energies

MDPI

Article

Artificial Intelligence-Based Temperature Twinning and Pre-Control for Data Center Airflow Organization

Na Huang [1,2], Xiang Li [1,2], Quanming Xu [3], Ronghao Chen [4], Huidong Chen [5] and Aidong Chen [1,2,6,*]

[1] Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China
[2] College of Robotics, Beijing Union University, Beijing 100101, China
[3] Vertiv Tech Co., Ltd., Shenzhen 518116, China
[4] College of Environment and Energy, Peking University Shenzhen Graduate School, Shenzhen 518055, China
[5] College of Urban Rail Transit and Logistics, Beijing Union University, Beijing 100101, China
[6] Research Center for Multi-Intelligent Systems, Beijing Union University, Beijing 100101, China
* Correspondence: aidongise@gmail.com

Abstract: Green and low-carbon has become the main theme of global energy development. Data centers are the core of the digital age, carrying huge arithmetic demand. Data centers must implement green low-carbon energy efficiency management to improve energy efficiency, reduce energy waste and carbon emissions, and achieve sustainable development. As a result, an intelligent management strategy for dynamic energy efficiency of data center networks with Artificial Intelligence (AI) fitting control is proposed. Firstly, a Long Short-Term Memory (LSTM) network is used for long sequence trend prediction to predict the temperature of the data center in the next sequence using the temperature of the past 15 sequences and the power consumption of the equipment as parameters. Then, based on the prediction results, the intelligent air conditioning controller based on Deep Q-Network (DQN) is designed to update the parameters by using the gradient of double-Q network and error backpropagation, and the optimal control action is selected by using the ε-greedy strategy to ensure that the prediction of the hotspot does not occur. Experiments show that the average absolute errors of temperature prediction for supply air, return air, cold aisle as well as hot aisle are 0.32 °C, 0.21 °C, 0.36 °C and 0.19 °C, respectively. The Power Usage Effectiveness (PUE) and Water Usage Effectiveness (WUE) decreased by an average of 2.6% and 2.5%, respectively. The method achieves the purpose of predicting future temperatures and intelligently controlling the output so that the data center can satisfy the premise of normal operation and thus achieve more efficient energy use.

Keywords: digital twin; temperature prediction; long and short-term memory networks; deep reinforcement learning

Citation: Huang, N.; Li, X.; Xu, Q.; Chen, R.; Chen, H.; Chen, A. Artificial Intelligence-Based Temperature Twinning and Pre-Control for Data Center Airflow Organisation. *Energies* **2023**, *16*, 6063. https://doi.org/10.3390/en16166063

Academic Editor: Ramesh Rayudu

Received: 22 July 2023
Revised: 8 August 2023
Accepted: 15 August 2023
Published: 18 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the increasing prominence of economic development and natural resources and environmental conflicts, green and low-carbon development has become an important trend in global development, with deepening integration with all areas of human society, economy and politics [1–5]. From a global perspective, countries are strengthening cooperation on environmental protection, setting stricter carbon emission targets, and promoting a shift to renewable energy and energy-efficient technologies. Enterprises are also integrating green concepts into their strategic planning, increasing environmental protection investment, and promoting the development of green industries [6].

Data centers are recognized as a high energy-consuming industry. "2023 Data Center Green Development Conference" data from the Ministry of Industry and Information Technology of the People's Republic of China shows that by the end of 2022, the total size

of China's in-use data centers exceeded 6.5 million standard racks, with an average growth rate of more than 25% over the past five years; 2022 Chat Generative Pre-trained Transformer (ChatGPT) opens the Artificial Intelligence Generated Content (AIGC), a new industry, pushing Artificial Intelligence (AI) development into the AI2.0 era featuring multimodal and large models, and pushing the construction of smart computing facilities into a new stage. The popularity of the Internet, the rapid development of mobile applications and the popularity of social media have led to the generation of a large amount of digital data in people's daily lives. The rise of the Internet of Things (IoT) and AI technologies has also put tremendous pressure on power demand. The rapid development of AI technologies requires massive computing resources and energy support. Many innovative technologies and solutions are emerging to meet these exploding computing and power demands. Nowadays, one of the focuses of research efforts is to transform data centers into green data centers to reduce energy consumption [7,8]. Data center power-consuming facilities include servers and computing equipment, network equipment, Refrigeration system, Uninterruptible Power Supply (UPS) systems, and equipment support systems. Among them, servers and computing equipment, refrigeration systems are the most important power-consuming facilities [9]. Many enterprises and scholars at home and abroad have conducted extensive surveys to study the proportion of energy consumption components of data centers, and although the results of different studies may differ, the overall energy consumption components and ranking are basically similar. Figure 1 shows that the energy consumption of servers and computing equipment as well as the energy consumption of refrigeration systems is the largest.

Therefore, efficient design and optimization of data centers are essential, and predicting the energy consumption of servers and computing devices is one of the keys to achieving this goal. By predicting server energy consumption, data centers can provide more reference information for intelligent tasking to achieve optimal energy usage; better assess and budget future power usage to improve economic efficiency; identify specific improvement measures to increase energy efficiency and assess the impact of energy consumption on the environment and take appropriate measures to reduce carbon emissions and other environmental impacts, and better fulfill our responsibility for sustainable development.

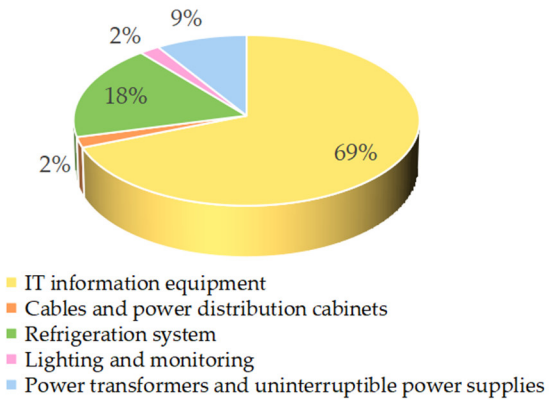


Figure 1. Percentage of energy consumption in data center.

At present, most refrigeration systems adopt the basic control mode, limited by the interface capacity of the control board, unable to access a large number of sensors, the distribution of sensors cannot meet the requirements of the precise collection of airflow temperature and generally use the empirical value to set the relevant parameters. In addition, the controller in the refrigeration unit adopts the traditional control algorithm, which lacks the intelligence of machine learning and cannot achieve the purpose of precise control. Especially in the same data center server room, multiple refrigeration units "go their own way" and lack the intelligence of group control, which results in a waste of

refrigeration resources and increases the overall energy consumption of the refrigeration system. AI control method can solve the above problems, unlike the basic control method, the latter adopts the intelligent collector in addition to collecting a large number of temperature data but also collects the energy consumption data of refrigeration equipment, power supply and distribution equipment, and IT information equipment for background management [10–15]. The AI control approach is based on AI technology, combining physical a priori, big data and IoT technology, with the help of historical data, real-time data, and algorithmic models, etc., to predict potential risks, optimize resource allocation, and achieve the purpose of predicting temperatures, intelligent management, as well as reducing energy consumption[16–19]. As a result, an intelligent management strategy for dynamic energy efficiency of data center networks with AI fitting control is proposed.

The main contributions include three aspects:

1. By normalizing the temperature and power consumption data of the data center room, using the log-mean-square error and the coefficient of determination as the evaluation indexes for the effect of the prediction model, and adjusting the network parameters through the indexes to obtain the optimal neural network prediction model.
2. A deep Q-network algorithm-based intelligent controller for air-conditioning is designed for air-conditioning control. The state input network is constructed through indoor environment indicators, feedback and control actions, and the parameters are updated using error backpropagation and the optimal actions are selected to achieve intelligent control.
3. According to the distribution of power consumption of information equipment load and partition calculation of the refrigeration module, the output of refrigeration equipment in a certain range is intelligently controlled to reduce the risk of local overheating and achieve effective group control operation.

The remainder of the article is organized as follows. Section 2 reviews and presents the background knowledge of data center cooling systems. Section 3 details the design and application of the proposed digital twin (simulation) module. Section 4 details the design and application of the proposed hotspot and fault pre-control module, as well as the group control optimization control module. Section 5 validates the proposed methodology of the article through experiments. Section 6 concludes the full paper.

2. Materials and Methods

The data center cooling system consists of underlying intelligent measurement devices and collectors that form the collection units of the subsystems, as shown in Figure 2. These units collect a large amount of decentralized data, which are transmitted via Ethernet to the DS distribution server and the DB database server. The servers process historical and real-time data, calibrate, clean, classify, and generate different databases. The AI algorithm server uses deep learning to mine, learn, and compute data to compare temperatures, controls, and events to provide optimized control algorithms. The Control Visualizer runs on the Windows operating system and is supported by B/S (SiteWeb6.0) and C/S (SiteMonitor 2.0) software architectures to present analysis curves, charts, and reports. The Data Center Infrastructure Management Platform uses Computational Fluid Dynamics Simulation (CFD) (6SigmaRoom, Release 16.2) and Building Information Modeling (BIM) software (AUTODESK Revitit 2020) to establish a cooling system model and apply neural networks to pre-control hot spots and faults and realize optimized group control of air conditioners.

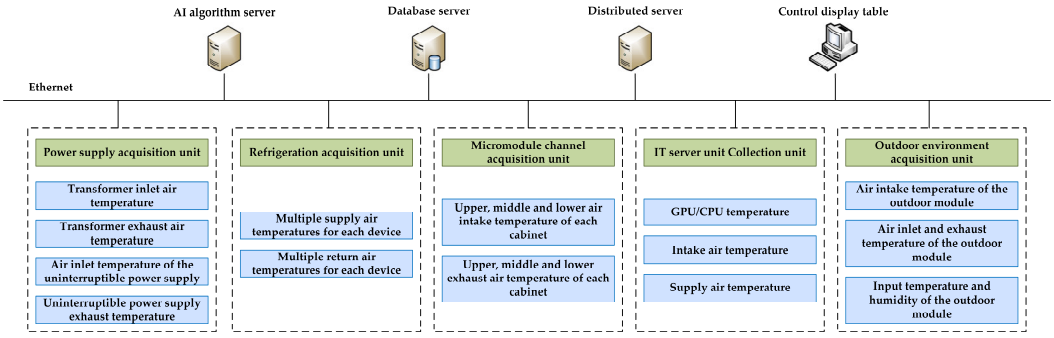


Figure 2. Network architecture for data center cooling systems.

The construction of the main module of the data center cooling system mainly includes the design of the digital twin (simulation) module, the hot spot and pre-failure control module and the group control optimization and control module. As shown in Figure 3.

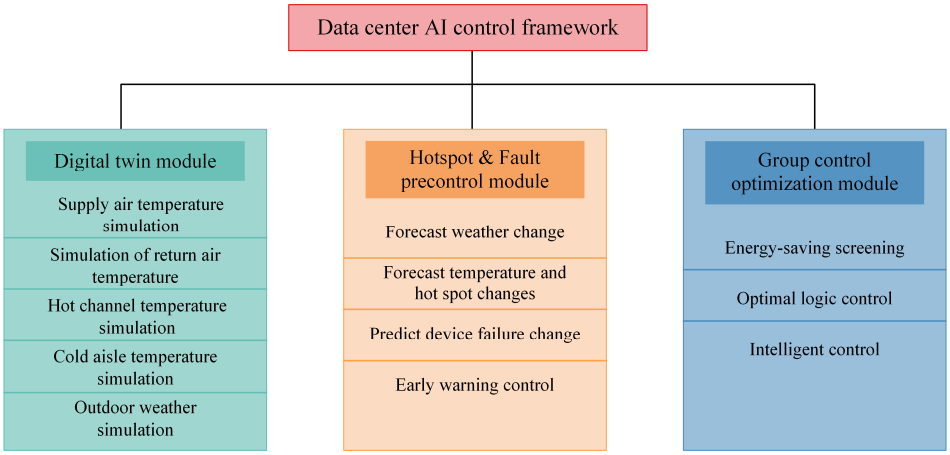


Figure 3. Data center cooling system main module components.

3. Digital Twin (Simulation) Module

3.1. Application of Digital Twin Technology in Data Center Cooling Systems

Digital twin is a technology that generates a visual twin model by digitally and accurately mirroring the physical system to the virtual world through sensors, IoT and other technologies [20,21]. The digital twin model of the data center cooling system is constructed to achieve two-way interactive feedback and iterative operation of information between the virtual and real systems so that the twin system model completely presents the operating state of the cooling system in the digital space. The airflow and temperature fields are then simulated. Moreover, the outdoor air temperature is simulated so that variables such as data center temperature can be simulated, verified and predicted with the help of historical data, real-time data and AI algorithms, which provide the basis for decision-making and control behavior of subsequent physical refrigeration objects throughout their life cycle.

3.2. Time Series Based Temperature Prediction Model

Recurrent Neural Network (RNN) is a neural network model for processing sequence data and time series data. A sequence is a collection of data arranged in a certain order. The data are arranged in their chronological or spatial order and there is a dependency between the data points, i.e., each data point is dependent on the information in the previous data points. It has a recurrent connection structure that models contextual information in sequences, capturing temporal correlations and contextual dependencies [22–24]. Long Short-Term Memory (LSTM) is a variant of recurrent neural networks [25], and the

network structure is shown in Figure 4. Compared with the traditional RNN, LSTM gives the network a stronger memory capability by introducing the "gate" mechanism, which is a learnable mechanism that can selectively save, update, or output information in the network. This allows LSTM to capture and exploit long-term dependencies in sequences more effectively without being affected by gradient vanishing or gradient explosion.

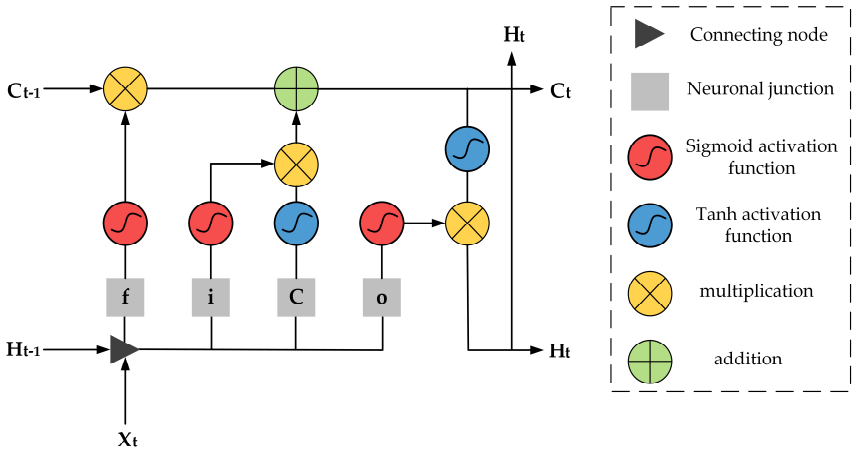


Figure 4. Long short-term memory network structure.

Represented in grey in Figure 4 are the neuron nodes, which are, in order, the Forget Gate, denoted as f, which decides whether to remove previously-stored information from memory; the Input Gate, denoted as i, which decides whether to incorporate the current input into memory; the cell state, denoted as C; and the Output Gate, denoted as o, which decides the current moment's output of the LSTM cell state. The individual variables are calculated as in Equations (1)–(6).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \cdot tanh(C_t) \tag{6}$$

h_{t-1} is the previous output, x_t is the current input vector, f_t, i_t are the output vectors, $\tilde{C}_t$ is the vector of candidate values W_f, b_f, W_i, W_C, W_o, b_i, b_C, and b_o are the learning parameters, and h_t is the vector of hidden units. σ is the Sigmoid activation function and $tanh$ is the Tanh activation function.

The flowchart of the neural network algorithm for data center temperature prediction is shown in Figure 5. According to the data acquisition of each collection unit, the 15-min time series is taken to divide the data, determine the input variables as the indoor temperature and power consumption of the machine room in the first 15 time series, and the output as the temperature data in the next time series, determine the topology of the neural network, and assign the optimal value to the LSTM by calculating the error of the hidden layer, solving for the gradient of the error as well as adjusting the weights and thresholds according to the learning function, and giving the optimal value to the initial weights and thresholds in the network. During the neural network training process, the error

between the real value of the actual result and the predicted value is used as the basis for judging whether the training is finished or not. Through iterative optimization and reinforcement learning of the training dataset, the parameters of the neural network are continuously adjusted so that the error is gradually reduced to reach the preset accuracy requirements. Once the error reaches the preset accuracy, the training is stopped and the final result is output, thus ensuring that the network achieves a high level of accuracy and performance in handling the task. The dataset is trained with repeated optimization and reinforcement learning until the error reaches the preset accuracy, the training is stopped and the results are output. Based on the available experimental data, the hidden layer of the neural network was trained by taking 22 neurons and fitted using a model with layers 1 to 5.

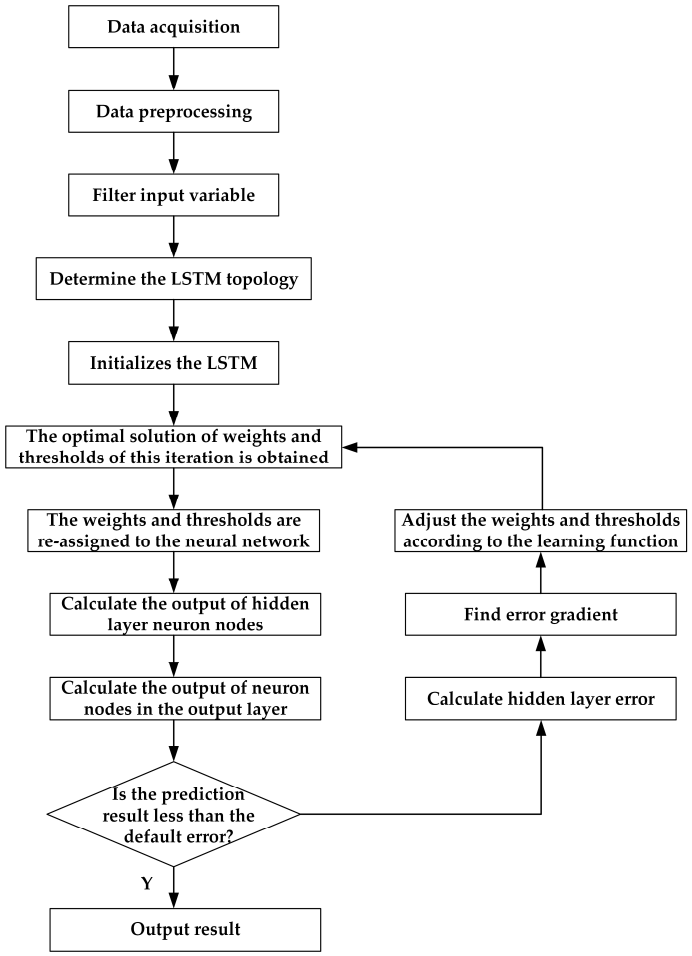


Figure 5. Flowchart of data center temperature prediction based on neural network.

4. Optimal Control Module: Hotspot and Fault Pre-Control and Group Control

4.1. Hot Spots and Fault Pre-Control

Deep Q-Network (DQN) is a deep neural network structure used to implement the Deep Q-Learning algorithm. DQN was proposed by DeepMind in 2013 [26], which solves the traditional Q-Learning challenges in dealing with high-dimensional state space by combining deep learning and Q-Learning. challenges when dealing with high-dimensional state spaces. DQN uses deep neural networks as estimators of the Q-value function and is able to effectively deal with high-dimensional state spaces deal with high-dimensional state spaces. The target network is used to compute the target Q-value, and the target network is separated from the main network, which reduces the volatility of the target value in training. DQN introduces the Experience Replay technique, which solves the problems of data correlation and non-static distribution and improves the efficiency

and stability of training by storing the experience of the agent's interactions with the environment in the buffer and randomly sampling the samples for training. The DQN algorithm pseudo-code is shown in Algorithm 1.

Algorithm 1: Introducing deep Q-networks for experience playback.

Input: State space S, Action space A, Discount rate γ, Learning rate α
Output: Q network $Q_\varphi(s,a)$
Initialise the experience playback buffer D with capacity N;
Initialise the Q-network with parameters φ;
Initialise the parameters of the target Q-network $\hat{\varphi} = \varphi$;
repeat
  Initialise the start state s;
  repeat
    In state s, the choice action $\alpha = \pi^{\epsilon}$;
    Perform action α, observe the environment, and get even the reward r and the new state s';
    Putting s, α, r, s' into D;
    Sample $ss, \alpha\alpha, rr, ss'$ from D;
    $y = \begin{cases} rr, & ss' \text{ is the termination state,} \\ rr + \gamma max_{\alpha'} Q_{\hat{\varphi}}(ss', a'), & \text{or else;} \end{cases}$
    Train the Q network with $\left(y - Q_\varphi(s,a)\right)^2$ as the loss function;
    s' to s;
    Every C steps, φ to $\hat{\varphi}$;
  until s is the termination state;
until $\forall s, a$, $Q_\varphi(s,a)$ convergence;

The control flowchart of the air conditioner intelligent controller based on the DQN algorithm for air conditioners is shown in Figure 6. The controller takes two consecutive time-series of indoor environmental indicators, the feedback of the previous time-series, and the control action of the previous time-series as the state input networks, constructs the error gradient of error back-propagation through the two Q-networks, updates the parameters of the Q-networks through the back-gradient propagation algorithm, and uses the Q-networks for prediction and selects the optimal output action under the current state and outputs the final control action using the ε-greedy (epsilon-greedy) algorithm to output the final control action.

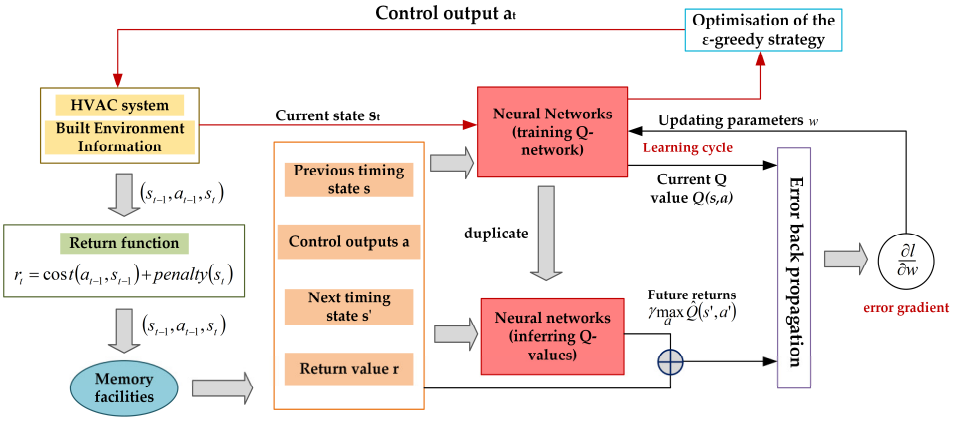


Figure 6. Intelligent control process for air conditioning based on DQN algorithm.

Similarly, when the refrigeration system has a fault, according to the established algorithmic control strategy, the maximum degree of opening the fault pre-control module

ensures that the safe operation of the data center is not affected. The data center is generally equipped with several refrigeration systems as standby systems, which are normally in the state of hot standby operation in order to save energy, and the system automatically notifies the relevant personnel to arrive at the scene at the same time, and according to the results of the AI algorithm, the system automatically adjusts to increase the refrigeration capacity of standby refrigeration equipment in its area to ensure that no overheating point occurs. After the elimination of faults, periodic tracking judgment, after the no-fault point appears, and restores the normal control logic.

4.2. Optimised Control Module for Group Control

According to the distribution of load and power consumption of the information equipment, under the premise of meeting the design requirements, the average temperature prediction model and power consumption prediction model are used to exclude dangerous and energy-consuming instructions from the AI instruction set, and then the deep reinforcement learning algorithm is used to select the optimal instructions to be issued, with priority to energy efficiency as the goal, taking into account safety, and continuous optimization.

The AI group control module collects the demand and temperature of the refrigeration equipment in group control mode and sends down the average demand after calculating according to the refrigeration group control module, which improves the temperature uniformity inside and outside the channel and enhances the overall energy efficiency. Based on the collected temperature data, it intelligently adjusts the output of the refrigeration equipment within a certain range to reduce the risk of local overheating. Realize independent control of different channels, adjusting the output of its own devices according to temperature changes, reducing energy consumption and hot spot frequency. Under the balanced thermal load, the airflow is the same, and the cooling capacity is output on demand, maximizing the use of natural cooling sources. As the outdoor temperature rises, it automatically allocates natural/traditional cooling sources to meet the equipment load demand and realize energy saving and emission reduction.

5. Experiments

5.1. Data Center Simulation Platform

The experimental site machine room has a floor height of 5.4 m and an anti-static floor height of 1.0 m. There are five micro-modules in the equipment area, 40 cabinets in each micro-module, a total of 200 cabinets, and the height of the cabinets is 2.2 m. Each module is arranged on both sides, with 20 cabinets on each side. The power consumption of a single cabinet load is 5 KW, and the equipment area adopts anti-static flooring with an air supply underneath and cable routing on top. The cabinet layout for hot and cold channel design, cold channel isolation; the channel with open floor air supply, and floor through the hole rate of 70%. The air conditioning area is separated from the main equipment area by a wall, and the air supply and return ports are opened on the upper and lower part of the partition wall. The air conditioners are distributed in the air conditioning rooms on the left and right sides, and seven air conditioners with 100 KW cooling capacity are installed in each air conditioning room, which adopts the mode of five mains and two standbys. Air conditioning rated air volume of 25,000 m^3/h, using air temperature control, set to 26 degrees.

The CFD simulation software for this experiment is 6SigmaDCX (6SigmaRoom, Release 16.2), an authoritative software in the data center industry. Based on the finite element method, 6SigmaDCX realizes the simulation of real physical phenomena by solving partial differential equations (single field) or partial differential equations (multi-field) and uses mathematical methods to solve physical phenomena. Before modeling and solving the data center, the following simplified assumptions are proposed for the model to

improve the calculation speed while ensuring the calculation accuracy. The boundary conditions are as follows:

1. The gas flow state is regarded as steady turbulence;
2. the heat of the cabinet does not change with time, is a constant value;
3. Ignore the heat radiation of solid walls and indoor surfaces and the heat transfer of the enclosure structure;
4. Ignore obstacles such as cables and pipelines in the machine room;
5. Ignore the human body heat dissipation and lighting heat dissipation.

The simplified model has a small deviation from the actual situation of the computer room and has little influence on the CFD simulation of the data center. The error between the simulation results and the design situation is within an acceptable range. The simplified model reduces the calculation amount and greatly improves the calculation speed. This paper focuses on the setting of the roof, wall, floor, air conditioning and IT import and export parameters. The purpose is to determine the rationality of sensors and other layouts and the uniformity of air supply through CFD simulation, so as to make the AI algorithm more accurate. The computer room simulation diagram drawn by CFD software is shown in Figure 7.

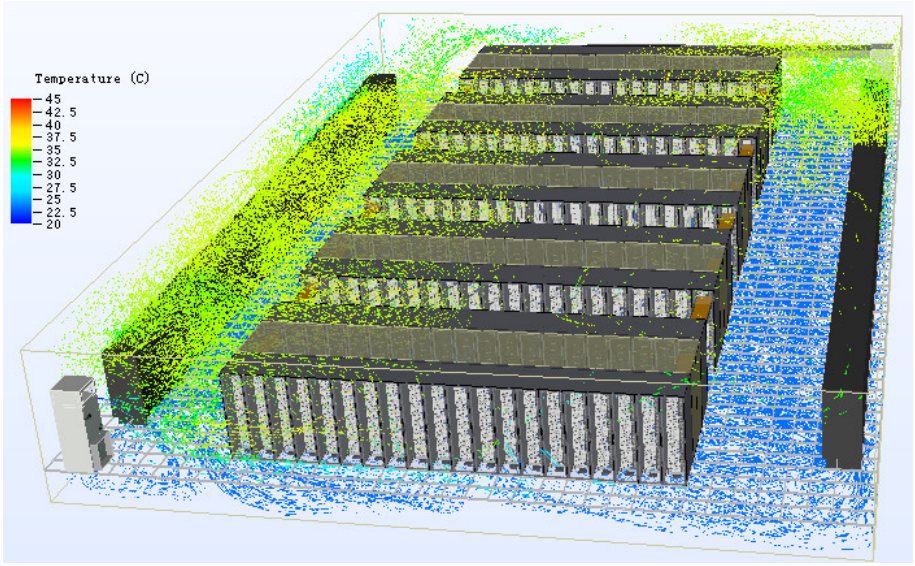


Figure 7. Server Room Simulation.

5.2. Experimental Data

The data used in this experiment came from the monitoring platform of a data room in Shenzhen. The data from 10 July 2021 to 10 May 2022 are used as training samples and provided to the algorithm model for training. Data from 10 May 2022 to 10 July 2022 are used as an estimated sample to verify the effect of the algorithm model. The reliability and landability of the algorithm are further verified by training and testing with real data from different time periods.

The data room is equipped with 728 data sampling points (sensors), and each data sampling point (sensor) sampling frequency is 15 min, the sensor technical specifications and description are shown in Table 1. Data types include outdoor meteorological parameters, air temperature of the cooling outdoor unit, air temperature of the cooling indoor unit, air temperature of the cooling indoor unit, cold aisle temperature, hot aisle temperature, cabinet load, cooling power consumption, and lighting power consumption in the equipment room. In order to avoid the interference of abnormal and missing data reception, it is necessary to preprocess and screen reasonable data before neural network prediction. Missing data are a common and important problem. Missing data can be caused by measurement equipment failure, incomplete data collection, or human error. One of

the methods to solve the problem of missing data is interpolation. The data set uses spline interpolation, which builds a continuous interpolation function by fitting multiple low-order polynomial segments between the data points. Spline interpolation has good robustness and smoothness, which can effectively fill in missing data and avoid overfitting problems.

Table 1. Sensor technical specifications and description.

Parameter	Value
Operating voltage	9 Vdc~28 Vdc
Measuring temperature range	−10 °C~+50 °C, accuracy ±0.5 °C
Energy meter accuracy	Classes 1
Power meter accuracy	Classes 0.5

All above meet the requirements of the national standard.

5.3. Experimental Setup

This experiment was conducted under Mint 20, using Python version 3.9.0. The experiment was conducted on a workstation with Intel(R) Xeon(R) Gold 5218 CPU @ 2.30 GHz, 128 G memory and NVIDIA GeForce GTX 2080 Ti graphics card. The model parameter settings as shown in Table 2 are used in this experiment. Adjusting the settings of the model parameters is a way to improve the model's performance and learning ability so that it performs better in facing the challenges in the network topology environment. Through iterative interactive training, the understanding and learning of the environment are used to optimize the selection of the distance between the task scheduling server and the data center to achieve the optimal scheduling strategy.

Table 2. Model parameter settings.

Parameter	Value
Epoch	9000
Discount factor	0.99
Learning rate	0.001
Batch size	128
Number of hidden layers	5
every T episodes	30
ε initial value	0.2
ε minimum	0.001
ε decay rate	0.0000199

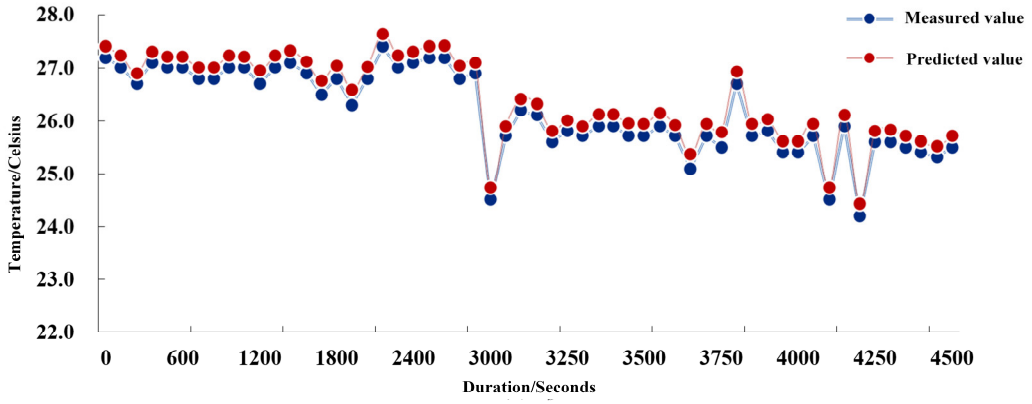
5.4. Experiments and Analysis of Data Center Task Prediction

As shown in the previous section, the prediction accuracy of this experiment is presented using the log root mean square error (log-RMSE) and coefficient of determination (R^2) evaluation metrics, with smaller log-RMSE and higher R^2 values indicating higher prediction accuracy. The relative errors of layers 1 to 5 in Table 3 are small and reach the prediction target. The log-RMSE in the table all reach below 0.36, which indicates higher accuracy. As well, the R^2 in the table all reach above 0.97, indicating a better fit.

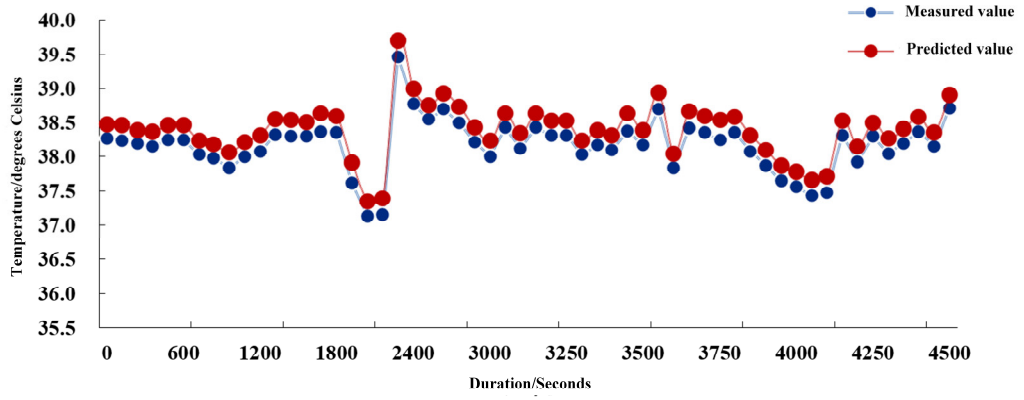
Table 3. log-RMSE and R^2 for different layers of neural networks.

Number of Layers	Log-RMSE	R^2
1	0.0333	0.9747
2	0.0267	0.9843
3	0.0301	0.9804
4	0.0314	0.9822
5	0.0356	0.9810

The average values of the supply and return air temperatures were collected and calculated for a 10-min cycle of the air conditioner. From Figure 8a, it can be seen that the air-conditioning supply air set temperature range is 26 °C ± 1 °C. Figure 8b shows that the air conditioning return air setting temperature range is 39 °C ± 1 °C. In order to obtain a lower PUE while guaranteeing the normal operation of the IT equipment, its return air temperature as well as the temperature difference between the supply and return air are raised as much as possible to ensure a more energy-efficient cooling system. The error between the real value and the predicted value should be no more than ±0.4 °C. The method obtains an average absolute error of 0.32 °C in the prediction of the supply air temperature and 0.21 °C in the prediction of the return air temperature, which proves that the method can more accurately predict the future temperature of the supply air and the return air temperature.



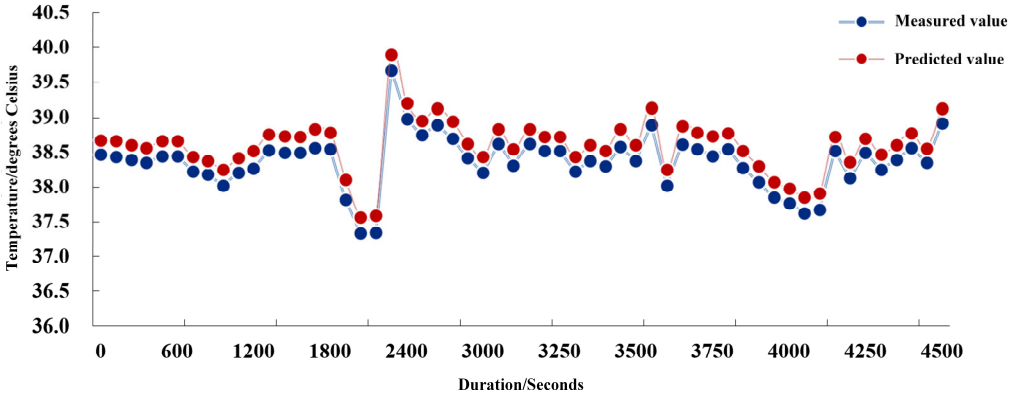
(**a**)



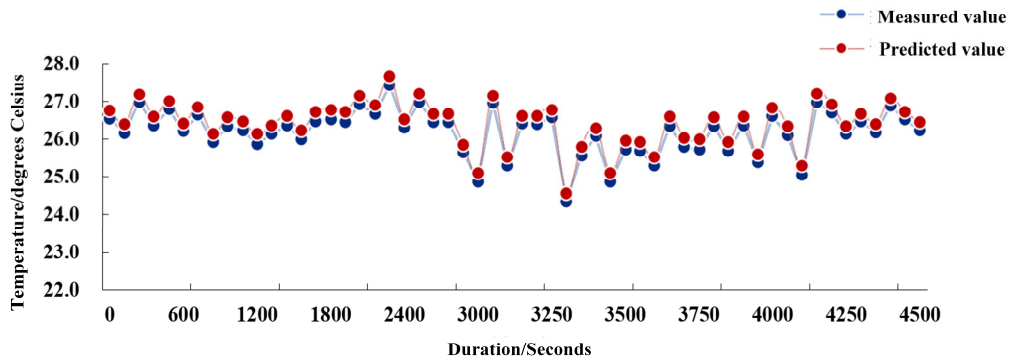
(**b**)

Figure 8. Comparison of predicted and real air-conditioning temperature trends. (**a**) Air supply temperature of air conditioner, (**b**) Air conditioning return air temperature.

The average of the server's 10-min hot-channel temperature and cold-channel temperature was collected and calculated. As can be seen in Figure 9a, the hot channel set temperature range is 39 °C ± 1 °C. From Figure 9b, it can be seen that the cold channel set temperature range is 27 °C ± 1 °C. In the case of guaranteeing normal operation of IT equipment, in order to obtain a lower PUE, its temperature is raised as much as possible to ensure the higher energy efficiency of the cooling system. Under the circumstance of guaranteeing the normal operation of IT equipment, in order to obtain lower PUE, the temperature difference between the cold channel and the hot channel should be made as small as possible to ensure the higher energy efficiency of the refrigeration system. The error between the true value and the predicted value should be no more than ±0.4 °C. The method obtains an average absolute error of 0.36 °C for the cold channel temperature prediction and 0.19 °C for the hot channel temperature prediction, which proves that the method can more accurately predict the future hot channel temperature and cold channel temperature.



(**a**)



(**b**)

Figure 9. Comparison of predicted and real temperature trends. (**a**) Hot channel temperature, (**b**) Cold channel temperature.

Figure 10 shows the comparison curve between the real and predicted values of some of the IT load variables derived by the algorithm, as the average of the load change of the server after 10 min, where the red curve represents the trend of the IT load (power) over time (per second) predicted by the neural network algorithm, and the blue curve represents the trend of the measured change. From the figure, it can be seen that the error between the real and predicted values of the 200 cabinets in the server room is within ±2 KW, which satisfies the error range of the algorithm and proves that the method can predict the IT power more accurately.

The red curve in Figure 11 represents the theoretically calculated temperature trend over time (hourly) and the blue curve represents the real-time outdoor temperature trend over time (hourly) for the month of August as measured. It can be seen that as human activities continue to intensify, the actual outdoor temperature is in an upward trend, which poses a greater challenge to energy conservation, so it is necessary to reduce the ambient temperature of its surroundings through spraying.

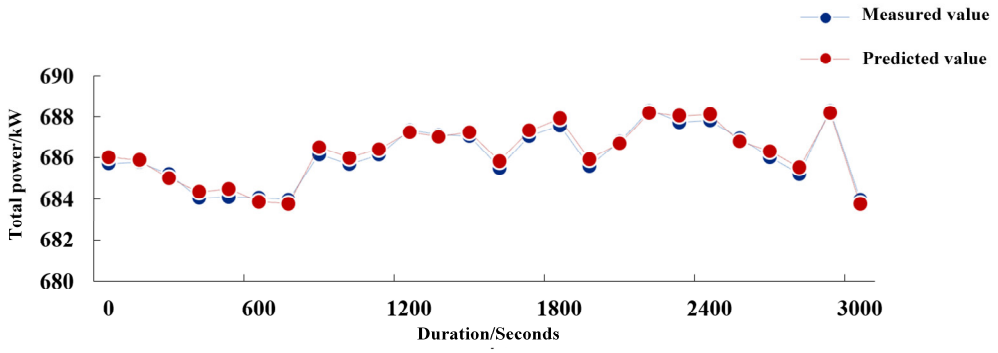


Figure 10. IT load trend of the equipment room.

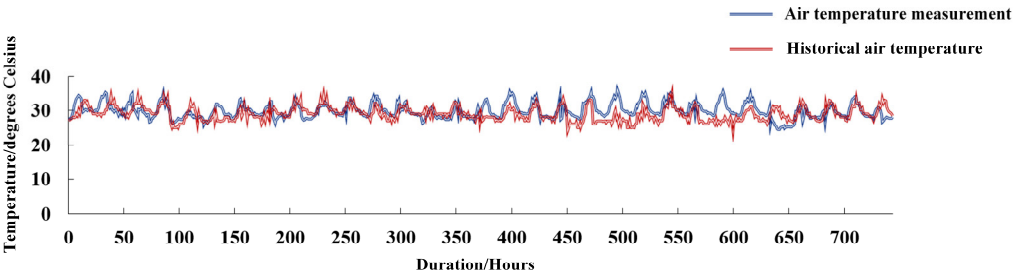
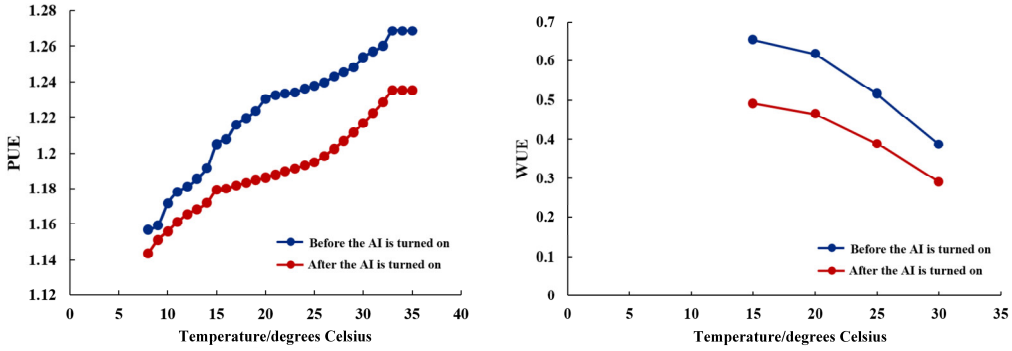


Figure 11. Trends in measured weather temperatures and theoretically calculated weather changes.

Power Usage Effectiveness (PUE) is an energy efficiency metric used to measure data centers, and Water Usage Effectiveness (WUE) is a water usage metric used to measure data centers. Lower PUE and WUE values indicate higher energy efficiency in the data center, both of which are as close to 1.0 as possible, enabling more efficient use of energy resources. The application of AI control of the data center ensures that a lower ambient temperature near the entrance of the cooling outdoor cooling module is obtained, guaranteeing that a lower PUE can be obtained even in a high-temperature environment. The red curve in Figure 12a represents the PUE trend with ambient temperature predicted by the neural network algorithm, and the blue curve represents the measured PUE trend with ambient temperature. As seen in the figure, under the premise of lower outdoor weather temperatures, the PUE can be close to about 1.1, and the highest PUE does not exceed 1.3. Comparative analysis before and after the AI is switched on shows that the energy-saving PUE is predicted to drop by an average of 2.6% for the whole year.



(**a**) (**b**)

Figure 12. Trend of energy-saving indicators before and after AI is switched on. (**a**) Trends in PUE, (**b**) Trends in WUE.

Through the application of AI control in the data center, according to the actual temperature of the outdoor weather and the wet bulb temperature of the automatic adjustment, such as more than 30 degrees to open the spray mode; more than 25 degrees to open the spray mode, and at the same time the difference between the dry and wet bulb temperature is greater than 3 degrees; more than 20 degrees to open the spray mode, and at the same time the difference between the dry and wet bulb temperature is greater than 3 degrees; more than 15 degrees to open the spray mode so that the refrigeration system is in the natural cold mode earlier and other measures to achieve the goal of reducing the ambient temperature near the entrance of the refrigeration outdoor cooling module and water saving and maximize water saving under the premise of guaranteeing the cooling of the refrigeration cooling module. The red curve in Figure 12b represents the trend of WUE metrics with ambient temperature change by the neural network algorithm, and the blue curve represents the measured trend of WUE metrics with ambient temperature change. As seen in the figure, AI opens the comparative analysis to achieve an average decrease of 2.5% in the data center water-saving WUE.

6. Conclusions

With the development of data centers towards large scale and high density, the problem of energy consumption is becoming more and more prominent. In order to reduce energy consumption and environmental impact, green data centers have become the focus of attention. This paper proposes an intelligent management strategy for energy efficiency in data centers with AI fitting control. Through basic control and model training, the temperature trend is predicted by using a temporal neural network and DQN, and the experiments show that the average absolute error of all kinds of temperature prediction is below 0.4, and the power usage effectiveness and water usage effectiveness are reduced by 2.55% on average, which realizes intelligent control to optimize energy consumption. The strategy can effectively cope with the trend of large-scale and high-density development, provide a sustainable solution for data center operators and researchers, and make an important contribution to the green transformation of the data center industry and environmental protection. The shortcoming of the research in this paper is that the object is limited by the building structure cannot be modified, the selected refrigeration and air conditioning supply type is simple, and only AI verification is done for the main room. In addition, the experimental data set is the static energy consumption data set collected. In real scenarios, user tasks change dynamically, and the characteristics of energy consumption will also change accordingly. Therefore, it is necessary to further consider and study how to conduct intelligent energy efficiency management for data centers in real dynamic scenarios. The future work plans to extend this method to air-conditioning systems, power supply and distribution systems, lighting, and even IT (that is, information equipment, such as servers, storage, networks and other computing infrastructure) and other systems in

the distribution room to further improve the efficiency of the entire data center, while reducing the overall energy consumption index, so as to achieve the dual control goal of carbon emissions.

Author Contributions: Conceptualization, X.L. and Q.X.; methodology, N.H. and R.C.; software, X.L. and R.C.; validation, N.H. and Q.X.; formal analysis, Q.X.; investigation, N.H. and H.C.; resources, Q.X. and H.C.; writing—original draft preparation, N.H., X.L. and R.C.; writing—review and editing, N.H., X.L. and R.C.; visualization, X.L. and H.C.; supervision, N.H. and A.C.; project administration, N.H. and A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ullah, S.; Luo, R.; Nadeem, M.; Cifuentes, F, J. Advancing sustainable growth and energy transition in the United States through the lens of green energy innovations, natural resources and environmental policy. *Resour. Policy* **2023**, *85*, 103848.
2. Chenic, A.Ș.; Cretu, A.I.; Burlacu, A.; Moroianu, N.; Virjan, D.; Huru, D.; Stanef-Puica, M.R.; Enachescu, V. Logical analysis on the strategy for a sustainable transition of the world to green energy—2050. Smart cities and villages coupled to renewable energy sources with low carbon footprint. *Sustainability* **2022**, *14*, 8622.
3. Zhou, Y. Low-carbon transition in smart city with sustainable airport energy ecosystems and hydrogen-based renewable-grid-storage-flexibility. *Energy Rev.* **2022**, *1*, 100001.
4. Zhang, J.; Lyu, Y.; Li, Y.; Geng, Y. Digital economy: An innovation driving factor for low-carbon development. *Environ. Impact Assess. Rev.* **2022**, *96*, 106821.
5. Delgado-Alvarado, E.; Elvira-Hernandez, E.; Hernandez-Hernandez, J.; Huerta-Chua, J.; Vazquez-Leal, H.; Martinez-Castillo, J.; Garcia-Ramirez, P.; Herrera-May, A. Recent progress of nanogenerators for green energy harvesting: Performance, applications, and challenges. *Nanomaterials* **2022**, *12*, 2549.
6. Wu, Z.; Huang, X.; Chen, R.; Mao, X.; Qi, X. The United States and China on the paths and policies to carbon neutrality. *J. Environ. Manag.* **2022**, *320*, 115785.
7. Katal, A.; Dahiya, S.; Choudhury, T. Energy efficiency in cloud computing data centers: A survey on software technologies. *Clust. Comput.* **2022**, *26*, 1845–1875.
8. Bharany, S.; Sharma, S.; Khalaf, O, I.; Abdulsahib, G, M.; Al Humaimeedy, A.; Aldhyani, T.; Maashi, M.; Alkahtani, H. A systematic survey on energy-efficient techniques in sustainable cloud computing. *Sustainability* **2022**, *14*, 6256.
9. Zhang, X.; Lindberg, T.; Xiong, N.; Vyatkin, V.; Mousavi, A. Cooling energy consumption investigation of data center it room with vertical placed server. *Energy Procedia* **2017**, *105*, 2047–2052.
10. Tong, X.; Wang, J.; Liu, W.; SAMAH, H.; Zhang, Q.; Zhang, L. A time-varying state-space model for real-time temperature predictions in rack-based cooling data centers. *Appl. Therm. Eng.* **2023**, *230*, 120737.
11. Jin, Shu.; Li, N.; Bai, F.; Chen, Yu.; Feng, X.; Li, H.; Gong, X.; Tao, W. Data-driven model reduction for fast temperature prediction in a multi-variable data center. *Int. Commun. Heat Mass Transf.* **2023**, *142*, 106645.
12. Athavale, J.; Yoda, M.; Joshi, Y. Comparison of data driven modeling approaches for temperature prediction in data centers. *Int. J. Heat Mass Transf.* **2019**, *135*, 1039–1052.
13. Zhang, Z.; Zeng, Y.; Liu, H.; Zhao, C.; Wang, F.; Chen, Y. Smart DC: An AI and digital twin-based energy-saving solution for data centers. In Proceedings of the NOMS 2022—2022 IEEE/IFIP Network Operations and Management Symposium, Budapest, Hungary, 25–29 April 2022; pp. 1–6.
14. Yang, Z.; Du, J.; Lin, Y.; Du, Z.; Xia, L.; Zhao, Q.; Guan, X. Increasing the energy efficiency of a data center based on machine learning. *J. Ind. Ecol.* **2022**, *26*, 323–335.
15. Koot, M.; Wijnhoven, F. Usage impact on data center electricity needs: A system dynamic forecasting model. *Appl. Energy* **2021**, *291*, 116798.
16. Liu, Y.; Wei, X.; Xiao, J.; Liu, Z.; Xu, Y.; Tian, Y. Energy consumption and emission mitigation prediction based on data center traffic and PUE for global data centers. *Glob. Energy Interconnect.* **2020**, *3*, 272–282.
17. Li, Y.; Wen, Y.; Tao, D.; Guan, K. Transforming cooling optimization for green data center via deep reinforcement learning. *IEEE Trans. Cybern.* **2019**, *50*, 2002–2013.
18. Karthiban, K.; Raj, J. An efficient green computing fair resource allocation in cloud computing using modified deep reinforcement learning algorithm. *Soft Comput.* **2020**, *24*, 14933–14942.
19. Sun, P.; Guo, Z.; Liu, S.; Lan, J.; Wang, J.; Hu, Y. SmartFCT: Improving power-efficiency for data center networks with deep reinforcement learning. *Comput. Netw.* **2020**, *179*, 107255.
20. Attaran, M.; Celik, B.G. Digital Twin: Benefits, use cases, challenges, and opportunities. *Decis. Anal. J.* **2023**, *6*, 100165.

21. Liu, X.; Jiang, D.; Tao, B.; Xiang, F.; Jiang, G.; Sun, Y.; Kong, J.; Li, G. A systematic review of digital twin about physical entities, virtual models, twin data, and applications. *Adv. Eng. Inform.* **2023**, *55*, 101876.
22. Jordan, M.; Serial order: A parallel distributed processing approach. *Adv. Psychol.* **1997**, *121*, 471–495.
23. Zhang, J.; Xiao, X. Predicting chaotic time series using recurrent neural network. *Chin. Phys. Lett.* **2000**, *17*, 88.
24. Zaremba, W.; Sutskever, I.; Vinyals, O. Recurrent neural network regularization. *arXiv* **2014**, arXiv:1409.2329.
25. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780.
26. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M. Playing atari with deep reinforcement learning. *arXiv* **2013**, arXiv:1312.5602.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.