Control of Air Free-Cooled Data Centers in Tropics via Deep Reinforcement Learning

Duc Van Le
Computer Science & Engineering
Nanyang Technological University
Singapore

Yingbo Liu
Computer Science & Engineering
Nanyang Technological University
Singapore

Rongrong Wang
Computer Science & Engineering
Nanyang Technological University
Singapore

Rui Tan
Computer Science & Engineering
Nanyang Technological University
Singapore

Yew-Wah Wong
Energy Research Institute
Nanyang Technological University
Singapore

Yonggang Wen
Computer Science & Engineering
Nanyang Technological University
Singapore

ABSTRACT

Air free-cooled data centers (DCs) have not existed in the tropical zone due to the unique challenges of year-round high ambient temperature and relative humidity (RH). The increasing availability of servers that can tolerate higher temperatures and RH due to the regulatory bodies' prompts to raise DC temperature setpoints sheds light upon the feasibility of air free-cooled DCs in tropics. This paper studies the problem of controlling the temperature and RH of the air supplied to the servers in a free-cooled tropical DC below certain thresholds to maintain servers' computing performance and reliability. To achieve the goal, a portion of the hot air generated by the servers is recirculated and mixed with the fresh outside air to adjust the RH of the supply air. To address the complex psychrometric dynamics, we apply deep reinforcement learning to learn the control policy that aims at minimizing the energy used for moving air and on-demand cooling. Extensive evaluation based on real data traces collected from an air free-cooled testbed and comparisons with hysteresis-based and model-predictive control approaches show the superior performance of our solution.

CCS CONCEPTS

• **Hardware** → **Enterprise level and data centers power issues**; • **Computing methodologies** → *Reinforcement learning*.

KEYWORDS

Data centers, air free cooling, deep reinforcement learning

ACM Reference Format:

Duc Van Le, Yingbo Liu, Rongrong Wang, Rui Tan, Yew-Wah Wong, and Yonggang Wen. 2019. Control of Air Free-Cooled Data Centers in Tropics via Deep Reinforcement Learning. In *The 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (BuildSys '19), November 13–14, 2019, New York, NY, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3360322.3360845

BuildSys '19, November 13–14, 2019, New York, NY, USA

ACM ISBN 978-1-4503-7005-9/19/11…$15.00
https://doi.org/10.1145/3360322.3360845

1 INTRODUCTION

Free cooling (a.k.a. economization) that utilizes outside cold air to cool the servers has been increasingly used to improve the energy efficiency of data centers (DCs) [10]. Free cooling reduces the use of traditional refrigerant-based cooling components such as chillers and compressors. In certain climates, free cooling can save more than 70% in annual cooling energy of DCs, which corresponds to a reduction of over 15% in annualized power usage effectiveness (PUE) [21]. For instance, a Facebook's air free-cooled DC in Prineville, Oregon reported an annualized PUE of 1.07 [22] whereas typical DCs have average PUEs of 1.7 [1].

DCs in the tropical climate with year-round high ambient temperature and relative humidity (RH) consume excessive energy in cooling. However, free cooling in tropics has been long thought infeasible. For instance, in the target tropic of this paper, the year-round average temperature is about 27°C with record instant maximum of 37°C; the average RH is about 70% with instant RH up to nearly 100% before/during rainfalls. If the servers cannot tolerate such high temperatures and RHs, the opportunity of utilizing outside air to cool servers will be very limited. Fortunately, to prompt DC operators to raise the temperature setpoints for better energy efficiency, the American Society of Heating, Refrigeration and Air-Conditioning Engineers (ASHRAE) has been working on extending the recommended allowable temperature and RH ranges of servers [23]. For instance, the servers that are compliant with 2011 ASHRAE Class A3 requirement [3] should be able to operate continuously and reliably with supply air temperature and RH up to 40°C and 90%, respectively. Many latest servers (e.g., all Dell's gen14 servers and all HPE's DLx gen9 servers) are compliant with the A3 requirement. Such wide allowable ranges for temperature and RH shed light upon the feasibility of air free cooling in tropics. However, ASHRAE's relaxed requirements are for traditional DCs with clean air that is recirculated within the enclosed DC buildings only. In tropics, the free cooling that continuously passes outside air through the server rooms will introduce extra challenges.

An immediate concern is the servers' potential computing performance throttling due to the high supply air temperature. To address this concern, we have conducted extensive measurements on a free-cooled DC testbed that allows us to maintain the supply air temperature in the range of [20°C, 37°C] through a cooling coil and an air heater. The testbed has a total of eight server racks and more than 200 measurement points to closely monitor the state of the testbed including the server room condition and server statuses. Our 8-month measurements with controlled server room condition and server workload in wide ranges show that the computing performance of the tested servers from four major manufacturers does not drop when the supply air temperature is up to 37°C (i.e., the record instant maximum in our region). Our measurements show that it is possible to apply air free cooling in tropics without degrading the servers' computing performance.

A major and challenging task in operating air free-cooled DCs in tropics is controlling the condition of the air supplied to the servers, which is the primary focus of this paper. Different from the traditional DCs that use filtrated and circulating air in the enclosed DC building to cool servers, air free-cooled DCs continuously inhale outside air that may contain corrosive gaseous and particulate contaminants. As these contaminants have deliquescent RHs (e.g., 65%) lower than the tropic's RH, they will absorb the moisture in the air to form corrosive pastes and acids that will undermine the servers' hardware reliability [4]. Extracting the contaminants from the continuously inhaled outside air will increase capital expenditure (Capex) and operating expenditure (Opex), offsetting or even negating the benefit of air free cooling. To address this challenge, we adopt an approach of mixing a controlled portion of the return hot air from the servers with the fresh outside air to form warm air that will be supplied to the servers. From psychrometrics, the RH of the warm supply air will be lower than that of the fresh outside air. With proper control of the air flows, warm supply air with RH always below the deliquescent RH of the contaminants is beneficial to the server hardware reliability. This RH control approach exploits the heat generated by the servers and their relaxed temperature requirement. When the outside air is too hot, the temperature of the mixed air to meet the RH requirement may exceed the servers' allowable range. In this case, the cooling coil should be used to cool the inhaled outside air.

We formally formulate a problem of minimizing the expected energy consumption of the server room fans and cooling coil over a long time horizon, subject to specified upper bounds for supply air temperature and RH. The control inputs include the supply air volume flow rate, the portion of the return hot air to be mixed with the fresh outside air, and the temperature drop achieved by the cooling coil when being used. A key challenge in solving this problem is the complex psychrometric dynamics. Specifically, there is no closed-form model to describe the supply air temperature and RH. In addition, the power consumption of the server room fans and cooling coil can have complex coupling with the system's psychrometric state. To address these challenges, we apply deep reinforcement learning (DRL) to learn the optimal control policy over a long time horizon. To avoid potential excursions causing thermal unsafety during the online learning phase of DRL's typical workflow, we perform offline learning based on computational models characterizing the psychrometric dynamics and ventilation/cooling energy consumption. The adequately trained DRL agent is then commissioned to control the air free-cooled DC. We extensively evaluate the performance of the proposed DRL-based control approach based on real data traces collected from the free-cooled DC testbed mentioned early and show its effectiveness through the comparison with hysteresis-based and model-predictive control approaches.

To the best of our knowledge, this is the first work that studies the server room condition control for air free-cooled energy-efficient DCs in tropics with year-round high temperature and RH. Our results provide an important basis for full implementations of air free-cooled DCs on the testbed and in production settings.

The remainder of this paper is organized as follows. §2 introduces the background of free cooling and reviews related research. §3 discusses the requirements of air free-cooled DCs in tropics. §4 formulates the control problem. §5 presents a DRL-based solution. §6 presents evaluation results. §7 concludes this paper.

2 AIR FREE-COOLED DC & RELATED WORK

Water chillers and direct-expansion compressor-based air conditioners are traditional cooling equipment used in DCs. In recent years, *free cooling* has emerged as an effective scheme to improve DCs' energy efficiency [10]. It obviates the need of power-intensive water chillers and compressors by passing outside air through some heat dissipation device. A recent study in [9] has shown that by combining free cooling and solar generation, DC's brown energy consumption can be reduced by up to 59%. There are two major free cooling forms [6], i.e., water-side and air-side methods. In the water-side method, an energy-free *heat exchanger* uses water to carry the heat to the outdoor cooling towers. Large fans blow the outside air through the cooling towers to dissipate the heat into the ambient. Differently, the air-side method uses fans to blow the outside air directly into the server rooms without the water intermediary. The hot return air carrying the heat is then guided back using fans into the ambient. To adjust the temperature and/or RH of the air supplied to the server rooms, a portion of the return air can be recirculated and then mixed with the fresh incoming air. Note that free cooling admits minimum use of traditional cooling systems when the control target for the supply air condition cannot be achieved by purely using the water-side or air-side methods. In this study, we focus on the air-side method (referred to as *air free cooling*) due to its simplicity and higher energy efficiency. In particular, we study the control of supply air condition in air free-cooled DCs in the tropical climate that imposes a number of unique challenges as we will discuss in §3.

In what follows, we review related work on air free cooling control and various applications of DRL for saving energy.

Air free cooling control. A few existing studies [7, 8, 18] focus on the supply air condition control in air free-cooled DCs. Goiri *et al.* [7] designed and implemented a real air free-cooled DC testbed called Parasol, which combines the air free cooling with a direct-expansion air conditioner to control the supply air temperature. The work in [8] proposed an air free cooling control approach called CoolAir to maintain the average and variation of supply air temperatures in desired ranges. Based on predicted ambient temperatures, it selects a proper temperature setpoint to limit the temperature variation and minimize the use of traditional cooling. The work in [18] presented an optimization framework that determines the optimal provision from the traditional cooling in an air free-cooled DC to reduce cooling-related Capex and Opex subject to temperature constraints. The above studies [7, 8, 18] focused on the temperature control to avoid server shutdown due to overheating. RH control is usually not considered because free cooling has been recommended for cold and dry locations only [14], where the ambient RH does not exceed 60% in general. Differently, in tropics, the ambient RH is high. From the study by Manousakis *et al.* [19] based on data collected from a number of Microsoft air free-cooled DCs, the hard disk drive (HDD) failure rate is 10x more correlated with RH than temperature. Therefore, in this paper, we jointly address temperature and RH controls in air free-cooled DCs such that the energy savings achieved by the air free cooling can be maximized while maintaining high server performance and reliability.

DRL for energy saving. Reinforcement learning (RL) [25] is a trial-and-error learning approach, in which an *agent* explores and learns the optimal control policy by interacting with its *environment* through a sequence of the environment's states, the actions applied on the environment, and the rewards. DRL that uses a deep Q-network (DQN) as the function approximation of the control policy for the agent, has emerged as an effective method for solving complex control problems with high-dimensional state and action spaces. DRL has been recently applied to develop control policies for the heating, ventilation, and air-conditioning (HVAC) systems of buildings [27, 29]. Wei *et al.* [27] developed a DRL agent for HVAC control in the presence of ambient dynamics. The DRL agent chooses the optimal air flow rates for different zones in the building such that the energy consumption is minimized subject to tenants' comfort requirements. Zhang *et al.* [29] implemented and evaluated a practical DQN agent for a radiant heating system that aims to improve a building's energy efficiency. A three-month experiment in [29] shows that the DQN agent resulted in up to 18.2% heating demand reduction, compared to the rule-based control.

Due to RL's trial-and-error nature, DRL has not been widely used for environment condition control in mission-critical DCs that often have tight requirements on temperatures. Google reported the adoption of DRL for cooling control in several of its DCs [12]. However, Google does not release any technical details. Yi *et al.* [28] applied DRL to allocate computing jobs and reduce servers' energy consumption. To avoid potential unsafety caused by DRL's trial-and-error, the DRL training is performed offline using computational models capturing servers' power and thermal dynamics. This paper applies DRL for controlling air free-cooled DCs in tropics. DRL well addresses the complex thermal and psychrometric dynamics. Similar to [28, 29], we also adopt an offline training approach to preclude the risk caused by the trial-and-error nature of the learning phase. While our paper and the existing study [28] share the same control objective (i.e., to reduce energy consumption), we address different physical dynamics and constraints of the air free cooling design.

3 AIR FREE-COOLED DC IN TROPICS

In this section, we present the design of an air free-cooled DC testbed located in the tropical zone (§3.1). Then, we discuss the temperature requirement (§3.2) and RH requirement (§3.3) for operating air free-cooled DCs in tropics.

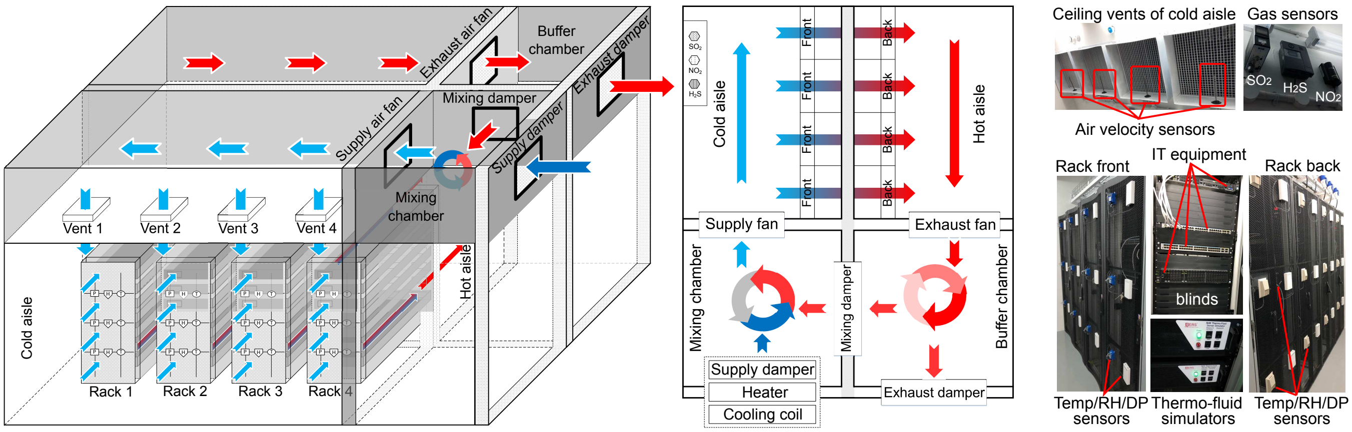
3.1 Air Free-Cooled DC Testbed

To study the feasibility of air free cooling in tropics with high ambient temperature and RH, we designed and instrumented an air free-cooled DC testbed located in the tropical zone. The testbed consists of two identical side-by-side server rooms that are located within the premise of a DC operator. In what follows, we briefly describe the design of a server room. More details can be found in [16].

Figs. 1a and 1b show the 3D and top views of the server room, respectively. The room has two layers with each divided into four chambers. A cooling coil and an air heater are installed on the top layer to process the fresh air inhaled into the test room. Note that the air heater is used only in a set of tests investigating the performance of the servers in high temperatures (cf. §3.2.2). The air free cooling control does not use the heater. Two fans (i.e., supply fan and exhaust fan) are installed on the top layer to move air. Moreover, there are three dampers (i.e., supply damper, exhaust damper, and mixing damper) as shown in Fig. 1. By setting their openness, we can control the air flow paths. The three dampers together are referred to as *damper system*. After the supply fan, the air enters a chamber and then goes down to the cold aisle chamber on the bottom layer through four vents. This design improves the evenness of the cold air volumes passing through the vents. Four 42U server racks are installed on the bottom layer, sitting between the cold aisle and hot aisle chambers. Our design well separates the cold air supplied to the servers and the hot air generated by them. This facilitates the control of the condition of the air supplied to the servers. The hot air is moved by the exhaust fan into a buffer chamber. Depending on the damper system's setting, the hot air is exhausted and/or recirculated to the mixing chamber.

Fig. 1c shows the deployment of some IT equipment and sensors. The racks in each server room host a total of six servers and five 1Gbps switches made by four different manufacturers. All these IT devices were new when they were deployed. To generate more heat and improve the realism of the testbed, for each server room, a total of six thermo-fluid server simulators are mounted on the racks. Their power consumption can be configured and can reach 30 kW totally that is comparable to that of about 100 servers. To well separate the cold and hot aisles, we deploy blinds for the rack slots not mounted with IT equipment and thermo-fluid server simulators. We also install a total of 85 sensors of various modality in each room to monitor the environmental condition as well as the powers consumed by the room facility and IT equipment. Specifically, we deploy the following sensors: (1) a combined temperature and RH sensor outside of the server room to monitor the ambient condition; (2) a temperature sensor in each of the mixing, cold aisle, hot aisle, and buffer chambers; (3) an air velocity sensor at each of the four cold vents to estimate the air volume speed in m^3/h; (4) sensors in the cold aisle for monitoring differential pressure (DP) with respect to atmospheric pressure and concentrations of corrosive gases (SO_2, NO_2, H_2S); (5) temperature, RH, and DP sensors at three heights on the front and back sides of each rack; (6) power meters to monitor the power of server rack, cooling coil, heater, and fans. The dense sensor deployment is for research only. In §5.4, necessary sensors for free cooling control will be discussed.

The real-time measurements of several sensors (e.g., temperature and air volume speed) are also used by various control algorithms to maintain the test room's environmental condition. For instance, the total air volume speed supplied to the servers can be maintained at a specified setpoint up to 12500 m^3/h by a proportional-integral-derivative (PID) controller for the supply and exhaust fans.



(a) 3D view. (b) Top view. (c) Equipment.

Figure 1: The air free-cooled DC testbed used in this work. Arrows represent the air flows.

3.2 Supply Air Temperature Requirement

3.2.1 Impact of temperature on server safety and reliability. Too high instantaneous supply air temperatures may cause permanent damages to server hardware components. To avoid the damage, most servers will automatically halt for self protection when the temperatures measured by the built-in sensors of the server enclosure exceed certain *safety thresholds*. For instance, a server deployed on our testbed has a safety threshold of 45°C for its inlet temperature sensor. For continuous operation of a DC, the safety thresholds of the servers must not be exceeded.

Besides the permanent damages caused by too high temperatures instantly, high temperatures are generally thought generating negative impact on the server hardware's long-term reliability that is often measured with annualized failure rate (AFR). A basis of this hypothesis is the Arrhenius equation that characterizes the temperature dependence of reaction rates in physical chemistry [13]. The electronics industry adopts this equation to predict that the failure rate of an electronic device increases exponentially with the temperature [11]. Based on this, ASHRAE, together with DC IT equipment manufacturers, provides the *x-factors*, which are the relative failure rates under certain temperatures, as a guideline for choosing DC temperature setpoint [3]. For instance, with a temperature of 37.5°C, the x-factor is 1.61, meaning that the failure rate at 37.5°C will be 1.61 times of the failure rate at the reference temperature of 20°C. For example, if the baseline AFR of HDDs at 20°C is 1.25% according to a cloud service provider's statistics [2], the AFR at 37.5°C is $1.25\% \times 1.61 \simeq 2\%$, i.e., two out of 100 HDDs fail over one year. Since the baseline AFR for any server component is low in general, the absolute increases of AFR due to higher temperatures are not significant. In particular, the recent advances in materials development and hardware design enable manufacturers to build more robust DC IT equipment that can tolerate higher temperatures and RHs. For example, many modern servers (e.g., all Dell's gen14 servers and all HPE's DLx gen9 servers) are compliant with ASHRAE Class A3 requirement [3]. Specifically, these servers can continuously and reliably operate under a temperature range of [5°C, 40°C] and RH range of [8%, 85%]. We call the temperature upper limit for a server's design reliability as *reliability threshold*, e.g., 40°C for ASHRAE Class A3 servers. Note that for a server, the reliability threshold is in general lower than the safety threshold, because the latter concerns about instant damages.

The ambient temperature of the tropical area that we are in has a record minimum of 19.4°C and maximum of 37.0°C. Thus, by using air free cooling only, it is possible to maintain the supply air temperature below modern servers' reliability thresholds. However, close monitoring and cautious control of the supply air temperature are still needed, because of the following. First, uncontrolled hot air recirculation due to imperfect separation of the cold air and hot air aisles may increase the supply air temperature. Second, as discussed shortly in §3.3, to reduce RH, an energy-efficient approach is to use controlled hot air recirculation to raise the supply air temperature. However, it reduces the buffer region from the safety thresholds. Thus, without cautious control, the system will have increased risk of server shutdown caused by overheating.

3.2.2 Impact of temperature on server performance. Another common concern is that high temperatures may cause degraded computing performance of servers. We conduct extensive controlled experiments over a duration of about eight months to investigate the impact of supply air condition (temperature and volume flow rate) on the server performance. We concluded that, when the supply air temperature is up to 37°C (i.e., the record maximum in our area), the temperature has no impact on the servers' computing performance if a sufficient air flow rate is maintained (e.g., $2,500\ \mathrm{m}^3/\mathrm{h}$ for one server room of our testbed). This section briefly summarizes the experiment methodology and results.

We separately benchmark the CPUs, HDDs, and main memories, which are the main components related to servers' computing performance. For each component, we vary the supply air temperature, the air volume flow rate, and the operating setpoint of the tested server component in their respective ranges. Table 1 summarizes the ranges of these parameters and the corresponding numbers of steps. Under each setting, we conduct a 1-hour experiment to measure giga floating point operations per second (GFLOPS) for CPU, input/output operations per second (IOPS) and response time

Table 1: Settings for server performance benchmark.

Parameter	Minimum	Maximum	Steps
Supply air temperature	25°C	37°C	13
Room air flow rate (m^3/h)	2,500	12,500	5
CPU utilization	10%	90%	7
HDD throughput (MB/s)	10	100	6
Memory block size (KB)	8	256	6

Table 2: Server performance benchmark results.

°C	GFLOPS	IOPS	RespTime (ms)	MemSpeed (MB/s)
25	303.97	6397.0	0.28	2873.12
27	301.48	6397.0	0.25	2990.79
29	282.41	6398.0	0.25	3539.04
31	300.01	6398.0	0.27	3191.80
33	308.63	6398.0	0.24	2977.22
35	295.19	6401.0	0.32	2792.49
37	304.15	6398.0	0.25	2798.15

CPU utilization: 90%; HDD throughput: 100 MB/s; memory block size: 256 KB

for HDD, and speed of data coping for memory. Benchmark results for a total of 1,235 net test hours have been collected. Table 2 shows the benchmark results for a CPU, an HDD, and a memory, under different supply air temperatures, and specific settings of CPU utilization, HDD throughput, and memory block size. We can see that the performance metrics remain stable when the temperature is up to 37°C. Other CPUs, HDDs, and memories also exhibit such stable trend. We also conducted experiments to jointly benchmark CPU, HDD, and memory, such that all these components generate heat simultaneously. Similarly, we observed no statistically significant impact of temperature on the computing performance within the test ranges specified in Table 1.

3.3 Supply Air RH Requirement

RH is the ratio of the amount of moisture contained in the air at a given temperature to the maximum amount of moisture that the air can hold at the same temperature. As discussed in §3.2, modern servers can operate reliably under high RHs (up to 85%) under typical DC settings. In typical air cooled DCs, the air is circulated within the DC building without admitting much fresh air from the outside; any admitted fresh air will be filtrated to control the concentrations of gaseous and particulate contamination [4]. Research has shown that, with clean air, RH has little impact on the IT hardware reliability [24].

Differently, in the air free cooling scheme, the outside air continuously passes through the server rooms. The solutions to control the concentrations of gaseous and particulate contamination will increase Capex for installing the filtration facility and Opex for filtration energy consumption and consumable component replacement. Thus, the design of our testbed chooses not to integrate the costly continuous air filtration solutions; it only applies a MERV 6 filter to remove PM10 or larger particles. Finer particles and corrosive gases (e.g., SO_2, H_2S, NO_2, and Cl_2) generated by transportation systems and industrial processes can negatively affect the reliability of the IT equipment. Specifically, if the RH of the supply air is higher than the deliquescent RH of the particles and gases, these contaminants will absorb the air moisture to form corrosive pastes and acids that will promote corrosion and/or ion migration of the IT hardware materials [4]. Corrosion can easily cause short circuits given today's dense layouts of printed circuit boards.

Existing studies have shown that the co-presence of high RH and air contaminants lead to reduced server hardware reliability. Svensson *et al.* [26] observed that the increase of RH from 75% to 95% results in about 9x higher corrosion rate of zinc at the same concentration level of SO_2. A study [24] showed that, in the presence of gaseous contamination, the copper corrosion rate of DC IT equipment increases with RH.

Therefore, it is important to maintain low RH for the supply air. This is a challenging requirement because the ambient RH in tropics is generally high. From our measurements, the ambient RH has an average of 71% with instant measurements up to 100%. Note that the deliquescent RH for many contaminants is about 65% [4].

4 PROBLEM FORMULATION

From §3.2 and §3.3, to achieve success of air free-cooled DCs in tropics, we will need to maintain the supply air temperature below the reliability thresholds of the servers and RH below a certain level (e.g., the lowest deliquescent RH of the particulate and gaseous contaminants present in the outside air). In this section, §4.1 overviews our approach to meeting the requirements; §4.2 presents a Markov Decision Process (MDP) formulation of the problem, which will be addressed by using DRL in §5.

4.1 Approach Overview

RH control is a challenging task in tropics. Traditionally, dehumidification is achieved by a cooling-then-reheating process. Specifically, a certain amount of moisture is condensed out from the humid air by cooling the air below its dew point. Then, the cold air is reheated to the desired temperature. However, the cooling and reheating processes consume significant energy. In this study, to reduce the supply air RH in an energy-efficient manner, we recirculate a portion of the hot return air and mix it with the fresh outside air to form supply air. The mixing can be implemented by controlling the openness of the three dampers as illustrated in Fig. 1. Note that, without condensation, the hot return air and the fresh outside air have the same absolute humidity. From psychrometrics, the hotter mixed air will have lower RH compared with the fresh outside air. However, when the fresh outside air is hot, the hotter mixed air to achieve the desired low RH may exceed the servers' reliability thresholds. In this case, the cooling coil should be activated to reduce the temperature of the incoming air.

This paper develops control algorithms for the supply and exhaust fans, the cooling coil, and the dampers such that the energy consumption of the non-IT facility is minimized subject to that the temperature and RH of the air supplied to the servers are below respective specified thresholds for the sake of IT hardware reliability.

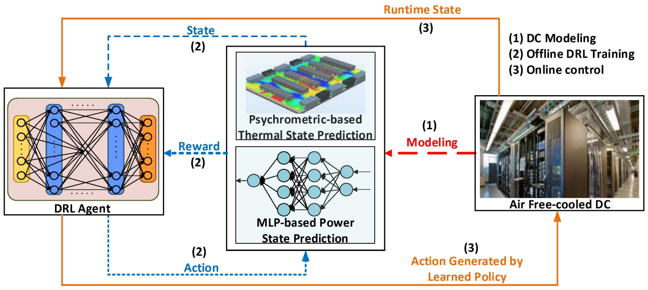


Figure 2: Design workflow of DRL-based DC control.

The system will operate in the presence of exogenous disturbances, i.e., the time-varying ambient condition and heat from servers.

4.2 MDP Formulation

Time is divided into intervals with identical duration of τ seconds. In this paper, we consider the secondary controls (i.e., adjustment of setpoints) for the actuators. The beginning time instant of a time interval is called a *time step*. Control action is performed at every time step. Thus, the τ is referred to as *control period*. In this paper, we do not consider the details of the primary controls of the actuators; we assume that the actuators can implement the setpoints decided by the secondary controls using their closed-loop primary controls and the system has reached the steady state by the end of every control period. In practice, the setting of the control period can be chosen with the consideration of the dynamics of the primary controls to ensure the above assumption. Under the above setting, the temperature and RH of the supply air at next time step depend only on the system's state (conditions of outside and supply air, servers' powers) and the control action at the current time step (cf. §5.2). Therefore, the control problem can be modeled as a Markov decision process (MDP). We now define the terminologies of the MDP formulation.

System state: The system state, denoted by x, is a vector $x = [t_s, \phi_s, p_{\text{IT}}, t_o, \phi_o]$, where t and ϕ respectively represent temperature and RH, the subscript s and o respectively represent supply air and outside air, and p_{IT} represents the total power consumption of all IT equipment in the server room. The p_{IT} determines the amount of heat generated in the server room.

Control action: The supply and exhaust fans admit air volume flow rate setpoints. To achieve steady state without control errors, the setpoints for the two fans should be identical; otherwise, the server room will be in the dynamic process of pressurization/depressurization or a steady state with control errors. Let $\dot{v}_s \in [0, \dot{v}_{\max}]$ denote the air volume flow rate setpoint for the two fans, where $\dot{v}_{\max}$ is the maximum achievable air volume flow rate. The cooling coil admits a setpoint Δt that represents the reduction of temperature, i.e., $\Delta t = t_o - t_p$, where t_p represents the temperature of the processed air leaving the cooling coil. Let $\Delta t_{\max}$ represent the maximum temperature reduction that can be achieved by the cooling coil. Thus, $\Delta t \in [0, \Delta t_{\max}]$. Let $\alpha \in [0, 1]$ denote the setpoint for the damper system, which is the fraction of the recirculated hot air in the supply air. Thus, $1 - \alpha$ is the fraction of the outside air in the supply air. A setpoint α can be achieved by controlling the openness of the three dampers. For example, to achieve $\alpha = 0$, the supply and exhaust dampers should be completely open and the mixing damper should be completely closed; to achieve $\alpha = 1$, the supply and exhaust dampers should be completely closed and the mixing damper should be completely open. The control action, denoted by a, is a vector $a = [\dot{v}_s, \Delta t, \alpha]$.

Reward function: When a control action a is performed at the current time step with a system state of x, let $p(x, a)$ denote the average power consumed by the supply and exhaust fans to maintain the air volume flow rate $\dot{v}_s$ and the cooling coil to lower the temperature by Δt Celsius degree over the next control period of τ seconds; let $t_s(x, a)$ and $\phi_s(x, a)$ denote the supply air temperature and RH, respectively. We define a penalty function as follows:

$$q(x, a) = \lambda_1 \cdot \max\left(t_s(x, a) - t_{\text{th}}, 0\right) + \lambda_2 \cdot \max\left(\phi_s(x, a) - \phi_{\text{th}}, 0\right), \quad (1)$$

where t_{th} and ϕ_{th} are the temperature and RH thresholds for the long-term reliability of the IT hardware equipment; λ_1 and λ_2 are configurable weights. From the definition of $q(x, a)$, if the supply air temperature and RH do not exceed their respective thresholds, no penalty will be applied. The immediate reward, denoted by $r(x, a)$, is defined as

$$r(x, a) = -p(x, a) - q(x, a). \quad (2)$$

Thus, the reward is defined based on the weighted sum of the non-IT power consumption and the degrees of supply air temperature and RH requirement violations. The impact of λ_1 and λ_2 on the control performance will be evaluated and discussed in §6.

Air free-cooled DC control problem: At every time step, the system controller observes the system state x. Then, it decides and executes a control action a to operate the supply and exhaust fans, cooling coil and dampers in the next control period of τ seconds. At the end of the next control period, the system controller can receive an immediate reward $r(x, a)$ as a feedback signal. The control design objective is to find a control policy that determines a based on x to maximize the expected reward over a long run, i.e., $\mathbb{E}[r]$.

In general, it is difficult to design a closed-form control policy to maximize $\mathbb{E}[r]$ because the state evolution of the system (i.e., $t_s(x, a)$ and $\phi_s(x, a)$) is complex. Model-predictive control (MPC) is a widely adopted approach to solve MDP problems (e.g., [17] for HVAC control). However, the optimization of MPC is computationally expensive and often for a limited time horizon only. DRL is an emerging approach to deal with the above challenges. In the interactions between the DRL agent and the environment (i.e., the controlled system), the agent will learn the optimal control policy from the historical data including system states, control actions, and the resulted immediate rewards. With sufficient interactions, the DQN learned by the agent can well capture the highly complex system dynamics. Moreover, the learned control policy approaches optimality for a long time horizon comparable to the time duration of the training phase. In §5, we will present the detailed design of our DRL system to address the air free-cooled DC control problem.

5 DRL-BASED FREE-COOLED DC CONTROL

5.1 Design Workflow

Typically, DRL agent learns the optimal control policy during the online interactions with the controlled system. However, for free-cooled DC control, the online learning scheme has the following two issues. First, it may take a long time duration to converge, especially when the state and action spaces are large. Second, during the learning phase, excursions due to RL's trial-and-error nature may lead to overheating and server shutdowns. To address these issues, we adopt an offline training approach. Fig. 2 illustrates the workflow of the approach, which consists of three steps. First, we build psychrometric and multilayer perceptron (MLP) models based on meta information and real data traces collected from the DC to characterize the supply air temperature and RH as well as the non-IT power consumption. Second, we use the models built in the first step to drive the offline training of the DRL agent. Third, after the completion of the offline training, the DRL agent is commissioned to control the actual free-cooled DC.

5.2 Modeling Air Free-Cooled DC

This section derives the dynamic model that describes the evolution of the steady system state of the air free-cooled DC. We also build three MLPs to characterize the power consumption of servers, cooling system, and supply/exhaust fans. These models are used for the offline training of DRL agent.

5.2.1 Dynamic model of system state. In this section, we perform psychrometric analysis for the four steps of the air processing in the air free-cooled DC, i.e., *heating* in the server room, *buffering* in the buffer chamber, *cooling* by the cooling coil, and *mixing* by the damper system. Based on these models, we construct a Markovian computational model to characterize the psychrometric dynamics:

$$t_s[k+1], \phi_s[k+1] = f(t_s[k], \phi_s[k], t_o[k], \phi_o[k], \dot{v}_s[k], \Delta t, p_{\mathrm{IT}}[k], \alpha),$$

where $k \in \mathbb{Z}$ represents the index of time step.

We define the following notation: $\dot{m}$ is mass flow rate, h is enthalpy, w is moisture content; for the above psychrometric variables, we use the subscripts $\cdot_s$, $\cdot_h$, $\cdot_r$, $\cdot_p$, $\cdot_o$ to refer to the supply air in the cold aisle, the hot air generated by the servers, the recirculated hot air from the buffer chamber to the mixing chamber, the processed air leaving cooling coil, and the outside air provided to the cooling coil, respectively. The four steps are as follows:

(1) Heating: Servers generate heat and introduce no extra moisture. Thus, the air enthalpy at the hot aisle is higher than that at the cold aisle, while the moisture contents at the two aisles are identical. Denoting by η the servers' heat rate transfer coefficient, the psychrometics of the server room is

$$\dot{m}_s h_s + \eta p_{\mathrm{IT}} = \dot{m}_h h_h, \quad \dot{m}_s = \dot{m}_h, \quad w_s = w_h. \tag{3}$$

(2) Buffering: The hot aisle air is transported into the buffer chamber by the exhaust fan. Under the setpoint α for the damper system, the buffer chamber is characterized by

$$\dot{m}_r = \alpha \dot{m}_h. \tag{4}$$

(3) Cooling: The total energy of ideal gas is the sum of dry air's energy and water vapor's energy. Without condensation, the cooling coil does not change moisture content of air passing through. Moreover, it does not change mass flow rate. Thus, the condition of the air leaving the cooling coil is given by

$$h_p = c_p(t_o - \Delta t) + w_p(c_{pw}(t_o - \Delta t) + l), \; w_p = w_o, \; \dot{m}_p = \dot{m}_o, \tag{5}$$

where c_p and c_{pw} respectively represent the specific heat of dry air and water vapor which are constants; l represents the evaporation heat. Note that $c_p(t_o - \Delta t)$ is the enthalpy of the dry air leaving the cooling coil; $w_p(c_{pw}(t_o - \Delta t) + l)$ is the enthalpy of the water vapor leaving the cooling coil.

(4) Mixing: The air leaving the cooling coil and the recirculated hot air are mixed in the mixing chamber. Governed by the conservation of mass and energy, the psychrometrics of the mixing process can be characterized by

$$(1-\alpha)h_p + \alpha h_r = h_s, \quad (1-\alpha)w_p + \alpha w_r = w_s, \quad \dot{m}_p + \dot{m}_r = \dot{m}_s. \tag{6}$$

Taking the moisture contents of the two influxes as boundaries, Eq. (6) suggests that the outflow's moisture content will be in between, which is the basis of the RH control through adjusting α.

The above models in Eqs. (3)-(6) are for enthalphy, moisture content, and mass flow rate. These quantities can be converted to temperature, RH, and volume flow rate according to the equations presented in [5]. The aforementioned Markovian computational model is as follows. By initializing the h_s and w_s in Eq. (3) with the current state of the supply air condition (i.e., $t_s[k]$ and $\phi_s[k]$), we use the remaining equations in Eqs. (4)-(6) to update h_s and w_s. The updated values are then used to initialize the h_s and w_s in Eq. (3) again and then solve Eqs. (4)-(6). This process is iterated until h_s and w_s converge; the converged values are converted to $t_s[k+1]$ and $\phi_s[k+1]$. Thus, the Markovian computational model has no closed-form expression, presenting a challenge to the design of optimal control policy.

5.2.2 Power consumption models. We design three MLPs to model the following powers averaged over the next control period: (1) IT power $p_{\mathrm{IT}}[k+1]$, (2) total power of supply and exhaust fans $p_f[k+1]$, and (3) cooling coil power $p_c[k+1]$. The MLPs use the respective power measurements in the past K control periods as a part of the input to address the autocorrelation of power consumption. Moreover, the MLPs use additional inputs that will be discussed below. Note that the hyperparameters of the MLPs (e.g., the number of layers and neurons) will be designed in §6 based on real traces.

The first MLP (MLP1) modeling $p_{\mathrm{IT}}[k+1]$ additionally takes $t_s[k]$ and $\dot{v}_s[k]$ as inputs. This is because (1) higher temperatures lead to higher rotation speeds of server fans and CPU fans, (2) air flow generates forces on the fan blades. The second MLP (MLP2) modeling $p_f[k+1]$ additionally takes $\dot{v}[k]$ and $t_s[k]$ as inputs. This is because (1) fan power increases with fan speed, (2) with a higher temperature, materials exhibit higher strength, resulting in the increase of stresses on rotating components. The third MLP (MLP3) modeling $p_c[k+1]$ additionally takes $\Delta t[k]$ and $\dot{v}_s[k]$ as inputs. This is because (1) the setpoint Δt determines the cooling capacity needed, (2) the cooling coil consumes more power when it processes a larger volume of air.

The average non-IT power consumed in the next control period is $p[k+1] = p_f[k+1] + p_c[k+1]$, which is a part of the reward.

5.3 Offline Training of DRL Agent

We adopt the learning framework in [20] to train offline a DQN for the control agent to capture a good control policy to address the problem formulated in §4.2. Specifically, the DQN is trained through interacting with the computational model developed in §5.2 for N *episodes*, each of which consists of T time steps. An

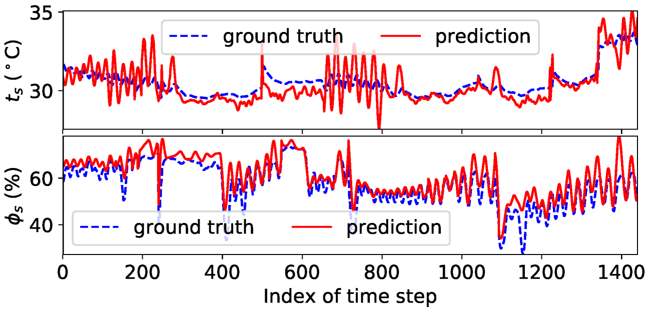


Figure 3: Prediction results by the psychrometric model (RMSEs for t_s and ϕ_s are 0.83°C and 5.3%, respectively).

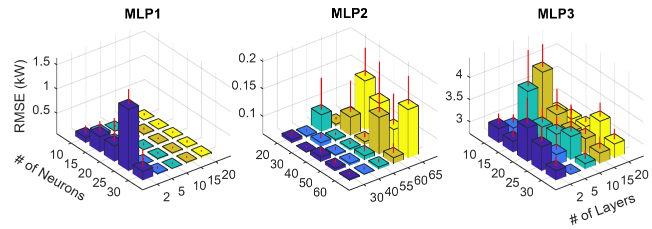


Figure 4: Impact of hyperparameters on MLPs' performance.

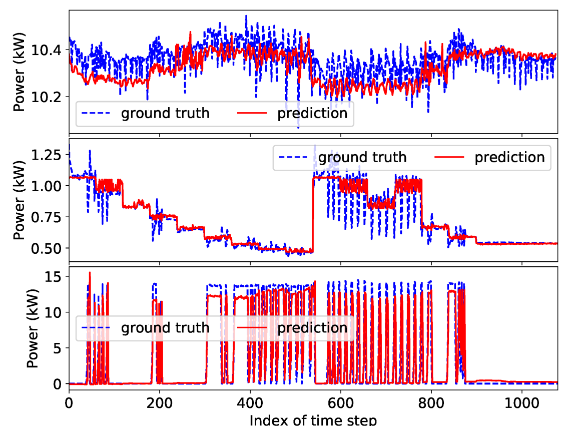


Figure 5: Prediction results of MLPs. Top: IT power; middle: fan power; bottom: cooling coil power.

episode starts with a state chosen randomly from the training data. Then, at the kth time step, an action $a[k]$ is selected for state $x[k]$ according to the ε-greedy algorithm [25] based on action-values given by the DQN. Given the selected action $a[k]$, the $t_s[k+1]$, $\phi_s[k+1]$ and $p_{\mathrm{IT}}[k+1]$ are estimated using the psychrometric model and IT power model (i.e., MLP1), where the outside air condition (i.e., $t_o[k+1]$ and $\phi_o[k+1]$) are taken from real traces. To calculate the immediate reward $r[k]$, powers of fans and cooling coil with respect to the selected $\dot{v}_s[k]$ and $\Delta t[k]$ are determined using MLP2 and MLP3, respectively.

During the learning phase, two mechanisms, i.e., experience replay and target Q-network, are used to update the weights of the DQN (denoted by θ) every time step. For the target Q-network mechanism, we use the soft target update method [15] to update the weights θ' of the target Q-network by setting $\theta' = \beta\theta + (1-\beta)\theta$ with $\beta \ll 1$. The soft target update often gives better learning stability than the hard target update of the original DQN training.

5.4 Sensor Requirement

The testbed presented in §3.1 is instrumented with many sensors to monitor the system state. To run the trained DRL agent, the essential sensors include: (1) temperature and RH sensors to monitor the outside air and supply air conditions; (2) a meter to monitor the total power consumption of the IT equipment. Moreover, to implement the primary controls of the supply/exhaust fans, the cooling coil, and the damper system, we need the following sensors: (1) air volume flow rate sensors to monitor the air entering the cold aisle and the air passing the mixing damper; (2) a temperature sensor measuring the air leaving the cooling coil. To collect training data for the offline learning of DRL, meters to measure the power consumption of supply and exhaust fans, as well as the cooling coil are needed in addition to the sensors mentioned above.

6 PERFORMANCE EVALUATION

This section evaluates the system state prediction and the DRL-based controller using simulations driven by real data collected from the air free-cooled testbed. The prediction and DRL are implemented in Python 3.5 with Keras 2.1.6 using TensorFlow 1.8.0.

6.1 Accuracy of Air Free-Cooled DC Modeling

6.1.1 Psychrometric model of system state. We use data traces collected during the controlled experiments on the testbed (cf. 3.2.2) to evaluate the psychrometric models presented in §5.2.1. The inputs to the model are t_o, ϕ_o, p_{IT}, $\dot{v}_s$, Δt, and α; the outputs are the predicted t_s and ϕ_s. We use root mean squared error (RMSE) between the prediction and the ground truth as the evaluation metric. Fig. 3 shows the prediction results over a time duration of 24 hours. We can see that the prediction by the psychrometric model well tracks the ground truth. The RMSEs for t_s and ϕ_s are just 0.83°C and 5.3%, respectively, over an evaluated period of 24 hours.

6.1.2 MLP-based power prediction. We evaluate the three MLP models presented in §5.2.2 for predicting IT power, cooling power, and fan power. Each MLP is trained, validated and tested using 1375, 700 and 1080 data samples, respectively. The settings of K (i.e., the respective power measurements in the past K control periods used for prediction) for the three MLPs are 5, 1, and 1. For all MLPs, the training batch size is set to 128; the training time is 3,000 epochs. The Adam optimizer with a learning rate of 0.001 is used for training. Moreover, we use the rectified linear units (ReLUs) as the activation function for input and hidden layers; we use linear units for output layer. We conduct extensive evaluation to choose the number of hidden layers and neurons for each MLP to minimize the prediction RMSEs. The evaluation for a certain combination of hyperparameter settings is repeated 5 times to account for the randomness of the training.

Fig. 4 shows the error bars for testing RMSEs with various hyperparameter settings of the number of hidden layers and the number of neurons. MPL1 achieves the smallest RMSE of 0.10 ± 0.07 kW

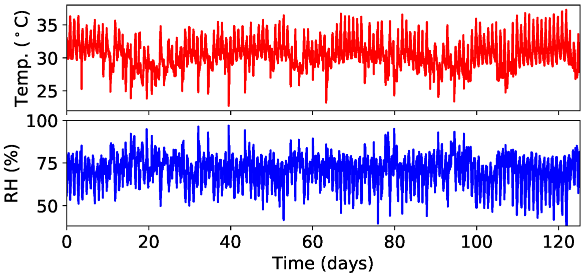


Figure 6: Outside air condition over 125 days.

with 5 hidden layers, each of which has 20 neurons. MPL2 achieves the smallest RMSE of 0.07 ± 0.002 kW with 60 layers, each of which has 40 neurons. MLP3 achieves the smallest RMSE of 2.73 ± 0.03 kW with 5 layers, each of which has 30 neurons. Fig. 5 shows the ground truth and the prediction by the three MLPs with the chosen hyperameters over a time duration of 18 hours. Overall, the predictions well track the ground truths.

6.2 DRL Agent Training and Execution

6.2.1 Settings. We build fully connected deep neural networks as the DQNs (i.e., the primary and target action-value functions). Each network consists of an input layer, four hidden layers and a linear output layer, where each hidden layer has 20 ReLUs. From our extensive trials, the choice of four-layer perception achieves satisfactory convergence performance for the control of the simulated testbed. The DRL agent admits a system state and chooses an action $a = [\dot{v}_s, \Delta t, \alpha]$ from a discrete action space: $\dot{v}_s$ is from 1000 m^3/h to 5000 m^3/h with step size of 500 m^3/h; Δt is from 0°C to 10°C with step size of 1°C; and α is from 0 to 1 with step size of 0.1. These step sizes are from the physical constraints of the supply/exhaust fans, the cooling coil, and the damper system. The size of the action space is $9 \times 11 \times 11 = 1089$. We set the RH threshold $\phi_{\text{th}} = 65\%$, which is the deliquescent RH of many contaminants [4]. We set $t_{\text{th}} = 45$°C, which is the reliability temperature of ASHRAE Class A4 servers. The control period is one minute. For the offline training of the DQN, we adopt the following settings: training batch size is 64; replay memory size is 50000; discount factor $\gamma = 0.99$; sort target update weight $\beta = 0.01$; Adam optimizer's learning rate is 0.001; the ε of the ε-greedy method reduces linearly from 1 to 0.1.

6.2.2 DRL agent training. Fig. 6 shows 125 days' outdoor air conditions of the testbed area. We use the first 95 days' data for training the DRL agent and the remaining data for evaluating the trained agent. The offline training is for $N = 5000$ episodes, each of which consists of $T = 1000$ control periods. At the beginning of each episode, we select a batch of 1,000 samples of outside air condition to drive the training. During the training, the system state is determined based on the action taken by the agent, the psychrometric model, and the power models in §5.

The weights λ_1 and λ_2 in Eq. (1) affect the trade-off between power consumption and compliance to the temperature/RH requirements. We evaluate the convergence of the DRL agent training under various settings for λ_1 and λ_2. Fig. 7 shows the training traces

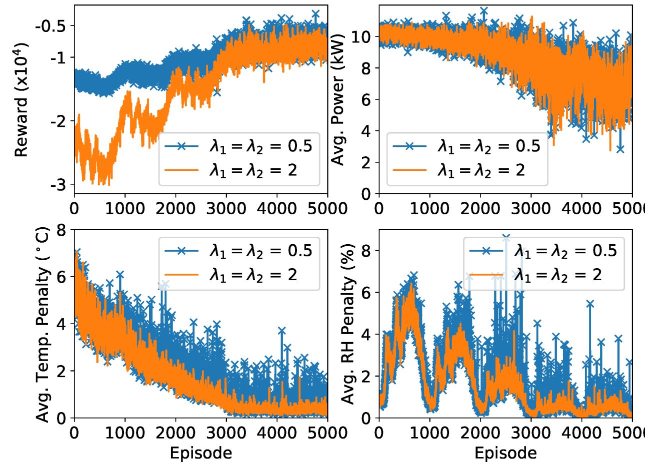


Figure 7: DRL training convergence and penalty factors.

of reward, average power, and average temperature and RH penalties (i.e., max $(t_s(x, a) - t_{\text{th}}, 0)$ and max $(\phi_s(x, a) - \phi_{\text{th}}, 0)$) over an episode of 1000 time steps. Along the training episodes, the reward becomes flat; the power consumption has increasing variance but decreasing overall trend. Both temperature and RH penalties drop during training. With $\lambda_1 = \lambda_2 = 2$, the penalties are close to zero after 5,000 episodes. Differently, with $\lambda_1 = \lambda_2 = 0.5$, the penalties are higher. We also train the DRL agent under various settings for temperature and RH thresholds as well as weights. The results show that the training of the DRL agent is convergent after a certain number of training episodes (e.g., $N = 5000$) with learning curves similar to those shown in Fig. 7.

6.2.3 DRL agent execution. We evaluate the execution of the trained DRL agent for controlling the system in trace-driven simulations over a period of 30 days. The last 30 days' outdoor air condition trace shown in Fig. 6 is used to drive the simulations. Fig. 8 shows the total energy consumption and boxplots for the distributions of supply air temperature and RH over the execution period of 30 days with the DRL agents trained with various λ_1 and λ_2 settings. From the 1st subfigure, the energy consumption increases with the weight. This is because, with smaller λ_1 and λ_2, the agent is trained towards saving more power. However, from the 2nd and 3rd subfigures, when the two lambdas are no greater than 1, the temperature and RH may exceed their thresholds. When the two lambdas are 2, the temperature and RH do not exceed their thresholds during the 30-day test period. The above results show the trade-off between the energy consumption and the temperature/RH requirement compliance. In practice, grid search can be applied to choose the settings of λ_1 and λ_2 based on training and validation data, to achieve compliance of the temperature and RH requirements.

6.2.4 Comparison with baselines. We compare our DRL-based approach with two baseline approaches: hysteresis-based and model-predictive control (MPC) approaches. The hysteresis-based approach adopts the maximum setpoints for the cooling coil and fans: $\Delta t = 10$°C and $\dot{v}_s = 5000 m^3$/h. The initial setpoint of the damper system α is 1. At the beginning of every control period, if the current supply air RH $\phi_s < \phi_{\text{th}} - \Delta\phi$, where $\Delta\phi \geq 0$, α is decreased by 0.1;

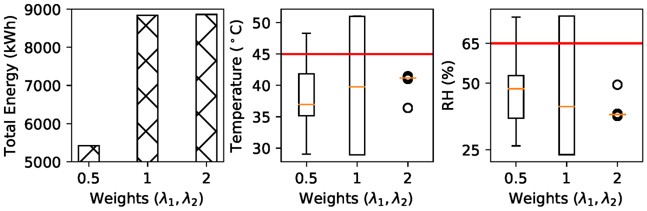


Figure 8: Execution results under various λ_1 and λ_2 settings. The red lines in 2nd and 3rd subfigures represent t_{th} and ϕ_{th}.

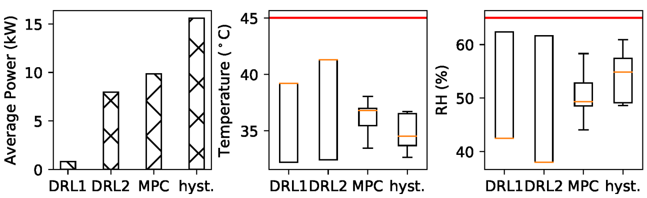


Figure 9: Comparison with baselines. $t_{\text{th}} = 45$°C, $\phi_{\text{th}} = 65\%$ (i.e., the two red lines); $\lambda_1 = \lambda_2 = 0.5$ (for DRL1), $\lambda_1 = \lambda_2 = 1.5$ (for DRL2), $\Delta\phi = 10\%$ (for hysteresis), $N_p = 10$ (for MPC).

otherwise, α is increased by 0.1. In MPC, the controller schedules the future N_p actions such that the predicted total non-IT energy consumption is minimized subject to the temperature and RH constraints (i.e., t_{th} and ϕ_{th}). The MPC algorithm is implemented using the nonlinear MPC toolbox in MATLAB 2019.

We conduct trace-driven simulations for the DRL, MPC, and hysteresis approaches over 1,000 control periods (i.e., about 16.7-hour simulated time). On a workstation computer with a 3.5 GHz CPU and 16 GB RAM, the three controllers need 0.014, 572.67, and 0.012 seconds on average to determine an action, respectively. Since each control period is 60 seconds only, the MPC approach violates the timeliness requirement. Thus, the MPC is merely used as a baseline for understanding the result of DRL; it cannot be applied in practice. Moreover, the excessive time for executing the MPC solver prevents us from running the simulations for long simulated time.

Fig. 9 shows the average power consumption and boxplots for the distributions of supply air temperature and RH under the three approaches. From the 2nd and 3rd subfigures, all three approaches can meet the temperature and RH requirements. From the 1st subfigure, the hysteresis approach results in the highest power consumption. The system with MPC consumes more power than that with the DRL-based control under various settings of weights, due to MPC's limited optimization horizon (10 control periods only). The above results show the superiority of the DRL-based control.

7 CONCLUSION

This paper developed an essential function for operating air free-cooled DCs in tropics – the control of the temperature and RH of the air supplied to the servers. It is based on an energy-efficient design of recirculating a controlled portion of return hot air to mix with the fresh outside air. This design leverages on the relaxed temperature upper limit of the latest servers and the heat generated by the servers to reduce the RH of the supply air. We formulated the control problem and proposed a DRL-based solution. Trace-driven simulations showed the effectiveness of the solution.

ACKNOWLEDGMENTS

This project is a collaboration between Info-communications Media Development Authority (IMDA) and Nanyang Technological University. This project is supported by the National Research Foundation (NRF), Prime Minister's Office, Singapore under its Green Data Centre Programme. Y. Wen's funded by NRF via the Green Data Centre Research (GDCR) administered by IMDA.

REFERENCES

[1] 2015. Data Center Industry Survey Results. https://bit.ly/2K1UDz4.
[2] 2018. Backblaze hard drive stats for 2018. https://bit.ly/2T8BiOa.
[3] ASHRAE TC 9.9. 2011. Thermal Guidelines for Data Processing Environments - Expanded Data Center Classes and Usage Guidance. *White Paper* (2011).
[4] ASHRAE. 2011. Gaseous and Particulate Contamination Guidelines for Data Centers. *White Paper* (2011).
[5] ASHRAE. 2013. Psychrometrics. In *2013 ASHRAE Handbook: Fundamentals (SI)*.
[6] Christy Sujatha D. and Satheesh Abimannan. 2011. Energy Efficient Free Cooling System for Data Centers. In *Proc. IEEE CloudCom*. 646–651.
[7] Íñigo Goiri, William Katsak, Kien Le, Thu D. Nguyen, and Ricardo Bianchini. 2013. Parasol and GreenSwitch: Managing Datacenters Powered by Renewable Energy. *SIGARCH Comput. Archit. News* 41, 1 (2013), 51–64.
[8] Íñigo Goiri, Thu D. Nguyen, and Ricardo Bianchini. 2015. CoolAir: Temperature- and Variation-Aware Management for Free-Cooled Datacenters. *SIGPLAN Not.* 50, 4 (March 2015), 253–265.
[9] Vani Gupta, Prashant Shenoy, and Ramesh K. Sitaraman. 2019. Combining Renewable Solar and Open Air Cooling for Greening Internet-Scale Distributed Networks. In *Proc. e-Energy '19*. 303–314.
[10] Hainan Zhang, Shuangquan Shao, Hongbo Xu, Huiming Zou, Changqing Tian. 2014. Free cooling of data centers: A review. *Renewable and Sustainable Energy Reviews* 35 (2014), 171–182.
[11] JEDEC. 2019. Arrhenius equation (for reliability). https://bit.ly/2ER6LiV.
[12] Will Knight. 2018. Google just gave control over data center cooling to an AI. https://bit.ly/2vbdLC4.
[13] Keith James Laidler and J Keith. 1965. *Chemical kinetics*. Vol. 15. McGraw-Hill.
[14] Kuei-Peng Lee and Hsiang-Lun Chen. 2013. Analysis of energy saving potential of air-side free cooling for data centers in worldwide climate zones. *Energy and Buildings* 64 (2013), 103 – 112.
[15] Timothy P. Lillicrap et al. 2016. Continuous control with deep reinforcement learning. In *Proc. 4th ICLR*. http://arxiv.org/abs/1509.02971
[16] Yingbo Liu, Duc Van Le, Rui Tan, Rongrong Wang, and Lek Heng Ngoh. 2019. Experiences with an Air Free-Cooled Data Center Testbed in the Tropics. In *Proc. BIM-WSN '19*.
[17] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, and P. Haves. 2012. Model Predictive Control for the Operation of Building Cooling Systems. *IEEE Trans. Control Syst. Technol.* 20, 3 (May 2012), 796–803.
[18] Ioannis Manousakis, Í nigo Goiri, Sriram Sankar, Thu D. Nguyen, and Ricardo Bianchini. 2016. CoolProvision: underprovisioning datacenter cooling. In *FAST*.
[19] Ioannis Manousakis, Sriram Sankar, Thu D. Nguyen, and Ricardo Bianchini. 2016. Environmental Conditions and Disk Reliability in Free-cooled Datacenters. In *Proc. of the 14th Usenix FAST*. 53–65.
[20] V Mnih et al. 2015. Human Level Control Through Deep Reinforcement Learning. *Nature* 518, 7540 (2015), 529–533.
[21] John Niemann, John Bean, and Victor Avelar. 2012. Economizer Modes of Data Center Cooling Systems. *Schneider Electric White Paper Library* (2012).
[22] Jay Park. [n.d.]. Designing a Very Efficient Data Center. https://bit.ly/2xBE7Oj.
[23] John Rath. 2011. Data Centers Can be Even Warmer. https://bit.ly/2ppu0sq.
[24] P Singh, Levente Klein, Dereje Agonafer, Jimil Shah, and Kanan Pujara. 2015. Effect of Relative Humidity, Temperature and Gaseous and Particulate Contaminations on Information Technology Equipment Reliability. In *Proc. InterPACK*.
[25] Richard S. Sutton and Andrew G. Barto. 1998. *Introduction to Reinforcement Learning* (1st ed.). MIT Press, Cambridge, MA, USA.
[26] Jan-Erik Svensson and Lars-Gunnar Johansson. 1993. A Laboratory Study of the Effect of Ozone, Nitrogen Dioxide, and Sulfur Dioxide on the Atmospheric Corrosion of Zinc. *Journal of The Electrochemical Society* 140 (1993), 2210–2216.
[27] T. Wei, , and Q. Zhu. 2017. Deep reinforcement learning for building HVAC control. In *Proc. ACM/EDAC/IEEE DAC*. 1–6.
[28] Deliang Yi, Xin Zhou, Yonggang Wen, and Rui Tan. 2019. Toward Efficient Compute-Intensive Job Allocation for Green Data Centers: A Deep Reinforcement Learning Approach. In *Proc. ICDCS*.
[29] Zhiang Zhang and Khee Poh Lam. 2018. Practical Implementation and Evaluation of Deep Reinforcement Learning Control for a Radiant Heating System. In *Proc. BuildSys '18*. 148–157.