

Article

Towards Sustainability and Energy Efficiency Using Data Analytics for HPC Data Center

Andrea Chinnici¹, Eyvaz Ahmadzada^{2,3}, Ah-Lian Kor², Davide De Chiara⁴, Adrián Domínguez-Díaz¹, Luis de Marcos Ortega¹ and Marta Chinnici^{3,*}

- ¹ Departamento de Ciencias de la Computación, Universidad de Alcalá, 28801 Madrid, Spain; andrea.chinnici@edu.uah.es (A.C.); adrian.dominguez@uah.es (A.D.-D.); luis.demarcos@uah.es (L.d.M.O.)
- ² School of Built Environment, Engineering, and Computing, Leeds-Beckett University, Leeds LS2 3AE, UK; e.ahmadzada2254@student.leedsbeckett.ac.uk (E.A.); a.kor@leedsbeckett.ac.uk (A.-L.K.)
- ³ ENEA Casaccia Research Center, Department of Energy Technologies and Renewable Sources, ICT Division-HPC Lab, 00123 Rome, Italy
- ⁴ ENEA Portici Research Center, Department of Energy Technologies and Renewable Sources, ICT Division-HPC Lab, 80055 Portici, Italy; davide.dechiara@enea.it
- * Correspondence: marta.chinnici@enea.it

Abstract: High-performance computing (HPC) in data centers increases energy use and operational costs. Therefore, it is necessary to efficiently manage resources for the sustainability of and reduction in the carbon footprint. This research analyzes and optimizes ENEA HPC data centers, particularly the CRESCO6 cluster. The study starts by gathering and cleaning extensive datasets consisting of job schedules, environmental conditions, cooling systems, and sensors. Descriptive statistics accompanied with visualizations provide deep insight into collated data. Inferential statistics are then used to investigate relationships between various operational variables. Finally, machine learning models predict the average hot-aisle temperature based on cooling parameters, which can be used to determine optimal cooling settings. Furthermore, idle periods for computing nodes are analyzed to estimate wasted energy, as well as for evaluating the effect that idle node shutdown will have on the thermal characteristics of the data center under consideration. It closes with a discussion on how statistical and machine learning techniques can improve operations in a data center by focusing on important variables that determine consumption patterns.

Keywords: data center optimization; high-performance computing; energy efficiency; machine learning; thermal management; predictive modeling



Citation: Chinnici, A.; Ahmadzada, E.; Kor, A.-L.; De Chiara, D.; Domínguez-Díaz, A.; de Marcos Ortega, L.; Chinnici, M. Towards Sustainability and Energy Efficiency Using Data Analytics for HPC Data Center. *Electronics* **2024**, *13*, 3542. <https://doi.org/10.3390/electronics13173542>

Academic Editors: Fernando De la Prieta Pintado and Jinkun Geng

Received: 25 July 2024

Revised: 31 August 2024

Accepted: 2 September 2024

Published: 6 September 2024



Copyright: © 2024 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

1.1. Background and Motivation

The motivation for analyzing the energy-intensive operations within a DC is to provide deeper insight into DC energy consumption and build reliable predictive models. This study presents AI-based modeling approaches and strategies for managing energy efficiency in DC with a focus on IT reliability and sustainability as well as thermal operations.

For the modern digital economy, data centers play critical roles in service provision such as cloud computing and big data analytics through different activities' management such as collection, storage, and distribution. To accomplish its objectives, a good system should have a sufficient computing capacity in terms of hardware such as high processing units for servers, which could support the efficient handling of large amounts of information.

1.2. Data Center Components

A data center's core is made up of servers, which perform multiple functions ranging from file storing to complex analysis procedures. They are classified into server racks depending on their functions (e.g., application servers, database servers, or web servers).

Blade servers provide a compact solution that optimizes real-estate utilization within the server room [1].

Storage infrastructures ensure availability and integrity within the organization's premises to improve operations. Data centers have several storage options such as Storage Area Networks (SANs) and Network Attached Storage (NAS), which provide scalability and flexibly [2].

The network infrastructure within a data center is designed to facilitate internal as well as external connections. It essentially includes devices such as routers and switches, which manage efficient data transmission. On the other hand, software-defined networking (SDN) enhances the agility and efficiency of this system [3].

To ensure the uninterrupted power supply of energy in the facility, data centers must be equipped with Uninterruptible Power Supplies (UPS) and backup generators. The facility's electrical energy is distributed using efficient Power Distribution Units (PDUs) that minimize energy wastage. Nonetheless, cooling systems are necessary to mitigate heat dissipation by computing devices, particularly, emergent immersion cooling systems which offer enhanced thermal management as well as energy saving [4].

1.3. CRESCO6 Cluster

At the ENEA HPC data center in Portici, the CRESCO6 cluster consists of 434 nodes, each having two Intel Xeon Platinum 8160 processors and 192 GB RAM. These nodes are interconnected over an Intel Omni-Path network that supports parallel simulations and big data processing applications. CRESCO 6 is the second-biggest HPC cluster in Italy and the possibility to provide AI analysis using a real energy/thermal/computing consumption dataset related to this cluster is a privilege and novelty with respect to the current state of the art.

1.4. Aim and Research Objectives

The study is expected to focus on the CRESCO6 cluster, with an emphasis on energy consumption and resource allocation optimization. The research objectives are categorized into four phases:

- Phase 1: Exploratory data analysis—Explore and understand the dataset;
- Phase 2: Inferential statistical analysis—Explore relationships between operational variables;
- Phase 3: Predictive modeling—Create and train models that can predict hot-aisle temperature from operational data;
- Phase 4: Impact of idle node shutdown—Investigate the impact of thermal characteristics and energy consumption.

1.5. Rationale

Network infrastructures within a data center consist of routers, switches, and other network devices handling data flow to ensure efficient data transmission. In contrast, advanced solutions such as software-defined networking (SDN) improve agility and efficiency [3].

Optimized energy consumption leads to reduced operational expenses and increased economic and environmental benefits. Technologies such as machine learning and big data analytics may be utilized to implement smarter resource allocation strategies for more energy-efficient facilities [5]. Despite efforts towards improving efficiency at data centers, there is a gap in the deployment of comprehensive, predictive analytics-driven approaches with incorporated optimization techniques. This study seeks to address this gap by identifying mechanisms that can be used to efficiently reduce energy utilization [6]. Data-center-related best practice recommendations could also benefit data center policy-makers [7].

1.6. Contribution of Research

The present study improves predictive analytics for energy demand within data centers by using and comparing different machine learning algorithms. These models are more accurate and reliable; hence, they can be used for better predicting resulting in informed decisions toward resource management [8]. This study presents a step forward in terms of AI techniques applied to real-energy-consumption data center datasets.

In addition to theoretical contributions, this article offers practical recommendations and guidelines for operators of data centers. The insights come from empirical analysis and are verified by simulations, making them valuable sources of information to practitioners who want to improve their energy efficiency and effectiveness [9].

2. Related Work

In this section, the authors discuss the state-of-the-art and emerging energy-efficiency measures that have reduced DC energy consumption. This study reviews various practices and methods for advanced DC energy management.

Energy efficiency in data centers is crucial due to the increasing demand for digital services and computing power [10], which could have negative implications on the environment. To keep up with high-density computing equipment, servers, storage devices, and networking equipment need an uninterruptible electricity supply. Policies aimed at improving energy efficiency will not only minimize GHG emissions but also reduce electric bills [11]. Moreover, it enhances reliability by reducing hardware failure due to excessive heat and higher power usage rates within the facility.

2.1. Historical Evolution and Current State of Data Centers

Data centers have undergone major transformations over time due to their initial dependence on large, centralized mainframe computers tailored to specialized conditions and then the server farms, occasioned by smaller affordable servers, that followed when cloud computing became common during the early 2000s, enabling the flexible, cost-effective delivery of computer resources over the internet. During this transformation, companies such as Google, Amazon, and Microsoft played a central role [12].

Virtualization technology involving multiple virtual machines running on one physical server increases resource utilization and reduces hardware needs [13]. Modern servers are more energy efficient, while innovations in cooling technologies have further enhanced efficiency.

2.2. Challenges in Achieving Energy Efficiency

There are numerous challenges to achieving energy efficiency. Since servers in data centers must operate continuously, the data center's power requirements remain constant [14]. There is also the issue of more heat being generated due to denser computing racks that could not be efficiently cooled by traditional cooling systems [15]. Additionally, older servers and networks consume more power, while over-provisioned assets result in energy wastage during off-peak periods. The need for data processing and storage continues to increase, thus escalating energy consumption.

2.3. Strategies for Improving Data Center Energy Efficiency

Progressing towards better energy utilization entails advancement in the technical and operational aspects of data centers. For example, modern servers are equipped with more advanced power management capabilities [16]. Cutting-edge cooling systems such as liquid cooling or free air-cooling are more effective compared to traditional air conditioning methods [17]. The deployment of multiple virtual machines on a single physical server increases the rate of resource utilization by reducing the number of hardware equipment needed per unit area [18]. Data Center Infrastructure Management (DCIM) tools help to optimize energy efficiency through resource monitoring and management. Solar and wind energies among other renewable sources are also being integrated into DC operations.

2.4. Thermal Management in Data Centers

Thermal management must be prioritized in data center designs since it would inevitably affect its availability and reliability. This helps to maintain optimal operating temperatures by preventing overheating, which eventually leads to system failure [19]. More advanced cooling processes such as liquid cooling aid in maintaining optimal temperature and reducing energy use, while free air cooling is a more efficient substitute for traditional air conditioning [20].

To manage airflow more efficiently, many data centers use hot/cold-aisle-containment strategies. In this setup, servers are arranged in alternating rows, or aisles, of hot and cold air. Cold aisles face the air conditioning output vents, ensuring that cool air is drawn into the server intakes. Hot aisles face the server exhausts where hot air is expelled [21].

Computational fluid dynamics (CFD) is another advanced technique used to simulate and analyze airflow patterns within data centers. The use of CFD involves developing detailed models of the data center, which include server placement, cooling systems, and physical barriers with such information that it is possible to tell how air will circulate and where hot spots could develop [22].

2.5. Machine Learning Applications in Data Centers

Machine learning (ML) helps to increase the efficiency and manageability of data centers. It can automate processes such as predictive maintenance and energy optimization [23]. Predictive maintenance uses machine learning to anticipate possible hardware failures and reduce downtime. Energy usage is also optimized by ML through the analysis of consumption patterns and system optimization [24].

Workload management is another area where machine learning is making a significant difference in data centers. The best way that tasks can be distributed across servers based on their workloads can be predicted by ML algorithms for performance optimization with minimum energy wastage. This balancing act helps to prevent overheating on servers, hence reducing the chance of hardware failure risk through discharge into low-power states [25].

2.6. Workload Placement and Optimization

Efficient workload placement enhances the performance of a data center and its energy efficiency also. Load balancing involves even-workload distribution to prevent over or under loading [26]. In dynamic resource allocation, resource provision is based on demand, while virtualization involves running multiple virtual machines on one server, thereby increasing resource utilization [27]. Machine learning algorithms help optimize workload placement for efficient resource use and thermal management [28].

3. Methodology

The study has four stages of data center performance analysis and optimization. Each phase employs specific tools and approaches.

3.1. Phase 1: Exploratory Analysis

In this phase, data cleaning, pre-processing, and descriptive statistical analysis are employed to explore the characteristics of the datasets. Data are collected from the ENEA Portici HPC data center for the period of 1 January to 31 December 2020 (Table 1).

Data cleaning involves handling missing values, removing duplicates, and normalizing data. For example:

- Missing values in the job dataset's directory column were filled with "Not Specified", and invalid job status values were corrected;
- The environmental data's "FAIL" values were replaced by averaging adjacent valid entries.

Table 1. Dataset columns.

Dataset	Columns
Job data	id, jobid, numcores, user, queue, directory, executable, jobstatus, start, stop, numhost, hostlist
Environmental data	timestamp, timestamp2, cold/hot-aisle temperature, cold/hot-aisle humidity
Cooling data	time, machine name, machine status, supply air, return air, relative humidity, fan speed, cooling, free cooling
Sensor data	nodename, tempo, timestamp_measure, sys power, cpu power, mem power, fan1a, fan1b, fan5a, fan5b, sys util, cpu util, mem util, io util

As the next step, summary statistics such as mean, median, and standard deviation are computed. Data visualizations in the form of histograms and box plots provide insights into data distributions and variability.

3.2. Phase 2: Inferential Statistical Analysis

In this phase, we explore the relationships between different parameters within the data center. Correlation analysis is conducted:

- Environmental and cooling parameters: We compute correlation matrices between environmental variables (e.g., average hot-aisle temperature) and cooling system variables (e.g., fan speed, cooling capacity) to identify significant correlations;
- Cooling system and node sensor measurements: Correlations between cooling parameters and node sensor data (e.g., CPU power, ambient temperature) are analyzed to understand system-load impacts;
- Job parameters and sensor data: To understand job-load impacts on the data center operations, correlations between job characteristics (e.g., cores used, job duration) and sensor measurements are studied.

Heatmaps are used to visualize the strength and direction of correlations.

3.3. Phase 3: Machine Learning Model for Cooling Optimization

In this phase, machine learning models are proposed for predicting optimal cooling parameters based on environmental and sensor data.

For data preparation, we merge sensor, cooling, and environmental datasets into a single data frame with imputed missing values where needed using mean values to maintain the integrity of the dataset.

Next, we train the following models:

- Temperature prediction: We use ridge regression ($\alpha = 1.0$) to predict average hot-aisle temperature based on cooling parameters. The data are split into training (80%) and testing (20%) sets, and features are standardized using StandardScaler;
- Optimal cooling parameters: Another ridge regression model predicts optimal cooling parameters (fan speed, cooling capacity) based on sensor and environmental data.

These models are then run against all the records in the dataset to predict cooling parameters as well as average hot-aisle temperature for robust performance across different scenarios.

3.4. Phase 4: Idle Node Analysis and Predictive Modeling for Thermal Impact

Phase four includes the analysis of periods when nodes are idle to compute energy waste and predict the thermal impact of turning off idle nodes.

The information about the time during which the nodes were idle is derived from job data, considering that any gaps between when a job starts and stops may imply an idle period longer than 3 h. Next, a dynamic calculation of the wasted energy during idle periods based on the sensor readings' time differences is carried out using the sensor data.

For the predictive modeling:

- The idle node shutdown is considered in the development of the ridge regression model, which predicts hot aisle average temperature values;
- The performance of the model is assessed using Mean Squared Error (MSE), and predictions are compared with actual recorded temperatures to determine how they impact the thermal state of DCs.

4. Results and Discussion

4.1. Phase 1: Exploratory Data Analysis

First, job submission patterns, job status distribution, and queue distribution were analyzed. Key results include:

- Daily job counts: Trends over time (Figure 1);
- Job status distribution: Visualized using a count plot (Figure 2).

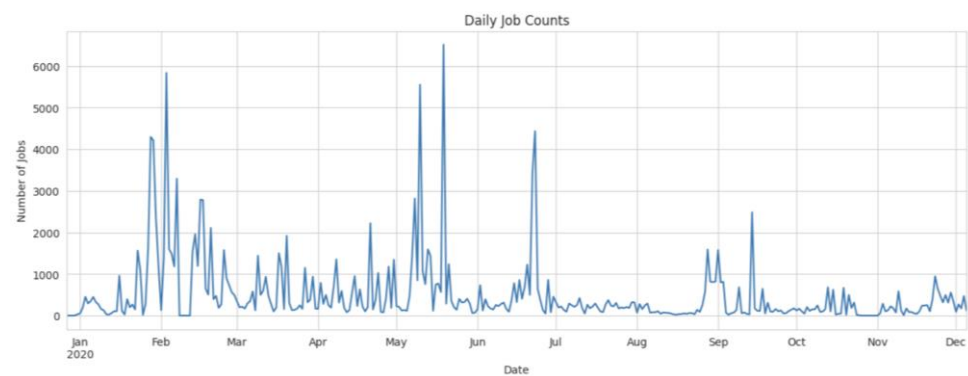


Figure 1. Daily job counts.

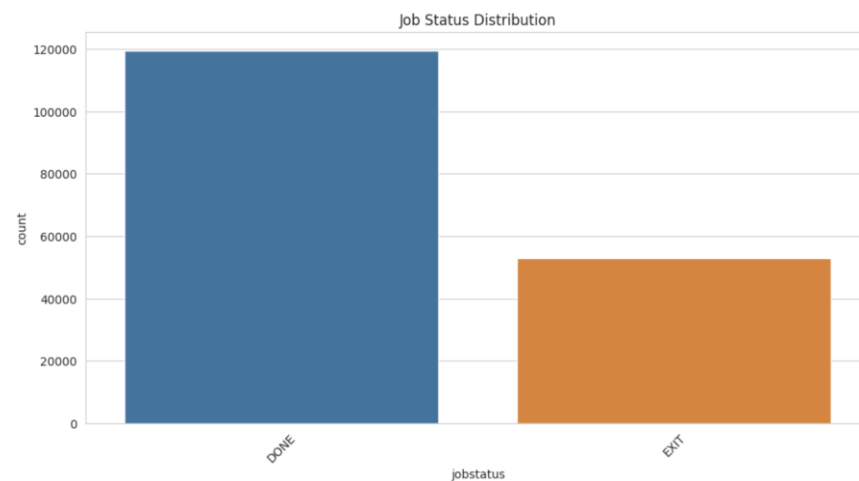


Figure 2. Job status distribution.

The analysis of daily job counts shows a generally consistent pattern with minor fluctuations observed on certain days. These fluctuations may be attributed to variations in workload demand or scheduling practices. Despite these occasional changes, the overall stability in job counts indicates a well-managed and balanced job distribution within the data center.

The job status analysis shows that many jobs have “exited”, meaning they did not finish as expected. This leads to wasted energy since the resources used for these jobs are not productive. Reducing the number of exited jobs is important to avoid unnecessary energy consumption.

Next, the seasonal trends in average temperatures and humidity for the hot and cold aisles are analyzed. The results are shown in Figure 3. As we can see, the aisles were overheating during the first 4 months of the year.

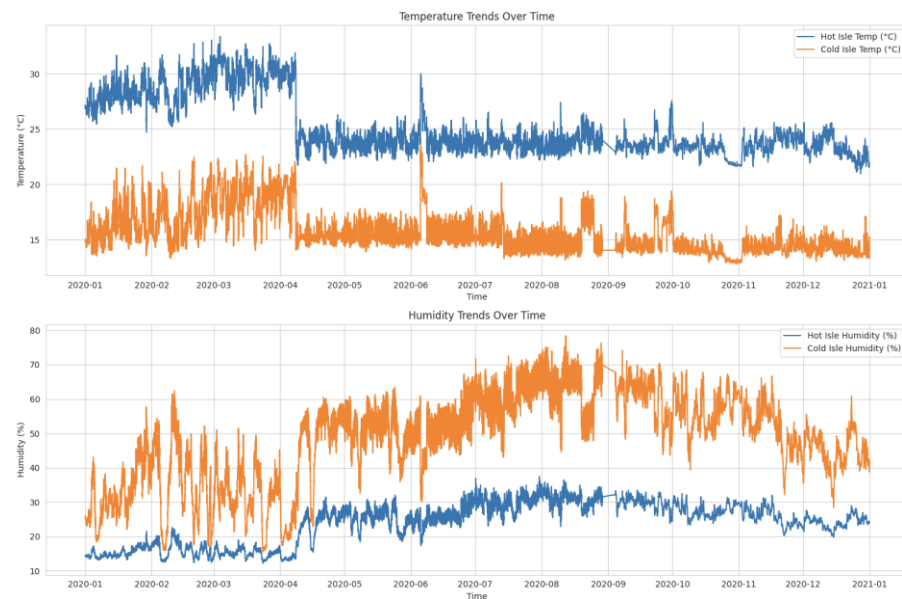


Figure 3. Temperature and humidity trends Over time.

To investigate the overheating issue, we analyze the temperature gradient at nodes and the average CPU temperature over time (Figure 4).

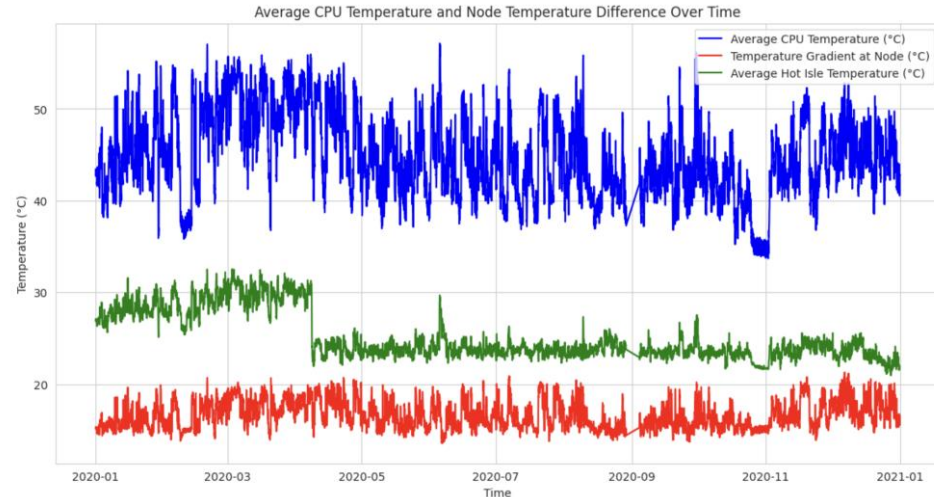


Figure 4. Average CPU temperature and node temperature differences over time.

The results suggest that there was no significant overheating in the CPU and inside the nodes, which means that there must have been an issue with the cooling system since the hot-aisle temperature was higher from the month of January till April.

4.2. Phase 2: Inferential Statistics

In this phase, we analyzed environmental conditions, cooling parameters, job characteristics, and sensor parameters to understand the correlations between them.

As we can see in the correlation matrix (Figure 5), there is a strong positive correlation between the average hot-aisle temperature and return-air temperature (0.94). This indicates that higher hot-aisle temperatures are associated with higher return-air temperatures. Conversely, there is a strong negative correlation with fan speed (-0.82), suggesting that,

as the hot-aisle temperature increases, the fan speed decreases. This might indicate that the cooling system is trying to adjust to maintain optimal conditions.

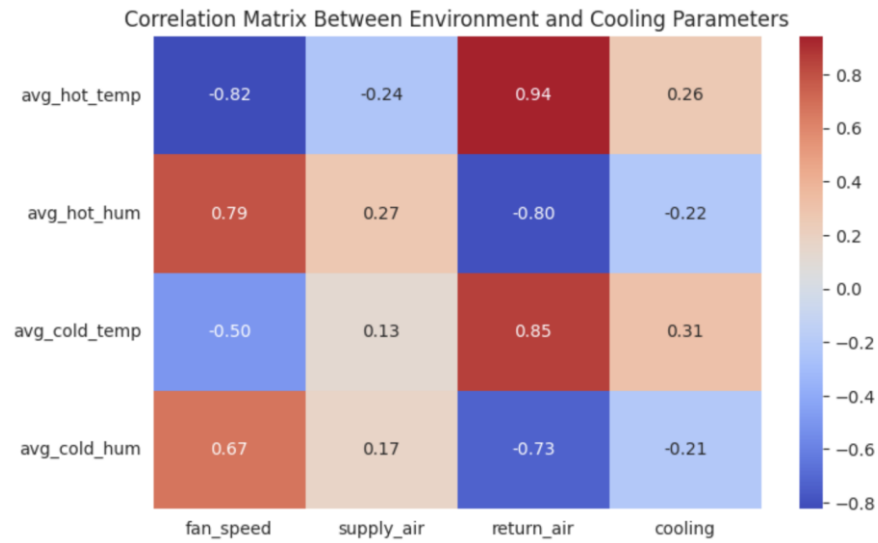


Figure 5. Correlation matrix Between environment and cooling parameters.

4.3. Phase 3: Predictive Modeling of Hot-Aisle Temperature and Cooling Parameters

First, we built a model to predict average hot-aisle temperature based on the cooling parameters. The model is tested on the entire dataset to predict the average hot-aisle temperature. The MSE for the full dataset was 1.3156, demonstrating consistent model performance across different data splits (Figure 6). The close alignment of data points along the diagonal line in the plot indicates that the model consistently performs well in predicting the average hot-aisle temperatures.

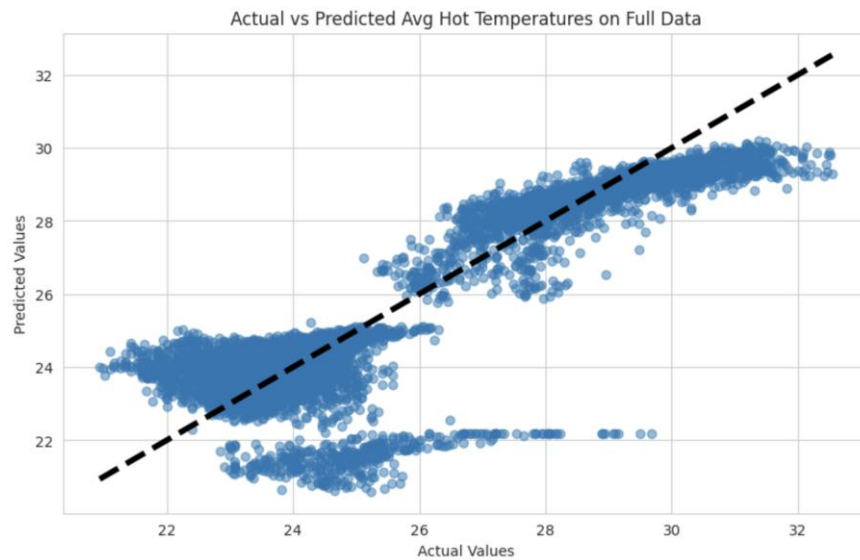


Figure 6. Actual vs. predicted average hot-Aisle temperature based on cooling parameters.

Next, we trained a regression model to predict optimal cooling parameters under ideal conditions, where the average hot-aisle temperature was between 18 °C and 27 °C. The prediction was performed based on the data center sensor readings. The model was applied to the entire dataset under ideal conditions. The MSE for the full dataset was 42.6744, which indicates variability in the data.

Figure 7 shows the actual versus predicted cooling fan speed values under ideal conditions. The proximity of the points to the diagonal line indicates good predictions, excluding a specific period. We found in Phase 1 that there was a period in the dataset with cooling system disruptions. Therefore, this behavior is correct since the model optimizes the cooling values during the disruption period.

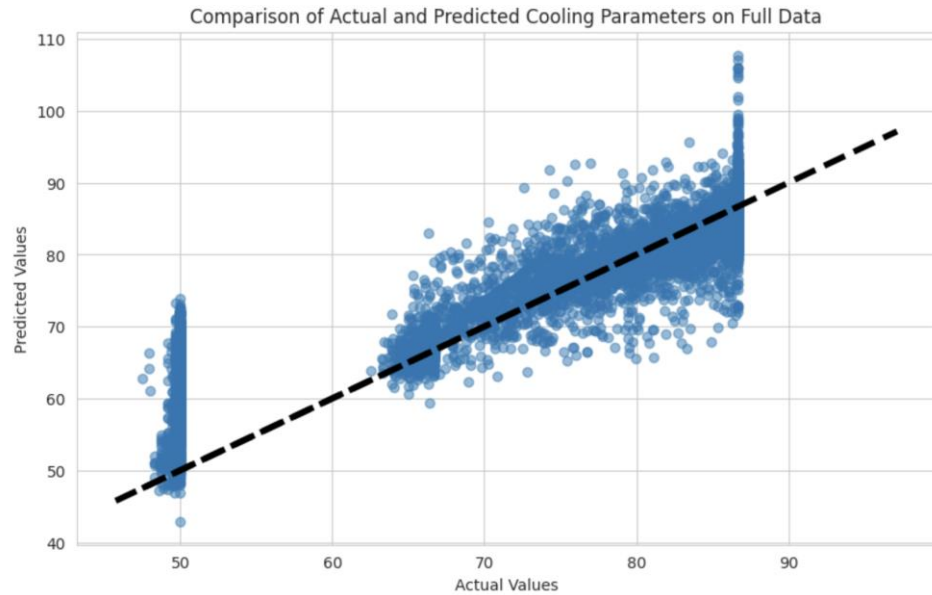


Figure 7. Actual vs. predicted fan speed values under ideal conditions.

Finally, to assess the impact of using ideal cooling parameters, the predicted cooling parameters determined by the ridge regression model were applied to the whole dataset. The results indicate how closely the predicted temperatures align with the actual recorded temperatures.

The scatterplot in Figure 8 shows the actual versus predicted average hot-aisle temperatures with the recommended ASHRAE maximum temperature of 27 °C highlighted. This comparison indicates whether the predicted temperatures remain within the recommended range under ideal cooling conditions.

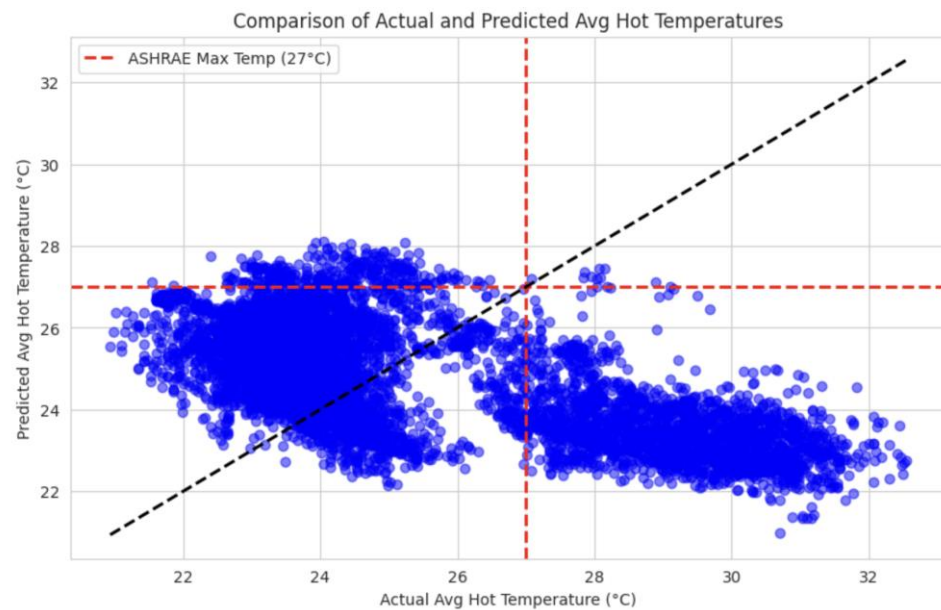


Figure 8. Actual vs. predicted (ideal) average hot-Aisle temperature.

When the ideal cooling parameters are applied, there is a noticeable decrease in the average hot-aisle temperature. This adjustment ensures that the cooling system operates at its most efficient settings, effectively reducing the thermal load within the data center. The results shows that most of the predicted temperatures fall below the maximum recommended temperature of 27 °C. This indicates that the optimized cooling parameters are successful in maintaining a lower and more stable temperature, which is crucial for the efficient and reliable operation of the data center's equipment.

4.4. Phase 4: Energy Optimization and Thermal Management

In this phase, we first identified the idle periods for each node over the entire year. These periods were defined as times when the nodes were not running any jobs. The results showed that there were significant idle periods throughout the year for many nodes. Then, we calculated the wasted energy during these idle periods. The wasted energy was computed by summing the power consumption of the nodes during the idle times. The initial calculation showed that a significant amount of energy was wasted when the nodes were idle.

As the next step, we trained a model to predict the data center average hot-aisle temperature based on the sensor parameters. The model's MSE was calculated to be 0.61, suggesting an accurate prediction (Figure 9).

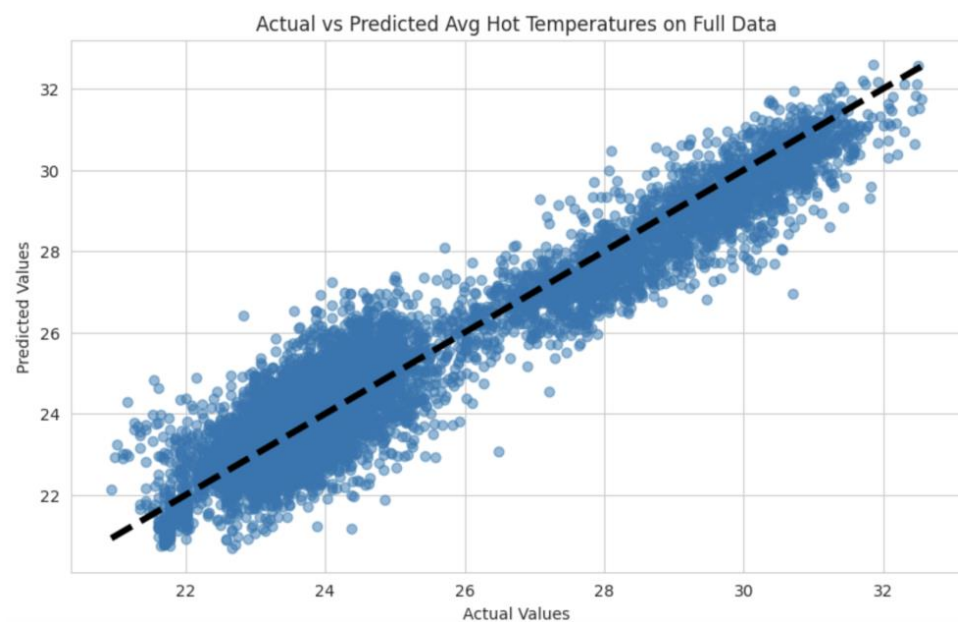


Figure 9. Actual vs. predicted average hot-Aisle temperature based on sensor parameters.

Finally, we applied the temperature prediction model created previously on the data where nodes were turned off during idle periods to predict the average hot-aisle temperature. The MSE for these predictions was higher at 9.30, indicating some deviations due to the changes in power consumption patterns.

As shown in Figure 10, the predicted temperatures are consistently lower than the recorded values across various periods. This indicates that, by turning off idle nodes, the data center could achieve a more stable and improved thermal state. The reduction in temperature highlights the potential effectiveness of this approach in enhancing the overall cooling efficiency and reducing unnecessary thermal loads within the data center.

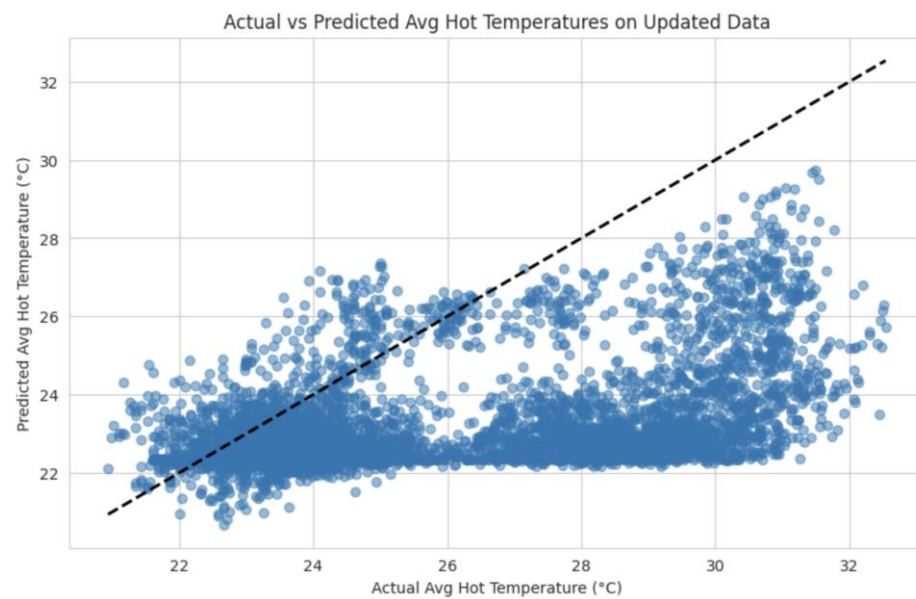


Figure 10. Actual vs. predicted average Hot-Aisle temperature on updated data.

5. Recommendations

Based on our findings, we make several recommendations to enhance the efficiency and sustainability of the ENEA HPC data center with a special focus on the CRESCO6 cluster. However, this study suggests that these models should be used to predict average hot-aisle temperature and adjust cooling parameters accordingly. Therefore, it can maintain optimum thermal conditions and reduce reactive cooling needs. It is important to regularly review and adjust cooling settings to comply with the ASHRAE guidelines concerning temperatures.

It is necessary to develop job-scheduling algorithms that consider the thermal effect of job allocations. These strategies ensure the equitable dispatch of computations across nodes leading to optimal thermal states while preventing hotspots from developing. Resource distribution can be adjusted in real time by monitoring systems based on current thermal characteristics and job requirements.

Establishing a comprehensive real-time monitoring system that tracks environmental conditions, cooling system performance, and job metrics is vital. This information will provide informed decision-making as well as trigger alerts when there are deviations from ideal environmental conditions. Continuously collecting high-quality data from sensors and monitoring systems, validating it regularly, and cleaning it up are all essential for ensuring accuracy.

Cooling parameters vs. energy consumption correlation analysis can be used to reduce energy use. Implementing energy-efficient cooling practices, as well as optimizing cooling system operations, reduces energy wastage. Furthermore, the integration of renewable energy sources such as solar/wind power would reduce environmental effects to enhance sustainable goals.

6. Conclusions and Future Work

This research aimed to improve cooling effectiveness and thermal management in the ENEA HPC data center, with a focus on the CRESCO6 cluster. The study was divided into four phases, each addressing one of the main aspects of operation principles in data centers.

In Phase 1, we looked at job patterns, environmental conditions, cooling system performance, and sensor data. We identified important trends and correlations such as those between hot-aisle temperatures and cooling parameters. The Gini coefficient further demonstrated that nodes had fairly equal job allocations, meaning the efficient use of resources.

Phase 2 of the research encompassed a detailed correlation analysis between environmental conditions, cooling parameters (fan speed and return-air temperature), power consumption, and job characteristics. This enabled us to identify very strong relationships between hot-Aisle temperatures versus fan speeds and/or exhaust-air temperatures to assess various influences on data center thermal characteristics.

The next stage involved the development of models that could estimate the average hot-aisle temperature based on cooling parameters. The ridge regression model has a Mean Squared Error (MSE) of 1.3191, indicating that it can accurately predict temperature changes, and most predicted temperatures were within the ASHRAE limits, guaranteeing safe operational temperatures.

Lastly, our focus shifted toward energy optimization for the CRESCO6 cluster. We switched off nodes during idle periods, thus minimizing energy waste, which also contributed to improved thermal characteristics. The correlation between the sensor data and average hot-aisle temperature showed the effectiveness of this approach, while the temperature prediction model worked well despite variations due to changing cooling needs. This analysis demonstrated that proactive energy management, together with real-time monitoring, is essential in improving efficiency in a data center.

The outcome of this research shows that data-driven approaches have a significant role in optimizing data center operations. We utilized relationships between various operational parameters to build models for the effective predicting and managing of thermal conditions. The approach taken in this thesis demonstrates that small changes like switching off unused nodes can lead to substantial improvements in energy efficiency and thermal stability. This would not only cut down on operating expenses but also support international efforts toward green data centers by reducing environmental effects.

Additionally, the study underscores the significance of continuous monitoring and adaptive strategies in maintaining optimal conditions within data centers. The predictive models developed herein serve as a basis for more advanced systems capable of automatically adjusting cooling parameters and resource allocation in real time. It is important to have this ability since it helps to manage the increased complexity and demands associated with modern high-performance computing environments, hence making them run efficiently and sustainably.

Further work could be carried out by considering several strategies. Future work can extend this study in several ways:

- By using real-time monitoring systems, we can make dynamic adjustments of cooling parameters based on current conditions. This applies the learned models in real time, thus improving efficiency;
- Exploring advanced machine learning techniques such as neural networks, ensemble methods, and deep learning could help to improve prediction accuracy for cooling and thermal conditions;
- Optimizing job scheduling algorithms considering both computational efficiency and thermal management could help to balance workloads while maintaining optimal thermal conditions;
- Incorporating new datasets from other HPC data centers, such as CRESCO7 and CRESCO8, could help to validate and enhance the models developed, ensuring broader applicability and robustness across different environments.

Addressing these aspects will enable further research to increase the sustainability and efficiency of, as well as the reliability of, data center operations towards broader environmental conservation objectives and cost reduction. This thesis has displayed how advanced data analytics and machine learning techniques can be leveraged to optimize data center operations, making way for highly efficient and sustainable high-performance computing environments.

Author Contributions: Methodology, E.A.; Formal analysis, D.D.C.; Investigation, A.C.; Writing—original draft, A.-L.K.; Writing—review & editing, A.D.-D. and L.d.M.O.; Supervision, M.C. All authors have read and agreed to the published version of the manuscript.

Funding: Marta Chinnici and Davide De Chiara were supported for this research by Project ECS 0000024 Rome Technopole—CUP B83C22002820006, the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2 Investment 1.5; they were funded by the European Union—NextGenerationEU.

Data Availability Statement: Data are contained within the article.

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

References

1. Barroso, L.A.; Clidaras, J.; Holzle, U. *The Datacenter as a Computer: An Introduction to the Design of Ware-House-Scale Machines*, 2nd ed.; Morgan & Claypool Publishers: San Rafael, CA, USA, 2013.
2. Mell, P.; Grance, T. *The NIST Definition of Cloud Computing*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2011.
3. Hamilton, J. Cooperative expendable micro-slice servers (CEMS): Low cost, low power servers for internet-scale services. In Proceedings of the Conference on Innovative Data Systems Research (CIDR), Asilomar, CA, USA, 4–7 January 2009.
4. ASHRAE. *Thermal Guidelines for Data Processing Environments*, 5th ed.; American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.: Atlanta, GA, USA, 2021; ISBN 978-1-936504-79-7.
5. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Science & Business Media: New York, NY, USA, 2009.
6. Georgiou, S.; Rizou, S.; Spinellis, D. Software Development Lifecycle for Energy Efficiency. *ACM Comput. Surv.* **2019**, *52*, 1–33. [[CrossRef](#)]
7. Manotas, I.; Bird, C.; Zhang, R.; Shepherd, D.; Jaspan, C.; Sadowski, C.; Pollock, L.; Clause, J. An empirical study of practitioners' perspectives on green software engineering. In Proceedings of the 38th International Conference on Software Engineering, Austin, TX, USA, 14–22 May 2016.
8. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.
9. Gebreyesus, Y.; Dalton, D.; De Chiara, D.; Chinnici, M.; Chinnici, A. AI for Automating Data Center Operations: Model Explainability in the Data Centre Context Using Shapley Additive Explanations (SHAP). *Electronics* **2024**, *13*, 1628. [[CrossRef](#)]
10. Berisha, B.; Mëzriu, E.; Shabani, I. Big data analytics in Cloud computing: An overview. *J. Cloud Comput.* **2022**, *11*, 24. [[CrossRef](#)]
11. Mondal, S.; Bin Faruk, F.; Rajbongshi, D.; Efaz, M.M.K.; Islam, M. GEECO: Green Data Centers for Energy Optimization and Carbon Footprint Reduction. *Sustainability* **2023**, *15*, 15249. [[CrossRef](#)]
12. Huerta, E.A.; Khan, A.; Davis, E.; Bushell, C.; Gropp, W.D.; Katz, D.S.; Kindratenko, V.; Koric, S.; Kramer, W.T.; McGinty, B.; et al. Convergence of artificial intelligence and high-performance computing on NSF-supported cyber-infrastructure. *J. Big Data* **2020**, *7*, 88. [[CrossRef](#)]
13. Gonzalez, N.M.; Carvalho, T.C.; Miers, C.C. Cloud resource management: Towards efficient execution of large-scale scientific applications and workflows on complex infrastructures. *J. Cloud Comput.* **2017**, *6*, 13. [[CrossRef](#)]
14. Panwar, S.S.; Rauthan, M.M.S.; Barthwal, V. A systematic review on effective energy utilization management strategies in cloud data centers. *J. Cloud Comput.* **2022**, *11*, 95. [[CrossRef](#)]
15. Manganelli, M.; Soldati, A.; Martirano, L.; Ramakrishna, S. Strategies for Improving the Sustainability of Data Centers via Energy Mix, Energy Conservation, and Circular Energy. *Sustainability* **2021**, *13*, 6114. [[CrossRef](#)]
16. Liu, J.; Yan, L.; Yan, C.; Qiu, Y.; Jiang, C.; Li, Y.; Li, Y.; Cérin, C. Escape: An Energy Efficiency Simulator for Internet Data Centers. *Energies* **2023**, *16*, 3187. [[CrossRef](#)]
17. Liu, C.; Yu, H. Evaluation and Optimization of a Two-Phase Liquid-Immersion Cooling System for Data Centers. *Energies* **2021**, *14*, 1395. [[CrossRef](#)]
18. Çağlar, I.; Altılar, D.T. Look-ahead energy efficient VM allocation approach for data centers. *J. Cloud Comput.* **2022**, *11*, 11. [[CrossRef](#)]
19. Guo, Y.; Zhao, C.; Gao, H.; Shen, C.; Fu, X. Improving Thermal Performance in Data Centers Based on Numerical Simulations. *Buildings* **2024**, *14*, 1416. [[CrossRef](#)]
20. Xu, S.; Zhang, H.; Wang, Z. Thermal Management and Energy Consumption in Air, Liquid, and Free Cooling Systems for Data Centers: A Review. *Energies* **2023**, *16*, 1279. [[CrossRef](#)]
21. Chen, H.; Li, D.; Wang, S.; Chen, T.; Zhong, M.; Ding, Y.; Li, Y.; Huo, X. Numerical investigation of thermal performance with adaptive terminal devices for cold aisle containment in data centers. *Buildings* **2023**, *13*, 268. [[CrossRef](#)]
22. Wibron, E.; Ljung, A.-L.; Lundström, T.S. Computational Fluid Dynamics Modeling and Validating Experiments of Airflow in a Data Center. *Energies* **2018**, *11*, 644. [[CrossRef](#)]
23. Chi, C.; Ji, K.; Song, P.; Marahatta, A.; Zhang, S.; Zhang, F.; Qiu, D.; Liu, Z. Cooperatively Improving Data Center Energy Efficiency Based on Multi-Agent Deep Reinforcement Learning. *Energies* **2021**, *14*, 2071. [[CrossRef](#)]

24. Mehta, Y.; Xu, R.; Lim, B.; Wu, J.; Gao, J. A Review for Green Energy Machine Learning and AI Services. *Energies* **2023**, *16*, 5718. [[CrossRef](#)]
25. Daradkeh, T.; Agarwal, A. Cloud Workload and Data Center Analytical Modeling and Optimization Using Deep Machine Learning. *Network* **2022**, *2*, 643–669. [[CrossRef](#)]
26. Malik, N.; Sardaraz, M.; Tahir, M.; Shah, B.; Ali, G.; Moreira, F. Energy-Efficient Load Balancing Algorithm for Workflow Scheduling in Cloud Data Centers Using Queuing and Thresholds. *Appl. Sci.* **2021**, *11*, 5849. [[CrossRef](#)]
27. Sabyasachi, A.S.; Muppala, J.K. Cost-Effective and Energy-Aware Resource Allocation in Cloud Data Centers. *Electronics* **2022**, *11*, 3639. [[CrossRef](#)]
28. Grishina, A.; Chinnici, M.; Kor, A.-L.; Rondeau, E.; Georges, J.-P. A Machine Learning Solution for Data Center Thermal Characteristics Analysis. *Energies* **2020**, *13*, 4378. [[CrossRef](#)]

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.