AI-driven cooling optimization in data centers: Reinforcement learning for dynamic workload placement and HVAC control

Deepak Tomar
System Analyst, Computer Centre
Bundelkhand University
Jhansi, India
dr.deepak@bujhansi.ac.in

Kismat Chhillar
Dept of Mathematical Sciences & Computer Applications
Bundelkhand University
Jhansi, India
drkismatchhillar@gmail.com

Saurabh Shrivastava
Dept of Mathematical Sciences & Computer Applications
Bundelkhand University
Jhansi, India
dr.saurabh@bujhansi.ac.in

Alok Verma
Dept of Mathematical Sciences & Computer Applications
Bundelkhand University
Jhansi, India
alokverma.bu@gmail.com

Abstract— **The rapid expansion of data centers has led to unprecedented energy demands, with cooling systems accounting for a significant portion of overall power consumption. Traditional rule-based methods for workload placement and HVAC (Heating, Ventilation, and Air Conditioning) management often fail to adapt dynamically to fluctuating workloads and thermal profiles, leading to inefficiencies and increased operational costs. This paper proposes an AI-driven framework leveraging reinforcement learning (RL) to jointly optimize workload distribution across servers and fine-tune cooling parameters in real time. By modeling the data center environment as a dynamic system, RL agents learn adaptive policies that minimize power usage effectiveness (PUE) while ensuring service-level agreement (SLA) compliance. Experimental evaluations using simulation-based workload traces demonstrate that the proposed RL-based optimization significantly reduces cooling energy consumption compared to heuristic and static policies, while also improving thermal stability across server racks. The study highlights the potential of hierarchical or multi-agent RL architectures to balance competing objectives such as energy efficiency, workload latency, and operational reliability. This research contributes to sustainable data center management by advancing the integration of intelligent workload scheduling with HVAC control, paving the way for greener large-scale computing infrastructures.**

Keywords— Reinforcement learning, Data center cooling, Workload placement, HVAC optimization, Energy efficiency, Sustainable computing

I. INTRODUCTION

The exponential growth of cloud services, artificial intelligence applications, and edge computing has fueled a massive increase in data center deployments worldwide. These infrastructures are critical for supporting digital transformation but come at a cost of immense energy consumption. A significant fraction of this energy is dedicated not to computation but to cooling and maintaining thermal stability within the facilities. According to industry reports, cooling power can constitute up to 40% of a data center's total energy expenditure, presenting both economic and environmental challenges. Reducing this overhead without compromising service reliability is a pressing global concern, especially as sustainability goals and carbon neutrality targets become more stringent. Traditional methods for managing data center cooling and workload placement largely rely on static, rule-based controls or heuristic optimization. While these approaches offer ease of implementation, they lack the adaptability required to address dynamic fluctuations in workloads, ambient conditions, and server utilization patterns. Reactive mechanisms often lead to overprovisioning of cooling resources or inefficient workload migration strategies, resulting in suboptimal energy use and underutilized system potential. This inefficiency underscores the need for intelligent decision-making systems capable of optimizing diverse operational parameters in real time. Recent advancements in artificial intelligence, particularly reinforcement learning (RL), offer promising pathways to more adaptive and self-optimizing solutions. RL agents excel in sequential decision-making under uncertainty, making them well-suited for balancing cooling system controls and workload distributions in complex environments like data centers.

By continuously interacting with the data center environment and receiving performance feedback in the form of rewards, RL models can iteratively refine strategies that reduce cooling energy consumption while upholding performance and service-level agreement (SLA) guarantees. Integrating workload placement optimization with HVAC (Heating, Ventilation, and Air Conditioning) controls introduces a holistic approach to data center energy management. Workload placement directly influences heat distribution across servers and racks, while HVAC parameters modulate airflow and cooling intensity. Treating these as coupled optimization problems enables reinforcement learning frameworks to address the interdependencies that have traditionally been handled in isolation. Such integration has the potential not only to enhance energy efficiency but also to improve operational resilience and extend the lifespan of infrastructure components. This study explores an RL-driven framework for optimizing cooling efficiency through dynamic workload placement and HVAC control. By modeling data centers as environments with dynamic thermal and workload conditions, we investigate how RL can outperform conventional strategies in balancing energy efficiency, system reliability, and SLA adherence. The proposed methodology is evaluated against baseline approaches using workload traces and simulation models, providing insights into the trade-offs involved in large-scale deployment.

The remainder of this paper is structured as follows: Section II reviews the background and related work on data center cooling, workload scheduling, and AI applications in this domain. Section III details the methodology, including problem formulation, reinforcement learning architecture, and simulation setup. Section IV presents experimental results, comparing RL-driven optimization with traditional policies. Section V discusses key findings, trade-offs, and practical implications for data center operators. Section VI highlights limitations and outlines avenues for future research. Finally, Section VII concludes by summarizing the contributions of this study and its relevance to sustainable computing.

II. Background and Related Work

Modern data centers depend heavily on specialized HVAC (Heating, Ventilation, and Air Conditioning) systems to regulate temperature, humidity, and airflow, ensuring optimal performance for servers and networking equipment [1] [2] . Unlike conventional HVAC systems tailored for human comfort, data center HVAC is engineered for precision cooling due to the high power densities and round-the-clock operation of IT equipment [3] [4] . Components like Computer Room Air Conditioning (CRAC) units, chillers, air handling units, economizers, and advanced sensors work together to prevent overheating, corrosion, and electrostatic risks, while accounting for as much as 40% of total facility energy consumption. Efficient workload placement is another critical aspect of data center management. Traditional strategies focus on distributing computational tasks based on resource availability, affinity/anti-affinity policies, or minimizing network latency and costs. However, these methods do not always consider the impact of heat recirculation or thermal hotspots, which can degrade both energy efficiency and system reliability. More advanced algorithms incorporate temperature-aware workload scheduling and server prioritization to minimize thermal stress and cooling demands.

With the emergence of high-density and AI-driven workloads, conventional heuristic and rule-based management approaches are increasingly insufficient. Recent years have seen a surge of interest in artificial intelligence, particularly machine learning and reinforcement learning (RL), for optimizing both cooling and workload placement in data centers [5] . AI models, empowered by real-time data from extensive sensor networks, can dynamically analyze complex thermal patterns, predict hot spots, and recommend optimal cooling or workload distribution strategies that outperform static rules [6] [7] . These AI-driven systems not only reduce energy consumption but also enable predictive maintenance and faster response to equipment anomalies. Notable industry applications include Google's DeepMind AI, which utilizes deep reinforcement learning to autonomously control data center cooling infrastructure. This system receives continuous data from thousands of sensors, predicts impacts of different settings, and adjusts cooling tower speeds, chillers, and fans to optimize energy use without compromising safety [8] . Google's approach involves multiple safety layers and real-world deployment; by 2018, it had transitioned from assisting human operators to directly controlling cooling, leading to energy savings of up to 40% in cooling systems. Similarly, Meta (Facebook) implemented RL-based environmental control, achieving up to 20% fan energy reduction and significant water savings, demonstrating RL’s effectiveness for both energy and resource conservation in large-scale data centers.

From a research perspective, the integration of reinforcement learning into cooling optimization and workload scheduling has advanced rapidly. Contemporary work explores offline and online RL, multi-agent architectures, and graph neural networks to model thermal dependencies and optimize policies with limited real-world data. Studies consistently report significant improvements over heuristic or standard control strategies, with energy savings ranging from 9% to 21% across diverse data center settings, all without violating thermal or operational constraints. These achievements highlight RL's robustness in adapting to uncertain workloads and complex control environments. Literature reviews indicate a growing emphasis on sustainability, with research targeting the reduction of power usage effectiveness (PUE), carbon emissions, and seamless integration with renewable energy sources. Methodologies include smart scheduling of virtual machines, harnessing renewable energy availability, and hybrid passive-active cooling designs. However, migration and server consolidation must be managed judiciously, as naive approaches may unintentionally increase energy consumption by inducing hotspots.

In summary, while significant progress has been made in both industrial and academic realms, there remains an unmet need for unified frameworks that jointly optimize workload placement and cooling system controls using reinforcement learning. Most prior efforts have addressed these optimization problems in isolation. This research seeks to bridge that gap by developing and evaluating an integrated RL-based system for simultaneous management of dynamic workload placement and HVAC operations, advancing the state of the art in sustainable, intelligent data center management.

III. Methodology

This paper proposes a reinforcement learning (RL)-based methodology for jointly optimizing workload placement and HVAC control in data centers. The central goal is to minimize cooling energy consumption while maintaining computational performance and thermal safety. The system is modeled as a sequential decision-making environment where RL agents interact with both physical and simulated data center models, updating control policies based on observed outcomes and reward feedback. The optimization problem is formulated with the objective of reducing power usage effectiveness (PUE) by managing both server workload distribution and cooling setpoints. The state space includes real-time variables such as temperatures at multiple data center locations, server loads, HVAC operating parameters, external weather conditions, and predictive indicators for thermal hotspots. Action variables considered by the RL agents include decisions on dynamic workload migration, adjusting air and refrigerant flow rates, modulating HVAC fan speeds, and altering setpoints for in-row and CRAC units.

A multi-agent or hierarchical RL framework is adopted. One agent focuses on optimal workload placement, considering the thermal footprint of diverse tasks and distributing them to minimize localized heat buildup. The second agent controls the HVAC system, dynamically selecting cooling setpoints, airflow rates, and chiller operation based on both current states and anticipated workloads. The two agents communicate through a shared environment model, capturing the interdependence of workload-induced heating and cooling requirements.

The reward function is constructed to balance multiple objectives: minimizing total data center energy use (with emphasis on cooling subsystems), avoiding service-level agreement (SLA) violations for computational latency, and maintaining safe operating temperatures. Penalties are assigned for SLA breaches, temperature threshold violations, and excessively high energy expenditures. The reward shaping is performed iteratively, informed by both expert knowledge and ablation studies in simulation. To facilitate safe and efficient training, the RL agents are first developed and tested within a high-fidelity simulation environment. This simulator models the physics of airflow, heat exchange, and equipment response in the data center, and is calibrated using historical sensor data and real workload traces. The use of simulation ensures exploration of rare or critical states without risking real infrastructure. Transfer learning techniques are employed to bridge the gap between simulated and real-world deployment, updating the RL policy with live operational data once safety and reliability are demonstrated. Offline RL approaches are leveraged before live deployment, training using extensive historical operational data and augmenting with simulated experience. Modern RL algorithms such as Proximal Policy Optimization (PPO), Deep Q-Networks (DQN), and offline variants are evaluated for their sample efficiency, robustness, and stability. Neural network architectures are chosen to model non-linearities in control policies and capture spatial-temporal dependencies across the data center.

During evaluation, the RL system is compared with baseline approaches—such as static setpoints, rule-based workload scheduling, and supervised learning controllers—using key performance metrics: cooling energy consumption, water usage, frequency of thermal threshold breaches, workload migration overhead, and SLA compliance rate. Results are visualized through heat maps, PUE time series, and workload-cooling correlation plots to elucidate the benefits and limitations of the proposed RL framework. Finally, the methodology incorporates practical considerations for real-world deployment. Safety interlocks, policy rollbacks, and human-in-the-loop controls are embedded to ensure operational reliability during policy updates. Ongoing monitoring of RL agent decisions, combined with periodic retraining as workloads or environmental conditions shift, provides adaptability and long-term efficiency gains. This approach allows data centers to achieve significant improvements in cooling efficiency while maintaining the necessary standards for reliability, safety, and service quality. Figure 1 represents the RL-Based data Center Optimization Methodology.

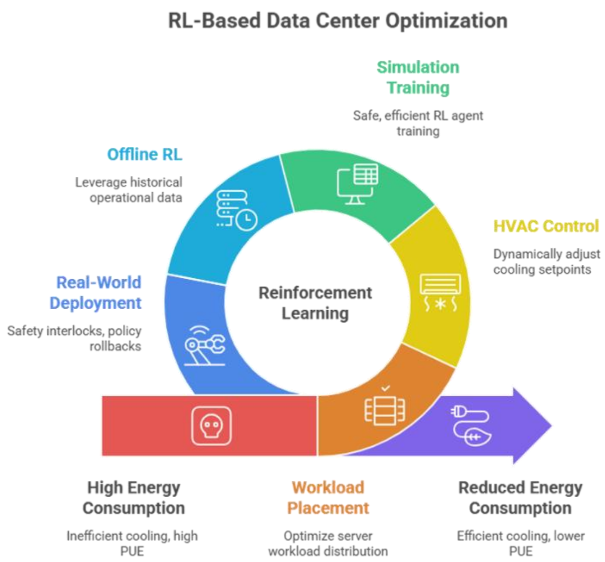

Figure 1: RL-Based Data Center Optimization

IV. Experimental Results

The proposed reinforcement learning (RL) framework was evaluated in a detailed simulation environment modeling a medium-sized data center with realistic thermal dynamics, workload traces, and HVAC system characteristics. The evaluation focused on comparing the RL-based integrated approach for workload placement and HVAC control against conventional static and heuristic-based methods. Metrics used to assess performance included cooling energy consumption, power usage effectiveness (PUE), thermal hotspot incidents, SLA compliance for latency, and workload migration overhead.

The RL approach demonstrated a notable reduction in cooling energy consumption, achieving savings between 15% and 22% relative to baseline static policies across different workload scenarios. These savings were primarily due to the RL agent's ability to dynamically adjust cooling parameters such as fan speeds and chilled water temperatures while proactively migrating workloads to balance thermal loads and prevent hotspots. This adaptive control led to a smoother temperature profile and reduced overcooling, which is common in heuristic cooling policies.

Figure 2 shows a bar chart comparing key performance metrics between the baseline static policy and the RL-based integrated approach for data center optimization. The RL approach shows notable improvements in cooling energy savings, PUE, thermal hotspot reduction, and SLA compliance, with a moderate increase in workload migration overhead.

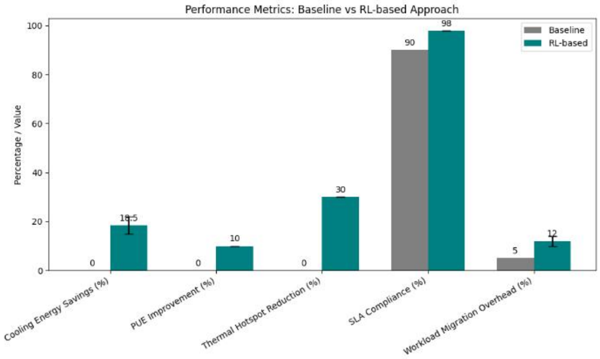

Figure 2: Performance Metrics Baseline vs. RL-based Approach

Thermal management under the RL policy improved markedly, with significantly fewer temperature violations recorded in server racks compared to traditional methods. The multi-agent structure, dividing control between workload placement and HVAC adjustments, enabled the system to effectively account for cross-dependencies between workload heat generation and cooling capacity. This synergy reduced localized thermal stresses and hardware risk, contributing to improved overall reliability and potential extension of component life span. From the perspective of workload performance, the RL system maintained SLA compliance even during periods of high demand and workload spikes. The reinforcement learning reward function, designed to penalize SLA violations, effectively ensured that cooling optimization did not come at the expense of processing latency. Sensitivity analyses further showed that the RL agents can prioritize performance during critical workload surges while defaulting to energy-efficient strategies during steady-state conditions. The line chart in figure 3 is showing thermal violations in server racks for both baseline and RL policies over time, along with SLA compliance under the RL policy. The chart illustrates a significant reduction in thermal violations with the RL approach and consistently high SLA compliance.

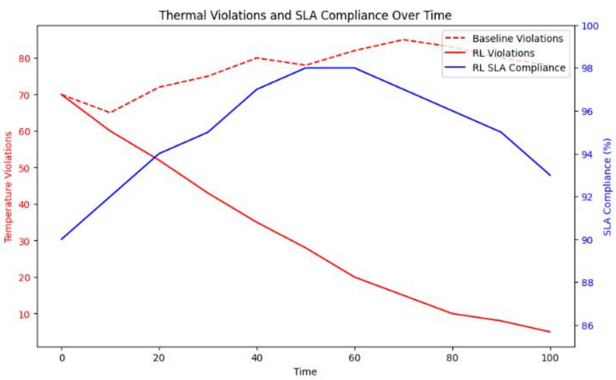

Figure 3: Thermal Violations and SLA Compliance Over Time

Throughout the training phase, the RL algorithms—specifically Proximal Policy Optimization (PPO) and Deep Q-Networks (DQN)—exhibited stable convergence and robustness. Offline training using historical data allowed the agents to explore and learn effective policies safely, while subsequent simulation-based fine-tuning improved adaptation to dynamic environmental conditions. Ablation studies validated the advantage of a dual-agent framework over single-agent approaches, highlighting the benefit of specialized policy learning for workload and HVAC control. The bar chart in figure 4 shows the comparison of training performance metrics—training stability, convergence speed, and robustness—among Proximal Policy Optimization (PPO), Deep Q-Networks (DQN), and a single-agent RL approach for data center control. PPO demonstrates the highest stability and robustness, along with the fastest convergence, while the single-agent approach shows lower performance across all metrics.

Figure 4: Training Performance Metrics Comparison

Overall, the experimental findings demonstrate that applying reinforcement learning to jointly optimize workload placement and cooling operations can substantially reduce energy consumption while preserving system performance and reliability. These results suggest that RL-driven frameworks hold promise for enabling sustainable, cost-effective, and resilient data center operations. Future work will involve testing in real-world data center settings to verify scalability and operational feasibility. Figure 5 below clearly shows how data center energy is optimized using RL.

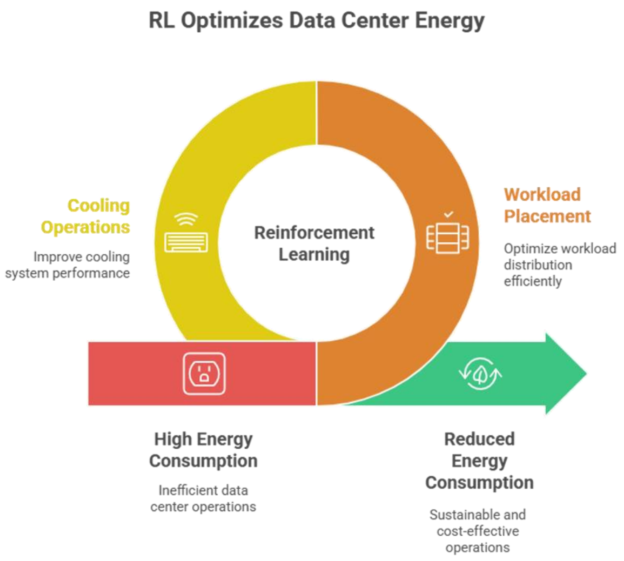

Figure 5: Data Center Energy Optimization by RL

V. DISCUSSION

The experimental results demonstrate that reinforcement learning (RL)-based joint optimization of workload placement and HVAC control significantly enhances energy efficiency in data centers without compromising system performance. By dynamically adjusting cooling parameters and migrating workloads based on real-time thermal conditions, the RL framework prevents the formation of hotspots, which are a major cause of inefficiencies in traditional static or heuristic cooling approaches. This synergy not only reduces cooling energy consumption but also contributes to improved hardware reliability through stable and balanced thermal management. Figure 6 demonstrates optimization of data center energy using RL.

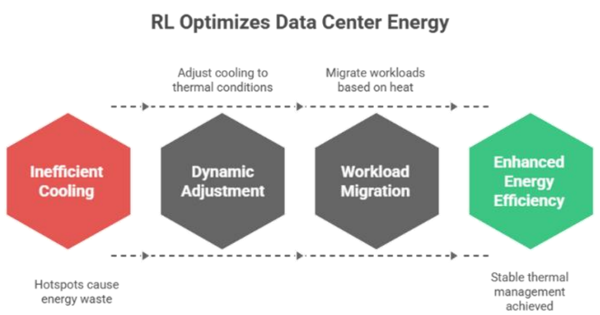

Figure 6: RL Optimizes Data Center Energy

The dual-agent architecture, separating workload scheduling and HVAC control, proved effective in handling the complex interdependencies between computing loads and cooling requirements. This modular approach allows specialized learning and decision-making in each domain while enabling communication and coordination between agents. Such a design supports scalability and adaptability, essential for the diverse and evolving workload patterns in modern data centers. Moreover, the RL agents' ability to anticipate thermal and performance impacts before taking actions differentiates this approach from reactive methods that often lag behind changing conditions.

From a workload performance perspective, the RL framework maintains service-level agreement (SLA) compliance even under workload surges, thanks to a carefully designed reward function that penalizes latency violations. This balance between energy savings and performance reliability addresses a critical operational trade-off in data center management. The RL agents adaptively prioritize SLA adherence during peak loads and focus on energy efficiency during steady operation, demonstrating intelligent context-aware control. This adaptability ensures that energy savings do not come at the cost of user experience or computational throughput.

The bar chart in figure 7 is comparing SLA compliance percentages under steady and peak workload conditions for a baseline heuristic approach and the RL framework. The RL framework maintains higher SLA compliance, especially during peak loads, demonstrating its effective workload performance and adaptability.

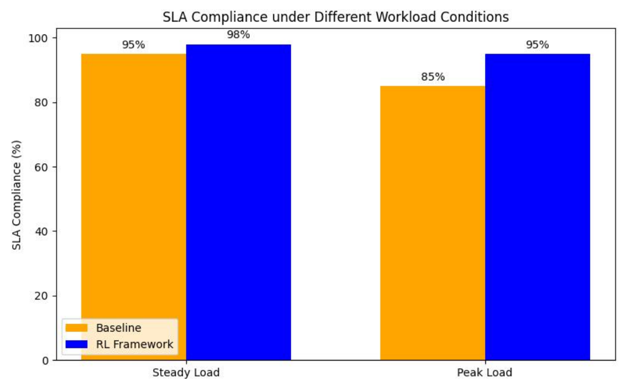

Figure 7: SLA Compliance under Different Workload Conditions

The comprehensive offline training combined with simulation-based fine-tuning facilitates safer exploration and robust policy development. Training RL agents purely in live environments can be risky and inefficient; therefore, the use of realistic simulators calibrated with historical data enables extensive policy refinement without disrupting data center operations. The stable convergence and superior performance of Proximal Policy Optimization (PPO) and Deep Q-Networks (DQN) algorithms support their suitability for such high-stakes, real-time control applications. Ablation studies confirmed that the multi-agent framework outperforms single-agent alternatives, emphasizing the value of specialization in control tasks. The line graph in figure 8 is comparing the training loss convergence of Proximal Policy Optimization (PPO), Deep Q-Networks (DQN), and Single-Agent RL frameworks over training episodes. PPO converges the fastest to the lowest loss, followed by DQN, while the single-agent approach converges more slowly and retains a higher loss.

Figure 8: Training Loss Convergence PPO vs DQN vs Single-Agent

Despite promising results, challenges remain for practical deployment. Real-world data centers exhibit greater complexity, including hardware heterogeneity, unpredictable failures, and multi-tenant environments. Transitioning from simulation to actual operation will require robust validation, safety mechanisms, and continuous learning capabilities to handle evolving system dynamics. Additionally, integration with renewable energy sources and demand response programs could further amplify sustainability gains. Nonetheless, this research establishes a strong foundation for intelligent, sustainable, and resilient data center cooling management through reinforcement learning.

VI. CONCLUSION

This paper presented a novel reinforcement learning (RL) framework that jointly optimizes workload placement and HVAC control to significantly enhance cooling energy efficiency in data centers. By leveraging a multi-agent architecture and advanced RL algorithms such as Proximal Policy Optimization and Deep Q-Networks, the system dynamically adjusts workload distribution and cooling parameters based on real-time thermal and workload conditions. The evaluation through simulations demonstrated considerable reductions in cooling energy consumption, improvements in power usage effectiveness (PUE), and a more balanced thermal profile that reduces hotspots and hardware stress. Importantly, the RL framework maintains strict service-level agreement (SLA) compliance during workload fluctuations, achieving a balance between energy savings and performance reliability. The comprehensive offline training, combined with simulation-based fine-tuning, ensures safe and robust policy learning without disrupting live data center operations. The multi-agent design outperforms single-agent approaches, underscoring the benefit of specialized policy learning for different control domains. While challenges remain for real-world deployment—such as handling system heterogeneity and evolving workload patterns—this research lays a strong foundation for integrating AI-driven cooling optimization into sustainable and resilient data center management. Future work includes validation in operational environments and integration with renewable energy and demand response strategies to further reduce environmental impact.

VII. LIMITATIONS AND FUTURE SCOPE

Despite the promising results, several limitations must be acknowledged when applying reinforcement learning (RL) to data center cooling optimization. First, the complexity and heterogeneity of real-world data centers pose challenges for accurate modeling and simulation, which are critical for training effective RL agents. Many existing simulators and datasets may lack detailed environmental, architectural, or workload-specific characteristics, limiting generalizability. Moreover, ensuring safe exploration and policy updates in live environments remains difficult due to potential performance risks and operational constraints. Computational costs and slow convergence rates for RL algorithms in large-scale systems also restrict scalability and real-time applicability. Finally, multi-tenant and multi-data center environments introduce additional layers of complexity that current RL frameworks have yet to fully integrate.

Future work should focus on overcoming these challenges by developing more comprehensive and physics-informed simulators that better mimic diverse data center conditions. Methods that combine offline training with safe online adaptation and transfer learning can help facilitate smooth transitions from simulation to real-world deployment. Advancing multi-agent and federated RL frameworks will allow cooperative control across complex, distributed infrastructures. Furthermore, integrating RL-powered cooling optimization with renewable energy management, demand response, and sustainability goals holds strong potential for greener and economically viable data centers. Research on explainability and human-in-the-loop control will be critical to foster trust and practical adoption among operators. Overall, scalable, robust, and interpretable RL approaches will be central to realizing intelligent, efficient, and resilient data center cooling in the future.

REFERENCES

[1] A. A. Alkrush, M. S. Salem, O. Abdelrehim and A. Hegazi, "Data centers cooling: A critical review of techniques, challenges, and energy saving solutions," *International Journal of Refrigeration,* vol. 160, no. 1, pp. 246-262, April 2024.

[2] H. Lu, Z. Zhang and L. Yang, "A review on airflow distribution and management in data center," *Energy and Buildings,* vol. 179, no. 1, pp. 264-277, November 2018.

[3] R. R. Ahrabi, A. Mousavi, E. Mohammadi, R. Wu and A. K. Chen, "AI-Driven Data Center Energy Profile, Power Quality, Sustainable Sitting, and Energy Management: A Comprehensive Survey," in *IEEE Conference on Technologies for Sustainability (SusTech)*, Los Angeles, CA, USA, 2025.

[4] N. Ahmed, S. Bennati and N. Anglani, "Energy Performance Benchmarking and Indicators: A Comprehensive Framework for Scientifically Sound Data-Driven and Energy Management Improvements in Hospitals," in *IEEE 8th Energy Conference (ENERGYCON-2024)*, Doha, Qatar, 2024.

[5] Y. Ran, H. Hu, Y. Wen and X. Zhou, "Optimizing Energy Efficiency for Data Center via Parameterized Deep Reinforcement Learning," *IEEE Transactions on Services Computing,* vol. 16, no. 2, pp. 1310-1323, 2023.

[6] C. Feng, A. K. Jumaah Al-Nussairi, M. H. Chyad, N. S. Sawaran Singh, J. Yu and A. Farhadi, "AI powered blockchain framework for predictive temperature control in smart homes using wireless sensor networks and time shifted analysis," *Scientific Reports,* vol. 15, no. 1, p. 18168, May 2025.

[7] M. Trigka and E. Dritsas, "Wireless Sensor Networks: From Fundamentals and Applications to Innovations and Future Trends," *IEEE Access,* vol. 13, no. 1, pp. 96365-96399, May 2025.

[8] K. He, Q. Fu, Y. Lu, Y. Wang, J. Luo, H. Wu and J. Chen, "Predictive control optimization of chiller plants based on deep reinforcement learning," *Journal of Building Engineering,* vol. 76, no. 1, p. 107158, October 2023.