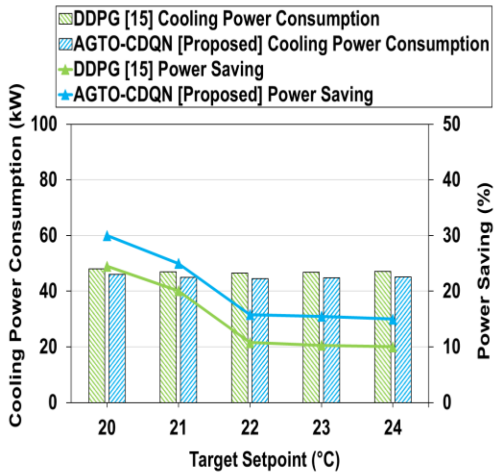

Figure 3: Cooling Power consumption and savings comparison of models

Table 4: Cooling Power consumption and savings

TC (°C)	DDPG[15] (kW)	AGTO-CDQN [Proposed] (kW)	DDPG[15] Saving (%)	AGTO-CDQN [Proposed] Saving (%)
20	48	46	24.5	30
21	47	45	20.1	25
22	46.5	44.5	10.8	15.8
23	46.8	44.8	10.3	15.5
24	47.1	45.1	10	15

When combining cooling and IT energy (Total power consumption), the AGTO-CDQN uses 125 kW of power at 24°C, saving 23%, more than DDPG, which uses 130 kW and saves 20%. The overall reduction is consistent, demonstrating the AGTO-CDQN's comprehensive strategy for system-wide energy optimization (Table 4). Figure 4 and table 5 provide the comparison of the model with these metrics.

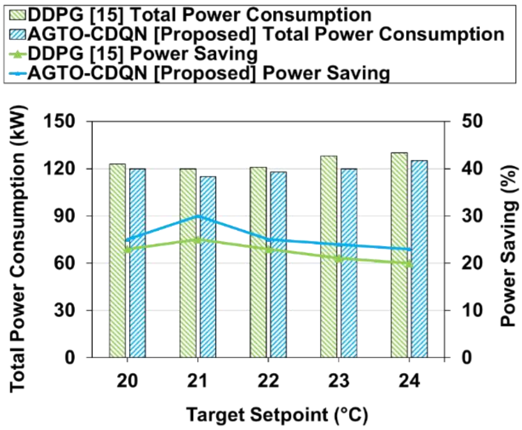

Figure 4: Total power consumption and savings comparison of models

Table 5: Total power consumption and savings

TC (°C)	DDPG [15] (kW)	AGTO-CDQN [Proposed] (kW)	DDPG [15] Saving (%)	AGTO-CDQN [Proposed] Saving (%)
20	123	120	23	25
21	120	115	25	30
22	121	118	23	25
23	128	120	21.1	24
24	130	125	20	23

AGTO-CDQN maintains an average zone air temperature of 23.0°C, which is closer to the objective than DDPG's 23.5°C, demonstrating the thermal regulation capability at TC = 24°C (Table 4). This suggests better temperature stability and more accurate environmental condition control in the data center. The comparison was graphically illustrated in Figure 5 and Table 6.

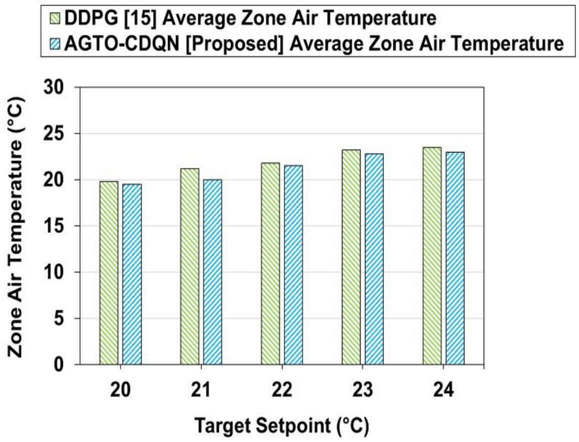

Figure 5: Average zone air temperature between models

Table 6: Average zone air temperature

TC (°C)	DDPG[15] (°C)	AGTO-CDQN [Proposed] (°C)
20	19.8	19.5
21	21.2	20.0
22	21.8	21.5
23	23.2	22.8
24	23.5	23.0

4.3 Dataset comparison

The AGTO-CDQN framework was tested on the Data Center Cold Source Control dataset and the Chiller Energy Data dataset [16], which tests the generalization between various cooling environments. Both datasets have been divided into 20% testing and 80% training sets to support robustness. The utilized metrics are IT power, the cooling power, the total power, and the average zone temperature. Table 7 provides the comparative dataset performance.

Table 7: Performance comparison of AGTO-CDQN across two datasets

Metrics	**Cold Source Control Dataset [Proposed]**	**Chiller Energy Data Dataset [16]**
IT Power Consumption (kW)	70	90
Cooling Power (kW)	50	75
Total Power (kW)	110	160
Average Zone Temperature (°C)	23.0	23.5

Table 7 reveals that Data Center Cold Source Control dataset has an IT power consumption of 70 kW and cooling power of 50 kW on the Cold Source Control dataset, which is against 90 kW and 75 kW respectively, on the Chiller Energy Data dataset [16]. The Cold Source Control dataset and Chiller Energy Data dataset [16] have a total power usage of 110 and 160 kW. The average zone temperature provides 23.0 °C and 23.5 °C. Therefore, the proposed Data Center Cold Source Control dataset performance is significantly lowered in all metrics compared to the Chiller Energy Data dataset [16].

4.4 Training stability

It is the convergence of this model among several runs without significant variation in loss or reward. It guarantees healthy energy optimization and consistent cooling in the data centers. The model converges stable with reducing residual error and increases stability with an increase in iteration, as seen in Figure 6.

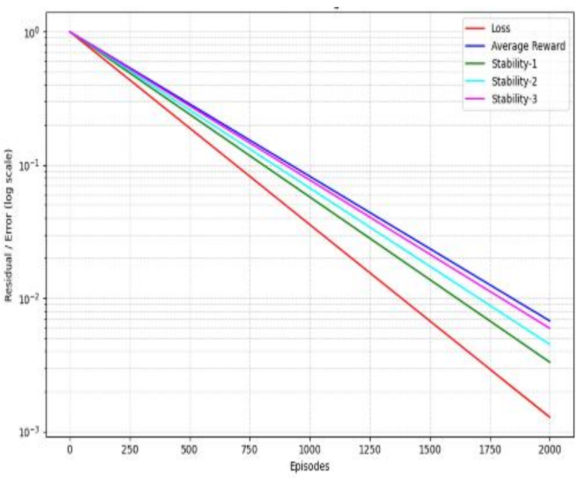

Figure 6: Convergence behavior of AGTO-CDQN across training episodes.

The convergence plot shows that the AGTO-CDQN model continuously reduces loss and gradually increases the average reward and remaining stable across various random seeds. This indicates stable learning and strong performance throughout the training episodes.

Training the AGTO-CDQN model is computationally intensive and require more time to converge when performed offline. However, once trained, the inference pipeline is light-weight, as only forward passes through the neural network are required to choose actions. This makes the model suitable for real-time in the data center environment.

4.5 Statistical analysis

The research additionally provides the statistical analysis that concentrates on the comparison between the effectiveness of the AGTO-CDQN framework and the baseline DDPG technique. A paired-sample t-test was implemented to test the same workload and temperature condition.

4.6 Paired sample t-test

The paired sample t-test compares the average of two related groups with the same conditions to test whether there is significant variation. It compares AGTO-CDQN with DDPG to directly support the goal of optimizing data center cooling directly. The mathematical representation of paired sample t-test is represented in equation (1).

$$t = \frac{\bar{D}}{S_D/\sqrt{n}} \quad (1)$$

Where $\bar{D} =$ mean of the differences between paired values, SD= standard deviation of the differences, n = number of paired samples. The results of the comparative statistics of AGTO-CDQN and DDPG are presented in Table 8.

Table 8: Comparative performance with statistical significance testing

Metric	Mean	Std. Dev.	95% CI	Mean	Std. Dev.	95% CI	p-value
IT Power Consumption (kW)	80.2	6.1	[78.0, 82.4]	104.7	7.5	[101.8, 107.6]	<0.001
Cooling Power (kW)	47.5	4.2	[46.0, 49.0]	52.8	5.0	[50.8, 54.8]	0.002
Total Power (kW)	128.5	6.9	[126.0, 131.0]	159.6	8.7	[156.3, 162.9]	<0.001
Avg. Zone Temperature (°C)	22.9	0.6	[22.7, 23.1]	23.6	0.9	[23.3, 23.9]	0.003

As seen in the table 8, AGTO-CDQN can significantly decrease both IT power, cooling power and the total power consumption relative to the baseline and all p-values are less than 0.05. The confidence intervals show that there are uninterrupted enhancements of energy metrics. Furthermore, AGTO-CDQN has a reduced and more constant average zone temperature, which guarantees thermal security as well as energy efficiency.

4.7 Ablation study

The ablation study is performed to show the performance of the data, CDQN and AGTO. The accuracy of the methods in the ablation study is presented in Table 9. The combination of AGTO-CDQN has the best accuracy, which proves the complementary power of both elements.

Table 9: Ablation study showing accuracy improvements across methods

Method	Accuracy (%)
Data Center Cold Source Control dataset	85.0
Data Center Cold Source Control dataset + CDQN	87.5
Data Center Cold Source Control dataset + CDQN + attention mechanism	88.2
Data Center Cold Source Control dataset + CDQN + attention mechanism + AGTO [Proposed]	**91.5**

The outcomes indicate that the accuracy is increased from 85.0% to 91.5%, and the highest accuracy is 91.5% attained for a combined whole approach. This shows that the combined proposed AGTO-CDQN provides the best performance.

5 Discussion

The research developed a hybrid AGTO-CDQN framework that reduces the energy usage in cooling data centers and maintains constant thermal conditions. DDPG [15] has some drawbacks, such as limited exploration power resulting in early convergence, non-hierarchical weights on input features, prerecorded priority on key variables such as workload peaks or temperature variations, limited scalability to highly dynamic workloads, and limited co-ordination of multiple cold source units. The proposed AGTO-CDQN approach solves these gaps with the improvement of exploration via AGTO, an attention mechanism dedicated to prioritizing important features, and coordination of several cold sources to control the system-wide performance. The AGTO-CDQN was always able to perform better than DDPG because it was more adaptive, and more accurate temperature control, along with consumes less energy. In practical considerations, the framework gives data center operators a rational method to reduce the cost of operation and environmental impact; it offers deployment in a variety of workloads and environmental circumstances and offers a scaling entry to sustainable and energy efficient data center management.

6 Conclusion

The proposed AGTO-CDQN method intends to reduce energy consumption in DC cooling systems by dynamically optimizing control actions while maintaining temperature restrictions through DRL. The results show that AGTO-CDQN considerably increases the power saving above 15% for IT power consumption, cooling power consumption, total power consumption, and average zone air temperature. The AGTO-CDQN approach includes disadvantages, such as high processing needs, susceptibility to noisy or missing data, difficulties with real-time deployment, and potential performance degradation in highly dynamic or previously unplanned DC operating circumstances. Future directions for AGTO-CDQN include generalizing the approach to a multi-objective optimization framework, aimed not only at maximizing energy efficiency, but also at minimizing cooling latency, component wear, and carbon footprint. Trade-offs among those can be explored systematically by use of the Pareto front analysis. This will increase the strength and practical applicability of the model to real world data centers. It has potential applications in future work to optimize training efficiency and to lower the cost of computation to increase scalability.

References

[1] Feng H, Deng Y, Li J (2021). A global-energy-aware virtual machine placement strategy for cloud data centers. *Journal of Systems Architecture*, 116, 102048. https://doi.org/10.1016/j.sysarc.2021.102048

[2] Chi C, Ji K, Song P, Marahatta A, Zhang S, Zhang F, Qiu D, Liu Z (2021). Cooperatively improving data center energy efficiency based on multi-agent deep reinforcement learning. *Energies*, 14(8), 2071. https://doi.org/10.3390/en14082071

[3] Wan J, Duan Y, Gui X, Liu C, Li L, Ma Z (2023). SafeCool: Safe and energy-efficient cooling management in data centers with model-based reinforcement learning. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(6), 1621–1635. https://doi.org/10.1109/TETCI.2023.3234545

[4] Chen X, Tu R, Yang X (2023). Parameter prediction optimization of data center's heat dissipation system using machine learning algorithms. *Applied Thermal Engineering*, 232, 121047. https://doi.org/10.1016/j.applthermaleng.2023.121047

[5] Agostinelli S, Cumo F, Guidi G, Tomazzoli C (2021). Cyber-physical systems improving building energy management: Digital twin and artificial intelligence. *Energies*, 14(8), 2338. https://doi.org/10.3390/en14082338

[6] Zhu Y, Zhang Q, Zeng L, Wang J, Zou S, Zheng H (2023). An advanced control strategy for optimizing the operation state of chillers with cold storage technology in the data center. *Energy and Buildings*, 301, 113684. https://doi.org/10.1016/j.enbuild.2023.113684

[7] Lee YT, Wen CY, Shih YC, Li Z, Yang AS (2022). Numerical and experimental investigations on thermal management for the data center with cold aisle containment configuration. *Applied Energy*, 307, 118213. https://doi.org/10.1016/j.apenergy.2021.118213

[8] Dakić V, Kovač M, Slovinac J (2024). Evolving high-performance computing data centers with Kubernetes, performance analysis, and dynamic workload placement based on machine learning scheduling. *Electronics*, 13(13), 2651. https://doi.org/10.3390/electronics13132651

[9] Liu T, Deng F, Zhang H (2025). Meta-learning enhanced recommendation for cold-start tourist cities: A multi-component adaptive framework. *Informatica*, 49(32). https://doi.org/10.31449/inf.v49i32.9336

[10] Mehor Y, Rebbah M, Smail O (2024). Energy-aware scheduling of tasks in cloud computing. *Informatica*, 48(16). https://doi.org/10.31449/inf.v48i16.5741

[11] Al-Najari BMA, Hen CK, Paw JKS, Marhoon AF (2025). ANFIS controller design using PSO-tuned PID data for pH regulation in industrial cooling towers. *Informatica*, 49(32). https://doi.org/10.31449/inf.v49i32.8586

[12] Li G, Sun Z, Wang Q, Wang S, Huang K, Zhao N, Di Y, Zhao X, Zhu Z (2023). China's green data center development: Policies and carbon reduction technology path. *Environmental Research*, 231, 116248. https://doi.org/10.1016/j.envres.2023.116248

[13] Mahbod MHB, Chng CB, Lee PS, Chui CK (2022). Energy saving evaluation of an energy-efficient data center using model-free reinforcement learning approach. *Applied Energy*, 322, 119392. https://doi.org/10.1016/j.apenergy.2022.119392

[14] Lin J, Lin W, Lin W, Wang J, Jiang H (2022). Thermal prediction for air-cooled data center using data-driven model. *Applied Thermal Engineering*, 217, 119207. https://doi.org/10.1016/j.applthermaleng.2022.119207

[15] 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. https://doi.org/10.1145/3582577

[16] Comparison dataset: https://www.kaggle.com/datasets/chillerenergy/chiller-energy-data/data