Physica A 624 (2023) 128966

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Advanced data analytics modeling for evidence-based data center energy management

Wania Khan [a], Davide De Chiara [b], Ah-Lian Kor [a,*], Marta Chinnici [c]

[a] Leeds Beckett University, Leeds, UK
[b] ENEA, C.R. Portici Piazzale Enrico Fermi 1, Napoli 80055, Italy
[c] ENEA, C.R. Casaccia, Via Anguillarese, 301 - 00123, Roma, Italy

ARTICLE INFO

Article history:
Available online 10 June 2023

Keywords:
Data center (DC)
Energy efficiency
Machine learning classification
Deep learning modeling
Prediction analysis
Time series forecasting

ABSTRACT

Over the past few decades, the demand for Data Center (DC) services has significantly increased due to the world's growing need for internet access, social networking, and data storage. Data Centers are among the most energy-intensive businesses, so optimizing IT operations in DC requires energy-efficient techniques. This paper presents AI based modeling strategies for effective energy management with a particular emphasis on DC's two most energy intensive systems (i.e., cooling and IT systems). This study addresses the issues of IT equipment performance degradation, inappropriate IT room thermal conditions, inefficient workload placement, and excessive energy waste. This research entails the application of machine learning for DC thermal classification, and deployment of deep learning models to predict resource utilization and energy consumption in DC. Furthermore, a comparative analysis is performed with existing relevant methods to demonstrate the effectiveness and accuracy of the proposed AI techniques. The findings of this study also provide evidence-based recommendations for DC efficient energy management.

1. Introduction

In today's data-driven cultures, Data Centers (DC) constitute a significant, mission-critical component of the computing infrastructure. They play an essential role in the IT industry worldwide and are continually expanding in size and complexity in terms of high-performance computing. The adoption of cutting-edge technologies in the field of digitalization has accelerated the growth of internet services such as big data analytics for businesses [1], Internet of Things (IoT), and cloud services [2], etc. Due to social networking, video streaming, conferencing, and online gaming, internet traffic has surged by more than 40% globally during the past two years [3]. The continual increase in demand for big data computing, processing, and storage [4] by a range of cloud service providers (i.e Google, Facebook, Twitter, etc.), is the key driver of DC's criticality [5]. Additionally, there has been an increase in electricity usage due to the expansion of IT systems. According to a recent report by the International Energy Agency [3], data centers are one of the most energy-intensive businesses, accounting for roughly 1% (200–250 TWh) of the world's electricity to support the rising demand for data-intensive technology [6].

The primary goal is to manage data center operations and control its associated energy consumption at various granularity levels [5]. This cooperative control of several DC components enhances both the stipulated Quality of Service

* Corresponding author.
E-mail address: A.Kor@leedsbeckett.ac.uk (A.-L. Kor).

https://doi.org/10.1016/j.physa.2023.128966
0378-4371/

(QoS) and the overall energy efficiency (Quality in Sustainability, QiS). Data centers are built differently based on their sizes [7]. The top priorities of data center deployment are good quality performance and also, energy efficiency. System control operations within a data center could be structured according to 3 levels: server/node level; rack level and data center level. The two broad categories within a data center are [8]:

1. The IT system: encompasses all IT equipment such as servers, storage devices, monitoring workstations, networking equipment.
2. The Facilities system: encompasses all the mechanical and electrical systems which are used to support the IT system such as uninterrupted power supply (UPS), power distribution equipment, cooling system/HVAC, computer room air conditioners (CRAC) units etc.

High server densities and correspondingly high-power consumption in DC result from expanding demand for computing resources [9]. Servers in DCs consume the most energy and account for more than 75% of the entire energy load of IT equipment. Storage devices are the second highest energy-consuming equipment making up 10%–15% of the total IT equipment energy load [8]. On the other hand, cooling infrastructure is the most energy-intensive facility, accounting for more than 50% of a DC's overall energy usage [9].

The purpose of this research is to apply appropriate AI techniques for efficient energy management in DC IT and thermal operations inclusively. **Organization of the paper**: This paper is organized as follows: Section 1 — Introduction; Section 2 — Related Work; Section 3 — Methodology (categorized into four phases); Section 4 — Results and Discussion; Section 5 — Conclusion; Section 6 — Recommendations and Future Work.

1.1. Background and motivation

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

The authors have summarized a review on projected DC increasing energy demand and excessive carbon emissions. The IT corporation Cisco's annual report has provided a global forecast that evaluates the digital transformation trend and has predicted that by 2023, 66% more people would be using the internet globally than in 2018 [10]. The DCs are the critical infrastructure to cope with such a rising trend. According to a survey by Synergy Research Group [11], there are over 400 major DCs worldwide in 2017. According to IEA estimation, data centers use 200–250 TWh of electricity in 2020, which represents 1% of the world's total electricity demand [3]. According to estimates by Cao [12], the DC industry is contributing up to 0.3% to the world's carbon emissions and it is projected to increase further in the next decade. The report by Kumar [13] states that the CO_2 emission from the ICT sector is anticipated to rise at a pace of 6% each year. Most research has investigated and evaluated the use of metrics for DC assessment in terms of sustainability and energy consumption [14]. A green DC has also been suggested to increase DC energy efficiency and reduce carbon footprint [15]. High energy efficiency in a DC mainly pertains to cooling and power supply systems [16]. Future challenges in greening DCs include maximizing energy efficiency and sustainability across all DC operations. Despite extensive research relating to DC energy efficiency, investigation of energy efficiency at different operational stages and their associated carbon footprint assessments remains limited in practice [17]. The following challenges need to be considered:

1. A lack of regulatory standard or framework for assessing DC sustainability that includes a comprehensive specification of particular measurements and methodology practices [18];
2. While considering energy-efficient solutions for server utilization, limited focus has been on integrating additional energy-consuming sources such as storage and network [19];
3. An in-depth evaluation of thermal characteristics analysis of IT rooms in real DCs [20];
4. Aborted jobs in DCs wastes resources and energy due to complex system dynamics. A more thorough examination of job disparities in correspondence to the operational and thermal characteristics of the compute node still requires due consideration [21].

The interaction between the computing and cooling systems is motivated by achieving overall energy efficiency in DC operations. The aim is to investigate factors that strongly affect DC energy efficiency. To achieve the aim of this research, we have implemented a three-phased methodology (see the methodology section). The research questions and objectives for this study are as follows:

RQ1: How could AI algorithms be applied to real DC efficient energy management?

To answer this research question, we have defined specific research objectives for each phase of our proposed methodology.

Phase I — IT Room Thermal Characteristics Analysis using Machine Learning

The goal of this phase is to analyze the DC IT room thermal conditions to maintain the equipment's operational environment as per ASHRAE 9.9 guidelines [22]. The research objectives for this phase are defined as follows.

- RO1.1: Implement supervised machine learning classification models to classify DC monitored data into thermal classes based on IT room environmental conditions;
- RO1.2: Evaluate results from RO1.1 using model's performance metrics.

Phase II — Prediction of Resource Utilization using Deep Learning Algorithm

In this phase, we aim to perform prediction analysis on DC real-time data using deep learning modeling techniques. The research objectives for this phase are as follows:

- RO1.3: Implement deep learning models on the monitored and estimated data to predict future DC resource utilization based on historical behavior;
- RO1.4: Evaluate results from RO1.3 and provide relevant recommendations.

Phase III — Future Forecast of DC Resource Energy Consumption and Waste Energy

This phase aims to forecast the active resource energy consumption and energy waste during jobs execution. The research objectives for this phase are defined as follows:

- RO1.5: Perform timeseries forecasting to predict active resource energy consumption and energy waste;
- RO1.6: Evaluate results from RO1.5.

2. Related work

In this section, the authors discuss the state-of-the-art and emerging energy efficiency measures that have brought about a reduction in DCs energy consumption. This study reviews various practices and methods for advance DC energy management.

2.1. Demand for data centers and electricity consumption

According to Dayarathna, Wen, and Fan [23], DCs are vital and energy-intensive computing infrastructures that host computer servers to run large-scale Internet-based computing jobs and provide services. The rapid adoption of data-intensive information, emerging technologies such as the Internet of Things (IoT), artificial intelligence (AI), smart and connected energy systems, big data analytics, blockchain, and 5G leads to increased demand for DC services. According to IEA research [3], increasing video streaming, social networking, and corporate operations digitization have contributed to a 15-fold rise in worldwide internet traffic since 2010 and an increase of over 40% in 2020. The annual report by Cisco [10] provides a global projection to evaluate the extent as well as trend of the digital revolution and predicts that the worldwide number of internet users will increase from 51% in 2018 to 66% by 2023. IEA [3] estimates that DCs use 200–250 TWh of power in 2020, representing 1% of the world's total electricity consumption.

2.2. Energy consumption outlook for data centers

Numerous research studies describe the energy consumed by various individual subsystems and components of the DC. The two main categories for DC energy use are computational and physical resources, as discussed in the Introduction Section. Statistics published by Rong and colleagues [9] have revealed that compute resources account for 50% of the overall DC energy consumption whilst physical resources contribute up to 40%. Other miscellaneous sub-systems (e.g., power supply system) use only 10%. A distribution of energy consumption of different components within a DC is presented in Fig. 2.1 [9]. The energy use of computing resources is further broken down into various subcomponents, such as servers' computation, which uses 40% of the energy, while communication and storage devices account for roughly 10%. Ahmed and colleagues [24] have also provided a component-based energy consumption model for the DC. According to their estimated energy consumption levels for various components, IT loads and cooling systems consume 86% of the total energy, whilst air conditioning and network equipment use 13%. The energy consumption of lightning facilities is roughly 1%, which is assumed to be negligible by most literature.

2.3. Data center carbon footprint

Due to the expansion of IT service providers' technological capabilities for high-performance computing, DCs significantly consume more electricity, as discussed in the previous section. This massive growth in DC energy consumption translates into a substantial increase in greenhouse gas emissions (mainly CO_2). This is because majority energy sources for DCs around the world are non-renewable resources and carbon-intensive fossil fuels (e.g. coal and natural gas) [25]. As a result of their high energy requirements for power transmission to large-scale computing infrastructure, DCs are now the concern for both environmental activists and governments. According to Cao and colleagues [12], the data center sector is estimated to account for 0.3% of global carbon emissions and is projected to increase further in the next decade. Additionally, excessive carbon emissions have a negative impact on the environment, and an estimated global DC carbon emission from 2018 to 2030 is depicted in Fig. 2.2. [12]. It can be seen in the graph how DC energy use and carbon emissions will grow over the next several years due to ongoing expansion of the world's digital economy.

In addition to an estimate of the DCs total carbon emissions, Fig. 2.3 also includes estimates of the CO_2 emissions for each ICT-information and communication technology-related category within DCs [13].

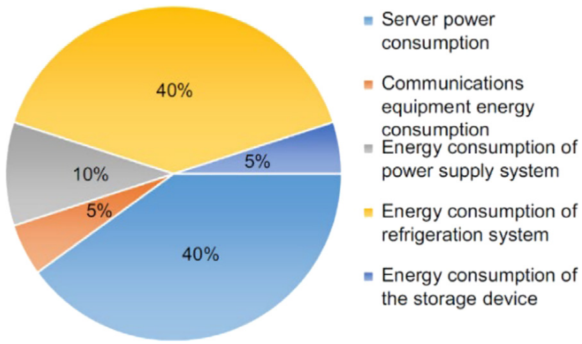

Fig. 2.1. Components-based energy consumption in a data center [9].

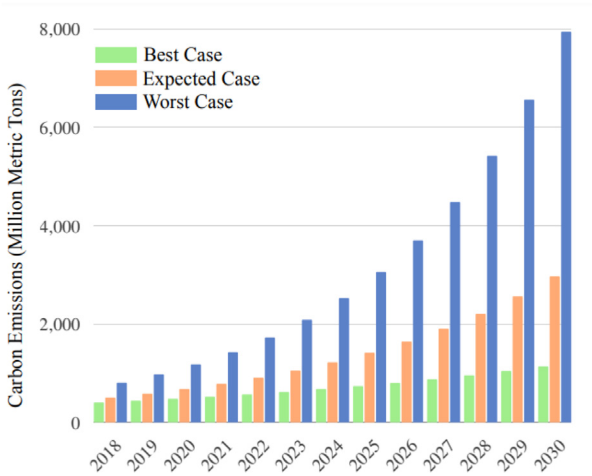

Fig. 2.2. Projected global carbon emissions from data centers [12].

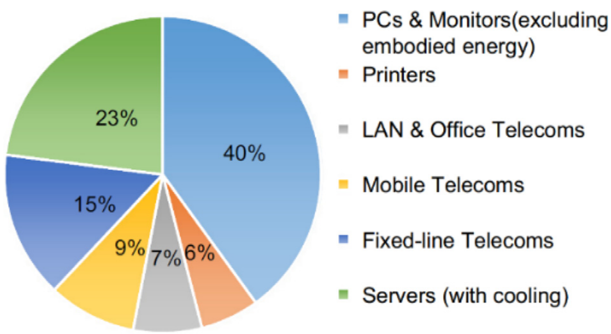

Fig. 2.3. Estimated CO2 emissions for different IT equipment [13].

2.4. Data center energy efficiency

One of the primary goals in all global economic sectors is energy efficiency. DCs represent an enormous and growing energy consumption industry with a substantial global impact. Several energy-saving strategies, policies, and frameworks are proposed to improve DC energy efficiency.

- **European Code of Conduct for DC Energy Efficiency**

The Joint Research Center (JRC) has launched this voluntary project in 2008 to restrict energy consumption in ICT sectors (including DCs) and reduce their associated negative impact on the environment and economy [26]. The goal is to educate and raise awareness among DC owners and operators to minimize DC's energy use without compromising its essential functions. The code of conduct for energy conservation focuses on the energy consumption of IT and facility loads. DC operators are advised to undertake an energy audit of their facilities to find potential energy savings. According to the 2021 European Code of Conduct on Data Center Energy Efficiency [27], DC operators must benchmark their efficiency and provide proof of efficiency growth over time.

• **Energy Efficiency Directives by European Parliament**

The EU Parliament's Directive (EU) 2018/844 lays out strategies for improving DC facilities energy efficiency considering various local and climatic factors. It provides guidelines and a standard framework for measuring parameters on consumption chain and management systems. It aims to promote smart technologies in DCs and cost-effective renovation of existing DCs. The European Commission has devised a methodological framework that the EU member states must adhere for measuring optimal energy performance requirements [28].

• **Energy Conservation and Optimization at different Granularities in DC**

This section summarizes existing approaches and strategies for maximizing DC energy efficiency while limiting its negative environmental impact. Additionally, several key elements are highlighted for energy efficiency while maintaining the required QoS – Quality of Service – to satisfy user expectations. The most significant aspects that may increase the energy efficiency of DC computing capability are instantaneous power usage and total energy consumption. Power can be conserved at different granularity levels, such as per job/task, per node etc.

◦ **Server Level**

According to research [29], the utilization rate of server resources in DCs only reaches 20% of its potential performance. Therefore, appropriate resource scheduling can effectively reduce server clusters energy usage and maintain a low idle rate [30].

<u>Resource Scheduling and Optimization:</u> In energy-saving research, dynamic voltage scaling (DVS)-based power-aware work scheduling is commonly employed. The effectiveness of conventional job schedulers may be further enhanced by real-time monitoring of energy consumption and forecasting power needs [9].

<u>Efficient Resource Management Systems:</u> Monitoring and reducing power use in accordance with task requirements also increases energy efficiency. The system performance is optimized by intelligently allocating workloads to available nodes based on the energy requirements of each application and the power capacity of computing resources [31].

<u>Low-Power Servers Design:</u> Improves performance due to the specific internal configuration of core components and optimization of processors and storage structure. It also consumes less energy than typical servers in DCs and offers sufficient computing power which is ideal for power-saving operations [32].

◦ **Rack Level**

Poor cooling systems in DCs often cause premature server failures and poor performance, eventually leading to a rise in energy consumption and operational costs. According to research by Nada and Elfeky [33], placing high power density servers in the middle rack may result in optimal thermal performance and energy efficiency. Tolia [34] describes a model-based approach that employs fan power management as a control technique for the server's energy utilization effectiveness. Thermal-aware task scheduling such as predictive modeling-based scheduling also optimizes system utilization [35]. An adaptive control system for the dynamic allocation of computing resources to IT loads to balance the computational load helps reduce energy consumption [36].

◦ **Data Center Level**

The primary approach for improving energy efficiency at the DC level is to divide the controlling systems into cooling and powering zones. The goal is to apply various appropriate controlling strategies in different zones and integrate all actions for the zone synergy [7]. A control method is presented by Rao and colleagues [9] that distributes the IT load across the servers with the lowest risk of high inlet temperatures. This approach enhances a cooling system's effectiveness, eventually reducing the cooling power usage. A thermal prediction model described in [37] enables pre-emptive over-cooling of DC to take advantage of time-varying power costs during the day [38].

2.5. Data center energy management

As discussed in the previous section, the four major categories of energy-consuming systems in the DC include IT, cooling, power, and other miscellaneous sub-systems. The contribution of this systems to total DC energy consumption is 45%, 40%, 10%, and 5% respectively [3]. Several cutting-edge technologies and methodologies for energy conservation and efficiency optimization in DC have been developed which are discussed in the next section.

• **Optimal Workload Management in DCs**

Numerous studies have been conducted on workload management, but only few of them have considered energy-efficiency at node-level processes. Zhu and colleagues [39] have provided a framework of delay-tolerant workload distribution which uses a Mixed Integer Linear Programming (MILP) model. This model calculates the load ratio of computational nodes and considers optimization of both the power supply side and the demand response side which results in saving 39.7% of operation cost. An energy-efficient task scheduling system is presented in [40] for UPS nodes

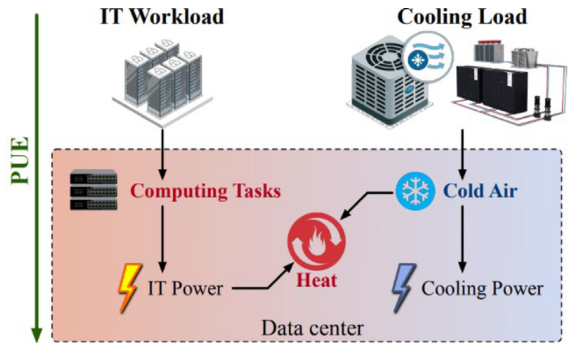

Fig. 2.4. Illustration of IT system and cooling system coupling in data center [50].

to reduce energy consumption by flexible allocation decisions. To anticipate overall DC power usage, a new framework for self-aware workload forecasting is introduced in a recent study [41] where the most relevant features of the ICT system are dynamically selected and applied to a model for the prediction of total power consumption. The study by Xue and colleagues [42] presents a neural network-based framework to forecast future resource utilization and power consumption. The proposed strategy efficiently and accurately predicts future loads and peak loads. An innovative technique to automatically determine the best model for an accurate prediction of DC resource utilization is proposed by Baig and colleagues [43]. The framework trains classifiers based on the pertinent statistical aspects of historical resource consumption. The suggested technique improves prediction results from 6% to 27%. Yi and colleagues [44] present a job allocation algorithm for long-lasting and compute-intensive jobs using deep reinforcement learning (DRL) which results in saving more than 10% DC computing power.

- **Optimal Thermal Energy Management in DCs**

A DC comprises several equipment including servers, storage, and networking devices etc. Heat dissipation from multi-core processing units or cooling machines contributes to increased facility temperatures in data centers. A recent study [45] experimentally investigates a compact two-phase loop cooling application system to deploy high-density rack-mount server cooling applications. With this methodology, the system can accommodate a high heat flux range of up to 22.22 W/cm2, showing excellent potential for cooling server rack-mount enclosures. Song and colleagues [46] have presented a framework of multi-tiers thermal intelligent workload placement which evaluates the thermal conditions of servers within the cluster using CPU DTS (Digital Thermal Sensor) thermal margin. Sarkinen and colleagues [47] have proposed a holistic air-cooling strategy to minimize DC energy usage by synchronizing DC server fans and facility fans. Their results reveal that lowering server inlet temperatures minimizes energy consumption and Power Usage Effectiveness (PUE). Authors in [48] have presented a data-driven machine learning method which uses several regression models to predict the host ambient temperature based on the host thermal behavior properties. The suggested approach has increased energy savings by 34.5% and decreased average peak temperature by up to 6.5%. In a recent study by De Chiara, Chinnici, and Kor [49], the researchers have proposed a data mining algorithm for locating hotspot areas in the server room. The procedure entails grouping nodes into clusters based on their thermal ranges (including hot, regular, etc.) and identifying hotspot zones.

- **Joint Workload and Thermal Management in DCs**

Zhang and colleagues [50] have identified heat as a component that links cooling systems and IT operations(see Fig. 2.4).

They have conducted a survey on joint optimization of DC IT operations and cooling system to provide a comprehensive discussion and comparison between different technologies. Based on their findings, they concluded that the learning-based approach is a promising framework for joint DC ICT and cooling management. In a recent study by Mirhoseini Nejad and colleagues [51], they have proposed a novel low complexity holistic DC model considering controlling parameters of cooling units in conjunction with the thermal effects of server's workload. Results have shown that combining workload scheduling and cooling factors saves more power than individually optimizing each of them. Another optimization approach based on Deep Reinforcement Learning (DRL) called DeepEE developed by Ran and colleagues [52], is for improving DC energy efficiency while concurrently considering IT and cooling systems. In contrast to baseline joint optimization methodologies, the findings reveal that this technique saves 15% energy consumption while maintaining service quality.

3. Proposed methodology

This section furnishes a holistic overview of the applied methods to achieve the goal of this research. As discussed in the previous section, DC sustainability is primarily attained by pursuing energy efficiency in all operational aspects, including workload processing efficiency, optimal resource allocation, and suitable cooling technologies. Hence, the

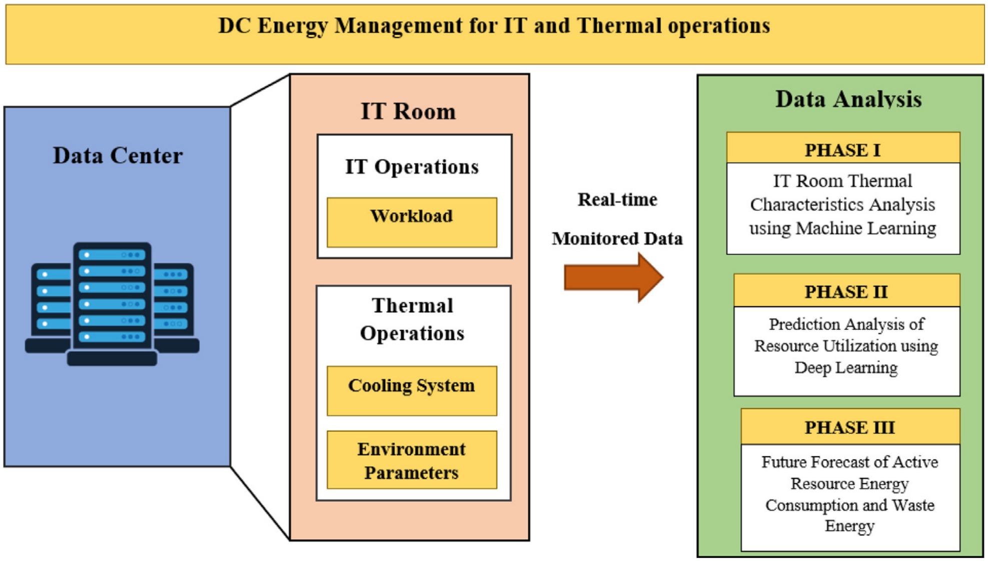

Fig. 3.1. Three phases methodology for DC energy management.

Fig. 3.2. HPC CRESCO6 in ENEA-R.C. Portici (courtesy of Davide De Chiara).

primary emphasis of this study is to analyze and investigate techniques for efficient DC energy management. To reiterate, the IT and cooling systems are viewed as two significant contributors to DC energy consumption. We have illustrated a holistic approach in Fig. 3.1 employed in this study where the DC facility monitored data has been used to build advanced analytical models. Phase 1 presents thermal characteristics analysis of the DC IT room to provide an optimal operational environment for the IT equipment. Phases 2 and 3 extends the analysis by experimenting on IT operations which involves resource utilization prediction and future forecasting of resources energy consumption based on system workload.

This research is conducted in collaboration with the ENEA-Research Center (RC), Portici in Italy. The real-time data is obtained from the High-Performance Computing (HPC) cluster CRESCO6 operating in the ENEA-R.C. Portici Italy (see Fig. 3.2). In this study, we have used four individual datasets collected from different sources within the DC IT room for the entire year of 2020. The analysis and preprocessing of all the datasets are provided in (Khan, De Chiara, Kor and Chinnici) [53] and are further extended in this work using AI techniques for the management of both IT and thermal operations [54,55].

The following sections provide an in-depth discussion of experiments performed in each phase. The work in each phase is organized as; 1. Problem identification and a proposed solution; 2. Data preparation for the experimentation; 3. Data analysis for insights; 4. Model application on the analyzed data.

3.1. Phase I: IT room thermal characteristics analysis using machine learning

In this phase, we examine the IT room thermal conditions to reduce excessive energy usage through thermal characteristics analysis followed by identification and control of high energy consumption related factors. For instance, it is known that high and long-term equipment utilization can effect IT equipment performance degradation especially

Table 1
Thermal guidelines by ASHARE for DC [22].

Class	IT equipment type	Recommended operating range	Allowable operating range	Maximum dew point	Environmental control
A2	Volume servers, storage products, personal computers, workstations	18° to 27 °C	10 to 35 °C	21 °C	Some control

when DC servers operate 24/7. For this reason, it is necessary to analyze the IT room thermal conditions to ensure IT equipment optimal use while maintaining their reliability.

• Data Preparation for IT Room Thermal Analysis

Based on the monitored data of CRESCO6 cluster at ENEA R.C. Portici, we have manipulated the available monitored data to produce several new features. The IT room in the DC has several installed temperature sensors, and the room's area is comparatively smaller compared to the entire DC. To address the complexity of handling multiple values for the same feature, we have used all the sensor mean values for the same timestamp. The description of all data features is provided in [53]. The calculations for new feature values are as follows:

1. There are ten built-in fans on the servers, we have considered their average speed for further analysis.
   So, average of node fan speeds:
$$\text{Avg. fan speed} = (\text{Fan1a} + \text{Fan1b} + \text{Fan2a}\ldots + \text{Fan5b})/10 \tag{1}$$
2. Average of CPU temperature:
$$\text{Avg CPU Temp} = (\text{Temperature of CPU1} + \text{Temperature of CPU2})/2 \tag{2}$$
3. Resource Utilization percentage:
$$\begin{aligned}\text{Resource Utilization percentage} &= \text{System utilization percentage} + \text{Network Utilization percentage}\\ &\quad + \text{Memory utilization percentage}\end{aligned} \tag{3}$$
4. Resource Consumption Power:
$$\text{Resource Consumption Power} = \text{System power} + \text{Memory power} \tag{4}$$

The same strategy is applied for the environment dataset provided in [53]. As there are several temperature and humidity sensors placed in the hot and cold aisles of the room, so we have considered the average of all sensor values.

• Guidelines by ASHRAE TC 9.9 on DC Operational Environment

The servers operating in DCs produces a large amount of heat during processing. If the environmental conditions are not monitored and maintained, it may cause equipment performance degradation. Hence, maintaining room temperature and humidity according to recommended operational environment is necessary for equipment performance efficiency. According to the guidelines by ASHRAE TC 9.9 [22], we have defined multiple thermal classes based on the recommended allowable ranges for an optimal processing environment for IT equipment.

1. We first analyze the class of the equipment installed in the DC and categorize them based on the classes specified by ASHRAE Thermal guidelines.
2. According to the class of equipment, we follow the recommended and allowable ranges for optimal data processing environment as defined by the ASHRAE TC 9.9 [22] (see Table 1).
3. We have created five thermal classes based on the ASHRAE guidelines as shown in Fig. 3.3.

• Thermal Analysis of DC IT Room

To visualize and analyze the room's thermal characteristics, multiple features including computing server node temperature, average processor temperature, and environment temperature are plotted. Additionally, we can observe from Fig. 3.4 that CPU temperature is significantly higher than the overall node temperature. It is because the CPU has comparatively larger range of operating temperatures and dissipates more heat during computation [56].

The operational environment of the IT room continues to remain in acceptable temperature ranges. However, it sometimes reaches over-temperature conditions as it crosses the acceptable threshold as shown in Fig. 3.5. We have

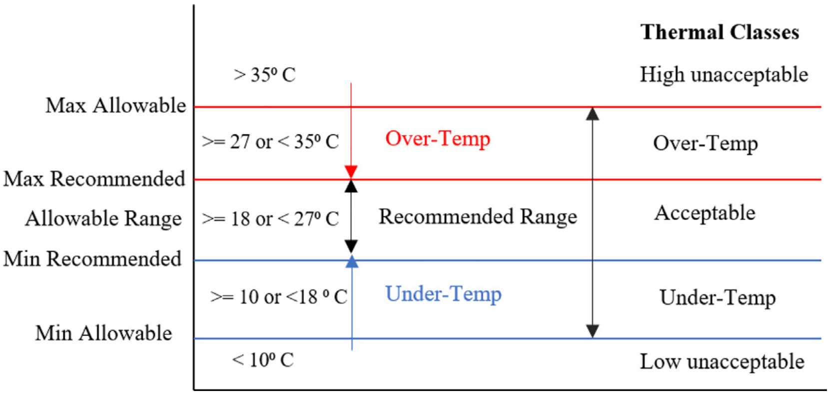

Fig. 3.3. Thermal Classes Guidelines by ASHRAE [22].

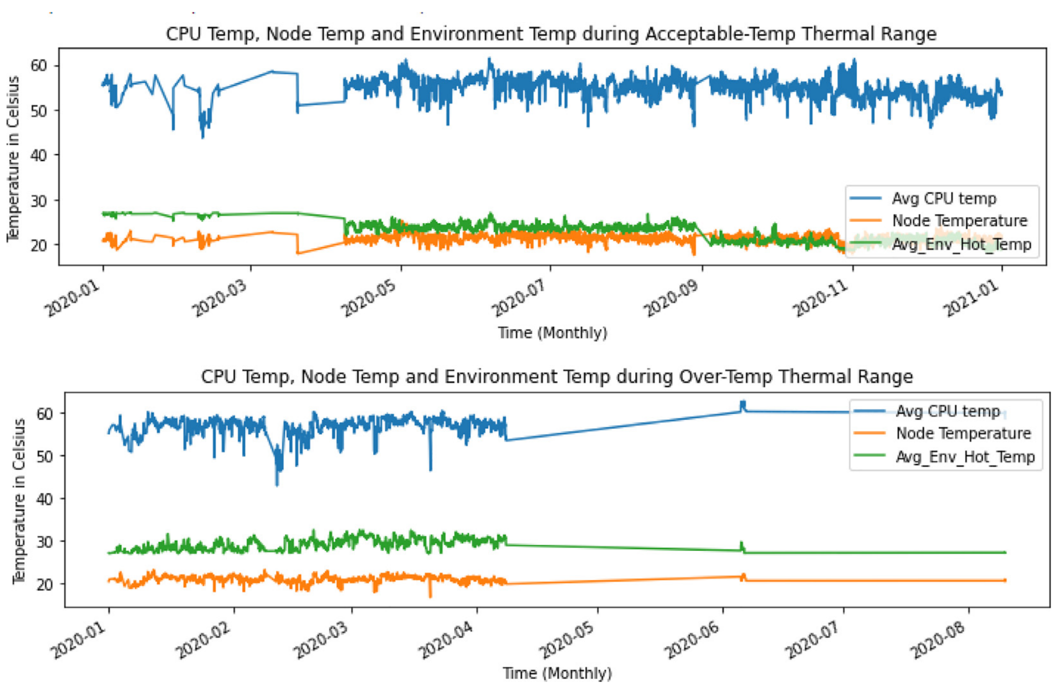

Fig. 3.4. Monthly analysis of thermal conditions of IT Room.

conducted a monthly analysis on servers, cooling systems, and environmental thermal parameters to determine the underlying cause.

Fig. 3.5 shows that even if the server is operating normally, the ambient temperature increases though the temperature of the compute node does not rise. Thus, we speculate that it could be due to the under-performing cooling system. A monthly analysis of the cooling system performance is conducted and depicted in Fig. 3.6.

We can observe in Fig. 3.6 that when the ambient temperature rises, the usage (percentage) of a cooling machine for the supply of cool air is low during that period. It demonstrates that the ineffective cooling in the IT room is the contributor to the unfavorable thermal conditions. Furthermore, fluctuations in the first few months (Fig. 3.6) shows that the AC machine supplies cool air to servers but for quite short periods of time which causes the peaks in the graph.

- **Machine Learning Modeling of IT Room Thermal Parameters**

We have observed the trend and patterns of IT room thermal conditions throughout the year. We have included thermal data from [53] and additional calculated data features as input features to the supervised machine learning classification models. The goal of this phase is to predict the thermal class of IT room for efficient DC IT operations. We have applied five different supervised machine learning classification models including Logistic regression, Decision tree, Support Vector Machine, Random Forest, and Gaussian Naïve Bayes for the prediction of thermal classes. For the evaluation of models,

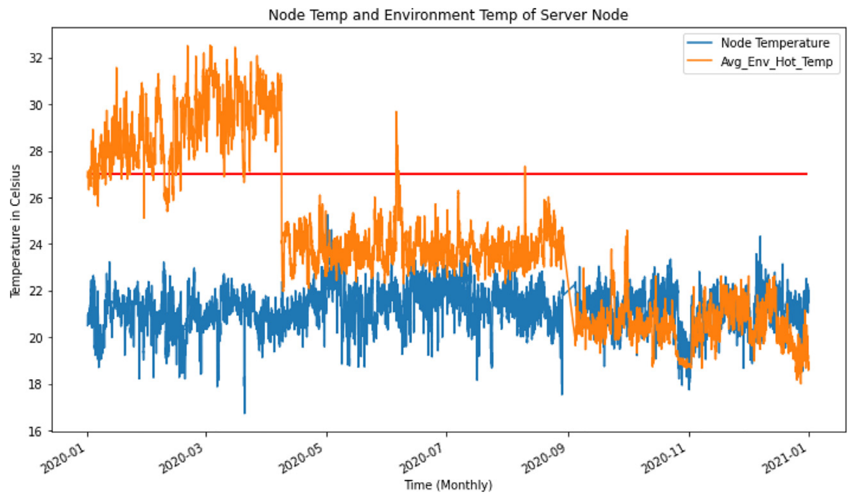

Fig. 3.5. Monthly analysis of thermal parameters of IT room.

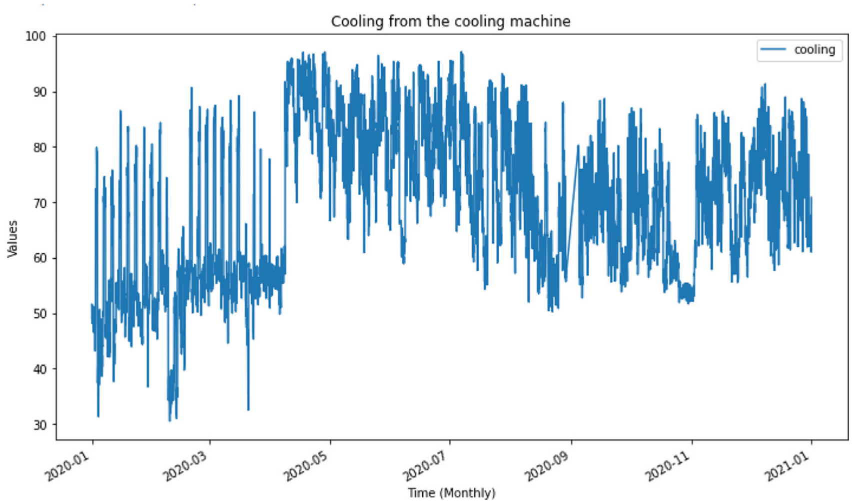

Fig. 3.6. Monthly analysis of cooling system performance.

we have used different performance metrics. Also, we have conducted a comparative analysis amongst the ML models. The results for the ML model are discussed in the Results and Discussion Section.

3.2. Phase II: Prediction of resource utilization using deep learning algorithm

In this phase, we focus on DC IT operations to examine relevant resource utilization behavior which includes CPU and memory usage, during job execution period and their corresponding energy consumption over a period of one year. In ENEA Portici HPC CRESCO6 cluster, LSF (Load Sharing Facility) scheduler is used. It is a framework for managing workloads and scheduling jobs in distributed HPC systems using FCFS — First come First Serve approach. Since CRESCO6 has more than 400 nodes, it distributes the work to a first available free node.

The simple linear regression model is a known technique for predicting a DC energy consumption. It is best suited for CPU-dominated servers with moderate usage and consistent power consumption. However, an advanced level predictive model is necessary for the consideration of predictions involving complex data. Thus, we have implemented two different models for advanced prediction of DC resources utilization.

- **Time Series Decomposition of Data Attributes**

The time series decomposition of data provides an insight into data variances [57]. To prepare data for time series decomposition, we first merge the monitored real-time data of CRESCO6 cluster with the estimated data from previous section.

```
   Augmented Dickey-Fuller Test on "Resource Utilization"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -9.6692
No. Lags Chosen       = 22
Critical value 1%     = -3.431
Critical value 5%     = -2.862
Critical value 10%    = -2.567
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.


   Augmented Dickey-Fuller Test on "Resource Consumption Power"
   -----------------------------------------------
Null Hypothesis: Data has unit root. Non-Stationary.
Significance Level    = 0.05
Test Statistic        = -8.7808
No. Lags Chosen       = 23
Critical value 1%     = -3.431
Critical value 5%     = -2.862
Critical value 10%    = -2.567
=> P-Value = 0.0. Rejecting Null Hypothesis.
=> Series is Stationary.
```

Fig. 3.7. Augmented Dickey–Fuller test on data attributes.

An Augmented Dickey–Fuller Test (ADF) [58] is performed for all the data attributes individually. It is a statistical test that determines whether the given time series is stationary or not. We have formulated a null hypothesis which determines the unit root. If the *p*-value is less than 0.05, we will reject the null hypothesis which shows the time series is stationary. From the results shown in Fig. 3.7, we observe that most of the time series data attributes are stationary and exhibit marginal variations in the values that repeat after each time interval.

- **Time Series Forecast using Vector Autoregression (VAR)**

An approach for statistical multivariate forecasting called vector autoregression is used to predict a time series vector [59]. It is often applied when time series data characteristics that need to be forecasted exhibit a correlation between values. It is best applied for variables that are stationary (i.e., mean and variance do not change over time). As we have previously completed time series decomposition to identify seasonality, residue, etc., we can then apply the algorithm on the data attributes. The results for the model application are discussed in the Results and Discussion section.

- **Time Series Forecast using LSTM**

Long Short-Term Memory (LSTM) is an artificial repetitive neural network (RNN) used in deep learning [60]. This model is the best suited for time series prediction as it can store information for a longer period of time than typical RNNs, thus enhancing its prediction accuracy. In our case, the future prediction of resource utilization is solely based on historical data. For this reason, we have used the LSTM model. Its structural elements include a cell, a learn input gate, a use output gate, a remember gate, and a forget gate [60].

3.3. Phase III: Future forecast of DC resource energy consumption and waste energy

In this phase, we experiment with IT operations and aims to predict energy consumption and energy waste for job execution by compute resources. The amount of energy used for server computations depends on the job execution time and percentage of resources usage. The server manages hundreds of jobs daily, thus, we are aiming to forecast overall server energy consumption for all jobs submitted over the course of a full year to the HPC CRESCO6 cluster. Based on the job execution status in the cluster, we have categorized the jobs into two classes i.e., done jobs, which indicate successful task completion, and exit jobs, which indicate unfinished task execution. The graph (Fig. 3.8) displays the frequency of both groups over the period of one year.

We have separated energy consumption into active resource energy consumption and energy waste in accordance with the job categories. This is because every executed job (albeit successful or not) consumes resources as well as energy. In this phase, we have considered three types of energy usage during task execution: energy waste from active resource energy consumption, energy waste due to exit jobs, and overall energy waste (due to missing job data).

- **Data Preparation for Prediction Analysis**

For prediction analysis, we first calculate active resource energy consumption and energy waste based on different features from different datasets provided in [53].

$$\text{Active Resource Energy Consumption} = \text{Resource Consumption Power} * \text{Active Job Execution Time} \tag{5}$$

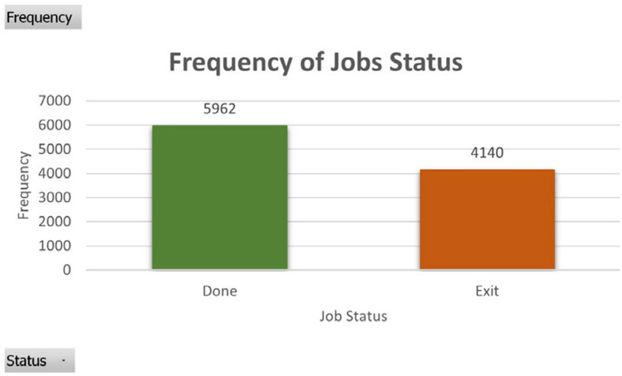

Fig. 3.8. Frequency of Done and Exit Jobs.

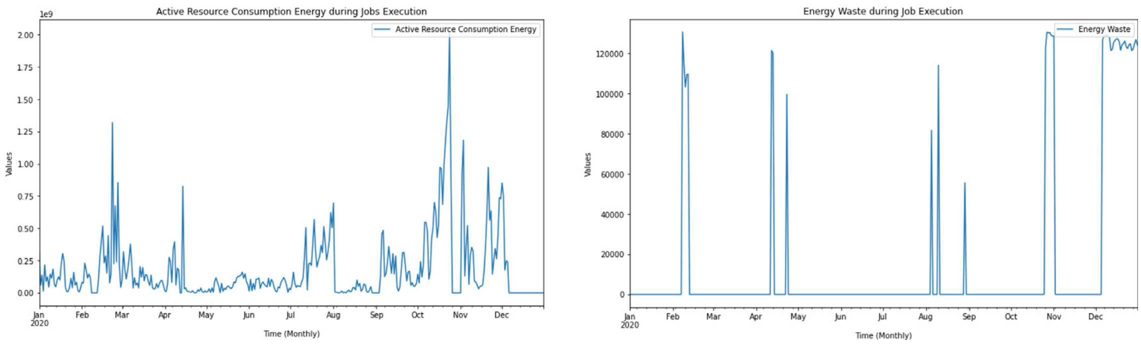

Fig. 3.9. Monthly analysis of active resource energy consumption and energy waste.

$$\text{Energy Waste by Exit} = \text{Resource Consumption Power} * \text{Execution time of Exit Jobs} \tag{6}$$

$$\text{Overall Energy Waste} = (\text{Total Resource Consumption Power} \times 24\ \text{h}) - \text{Active Energy Consumption} \tag{7}$$

The workload data on the compute nodes is provided for the entire year of 2020. However, job data submitted to the cluster is only recorded until early December 2020. As job data is unavailable for a short period, we have decided not to remove resource power consumption data for that period. Instead, we have assumed resource power is being used as usual during the period.

• **Time Series Decomposition of Data Attributes**

The time series decomposition is performed on data to examine active resource energy use and energy waste. As shown in Fig. 3.9, the exit jobs rarely occur compared to completed jobs. We observe some high peaks in the graph for energy waste. Additionally, we discover that most of the exit jobs take very long time (up to two days) for execution but exit in the end, hence, resulting in high futile energy consumption.

The time series analysis is also performed on data attributes which show a cyclic behavior in seasonality within data.

• **Monthly Analysis of Active Resource Energy Consumption and Energy Waste**

We have examined the feature's autocorrelation using the autocorrelation function (ACF) prior to the deep learning modeling. The ACF function provides the autocorrelation value of a series with its lag values [61]. It demonstrates how closely present values relate with the past values, which are crucial for predictions. The prediction accuracy of the model will be low if there is weak correlation within the data values. The data values exhibit non-stationary behavior at some points, as illustrated in Fig. 3.10, which are eliminated using differencing method.

• **Time Series Forecast using SARIMA**

Seasonal Autoregressive Integrated Moving Average, or SARIMA, is a method for forecasting future values based on historical data [62]. Our dataset has shown seasonality trends, thus, this has prompted the selection of the SARIMA model which considers the seasonality aspect of data as well as past values during the prediction of future values. The ADF test is initially run on the data attributes before tuning the model. Additionally, trend values and seasonal parameters are modified to evaluate the model's effectiveness in order to increase performance. The findings of this experimentation are discussed in the Results and Discussion Section.

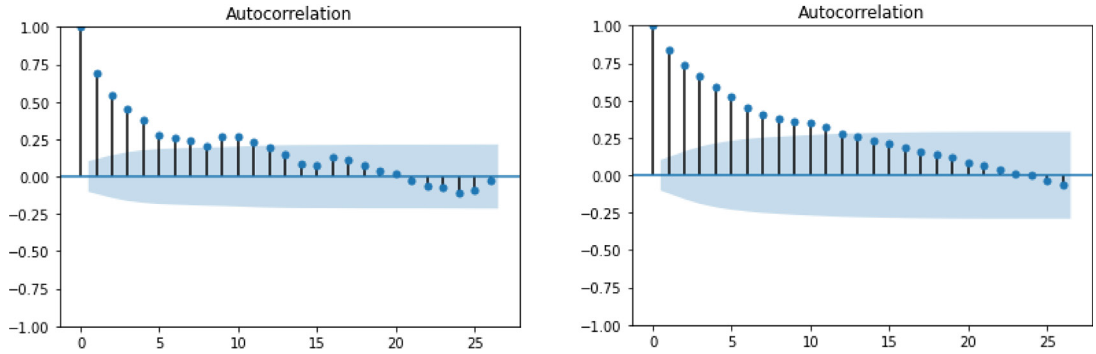

Fig. 3.10. ACF plot of active resource energy consumption and energy waste.

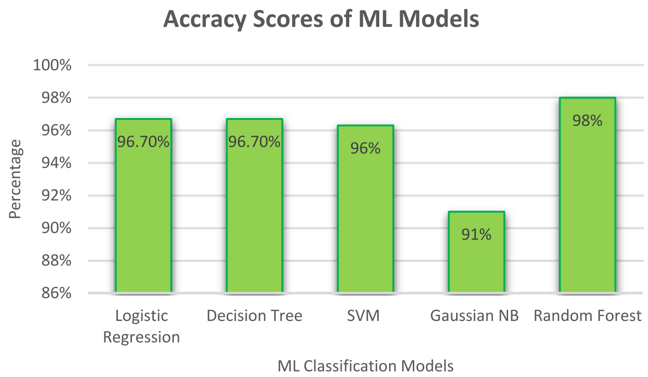

Fig. 4.1. Comparison of ML classification models accuracies.

4. Results and discussion

This section summarizes results for all phases in the research. Furthermore, a comparative analysis is conducted for this research experimental findings and other related existing models.

To reiterate, the IT and cooling systems are considered as two significant contributors to DC energy consumption. The goal of this research is to target energy efficiency operations and management within these two DC systems.

4.1. Results analysis of phase I

In this phase, we have performed thermal characteristics analysis of a DC IT Room using machine learning classification techniques. The thermal conditions of the DC IT room are first analyzed to identify different patterns in monitored data, and then classified into different thermal ranges based on guidelines provided by ASHRAE TC 9.9 [59].

As discussed in the methodology section, five machine learning classification models are applied on the real-time data for the prediction of thermal classes of DC IT room. We have included thermal data from [53] and additional calculated data features as input features to the supervised machine learning classification models. This combined dataset comprises both continuous and categorical data features. The datasets are split into training and testing datasets with the ratio of 80% and 20%. Furthermore, we have only selected input features which have shown high correlation with the target feature as given in [53].

The models' performances are evaluated based on prediction accuracies. To measure the model accuracy, we have used a confusion matrix and accuracy scoring metric. An overall comparison amongst models' accuracy scores is plotted in Fig. 4.1 while Fig. 4.2 depicts precision of all models for each individual thermal class. From the plotted graphs, we can see that Random Forest outperforms all the other four models with the highest true class predictions. However, Logistic Regression and Decision Tree have also shown good performances with the second highest accuracies in the prediction of thermal classes.

The confusion matrix and graph plots in Figs. 4.3 and 4.4 show the number of true predictions and false predictions made by each model. Random forest has only made 25 wrong predictions of thermal classes out of 1566 entities of a test sample. On the other hand, SVM and Decision Tree have made approximately the same number of false predictions. Implementation of an efficient ML modeling technique with an accuracy of 98% for classifying thermal conditions in the IT room concludes the second phase of this work and completes objectives RO1.1 and RO1.2.

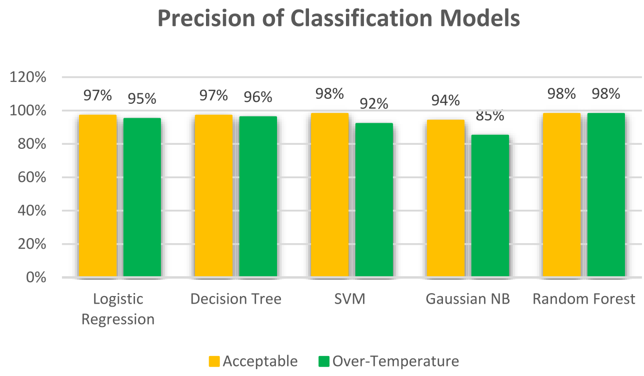

Fig. 4.2. Precision of classification models for thermal classes prediction.

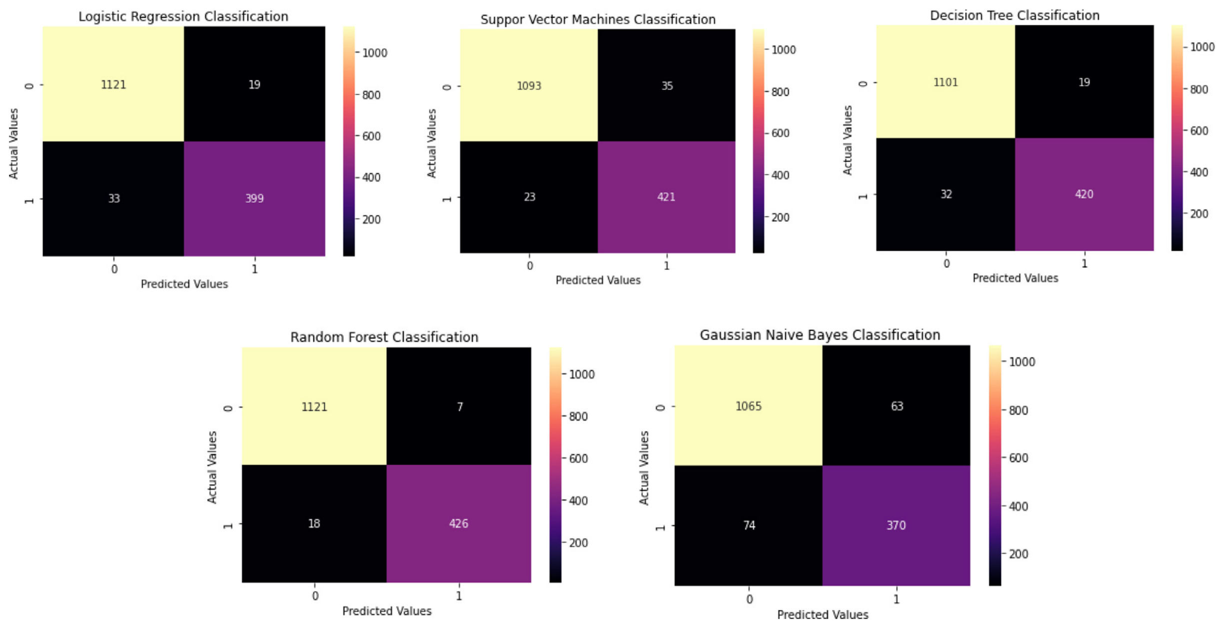

Fig. 4.3. Confusion matrix for ML models evaluation.

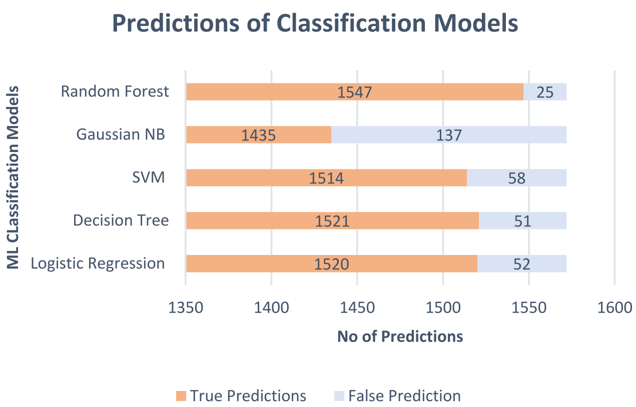

Fig. 4.4. True and false predictions of ML classification models.

In existing literature, Ilager and colleagues [48] have only considered CPU and inlet airflow temperature variations for DC thermal management assessment using Machine Learning algorithms which not only limits the scope of their

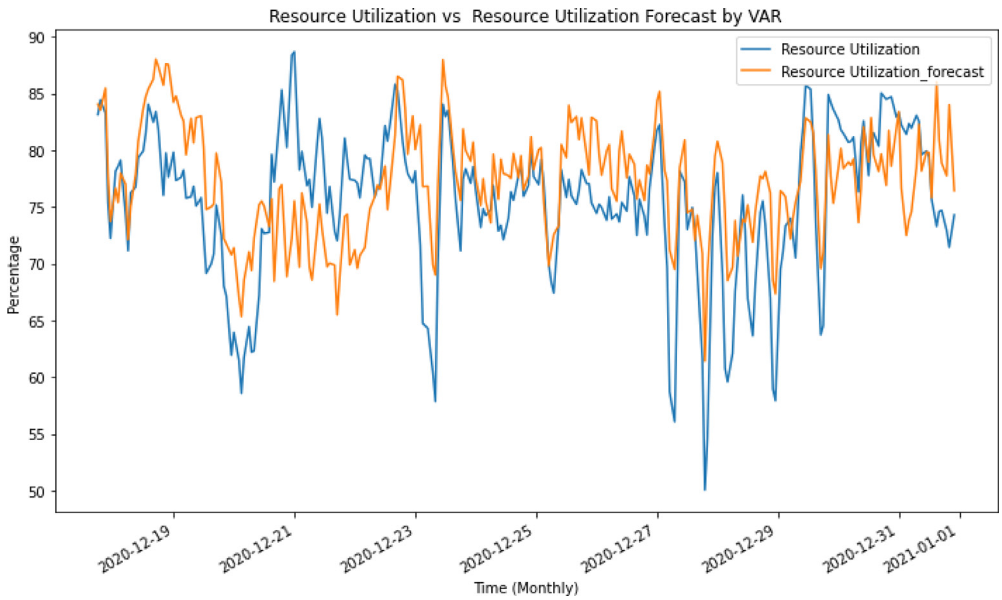

Fig. 4.5. Predicted values vs actual data values by VAR.

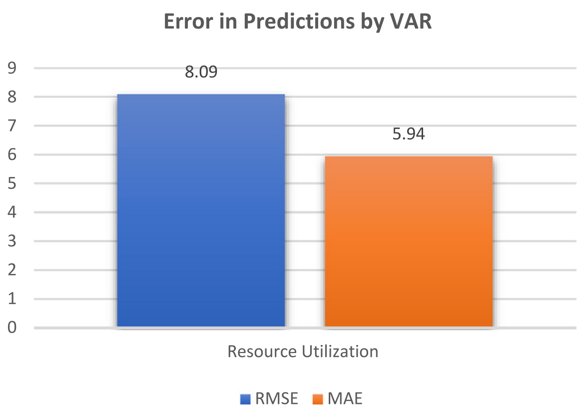

Fig. 4.6. Prediction error percentage by VAR.

study but also reduces prediction accuracy. On the other hand, in this research, we have considered all the parameters for experimentation including computing devices, environmental conditions, and cooling machines which directly impact the DC thermal characteristics. The integration of these features has provided a holistic approach for data-driven temperature estimation of IT room in a data center. As a result, optimal thermal management with accurate temperature prediction can reduce the operational cost of a data center and increase equipment reliability.

4.2. Results analysis of phase II

In this phase, we have aimed to perform advanced prediction of DC resource utilization for efficient workload placement. There are two techniques implemented for prediction analysis including deep learning modeling (i.e., long short-term memory (LSTM)) and statistical data modeling (i.e., vector autoregressive model (VAR)) to observe the variations in the results. To achieve efficient performance and accurate results, the models are trained with different tuning parameters. The performances of both models are evaluated using root mean square error and absolute error as the features to be predicted are continuous numerical data.

The Results for the prediction of resource utilization by the VAR model are shown in Fig. 4.5. Furthermore, the graph plot in Fig. 4.6 depicts the root mean square error for the prediction of different features.

As LSTM has a storage capability for longer periods and a feedback system that has made it best suited for such prediction problem. The model is only applied for prediction analysis of DC resource utilization. However, it can be used

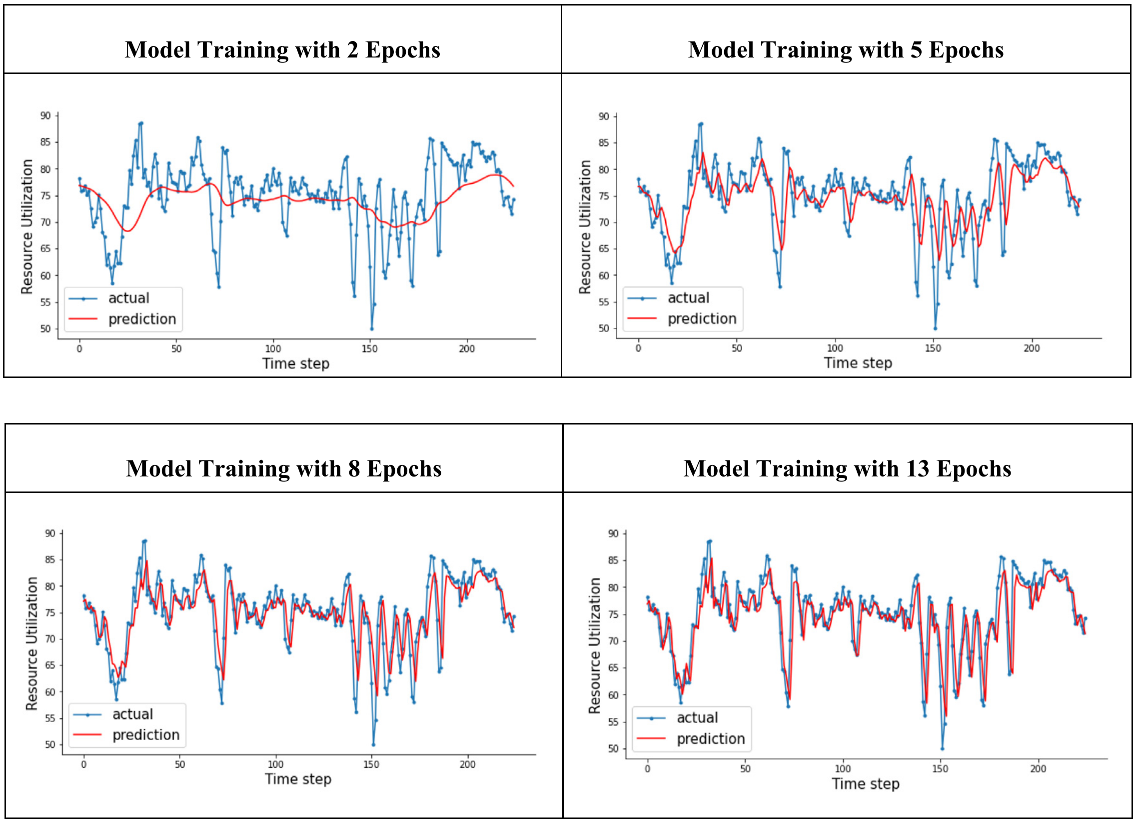

Fig. 4.7. Predicted Values vs actual Values using LSTM with different epochs.

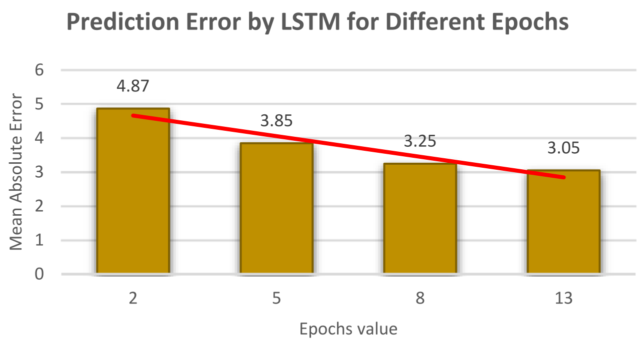

Fig. 4.8. Prediction error percentage by LSTM.

for other data features as well. The performance of LSTM model is enhanced further by epoch and batch size tuning. The batch size indicates how many data samples will be used for training while epochs is a hyperparameter that indicates the number of times the algorithm will run and iterate over the entire training dataset. Additionally, on testing the model with different values of epochs, the accuracy of the model has a positive trend. To prevent over-fitting, we have stopped at 13 epochs for model training. The graphs in Fig. 4.7 show the LSTM predicted values vs actual values with different hyperparameter tuning. The predicted values are relatively more accurate and closer to actual values as compared to previous VAR results. This model performance is also evaluated using mean absolute error, as seen in Fig. 4.8. The LSTM model prediction error reduces with increasing epochs which evidences enhanced model performance.

In existing literature, Yi and colleagues [44] have performed advance prediction of resources utilization using LSTM. However, the study merely focuses on CPU utilization excluding memory and network utilization. In this research, we have considered all possible parameters including memory and network utilization along with CPU utilization to analyze the overall behavior of resource consumption. Additionally, they [44] have performed prediction analysis by clustering the dataset which causes inconsistent accuracy results for different clusters. On the other hand, the experiments conducted in

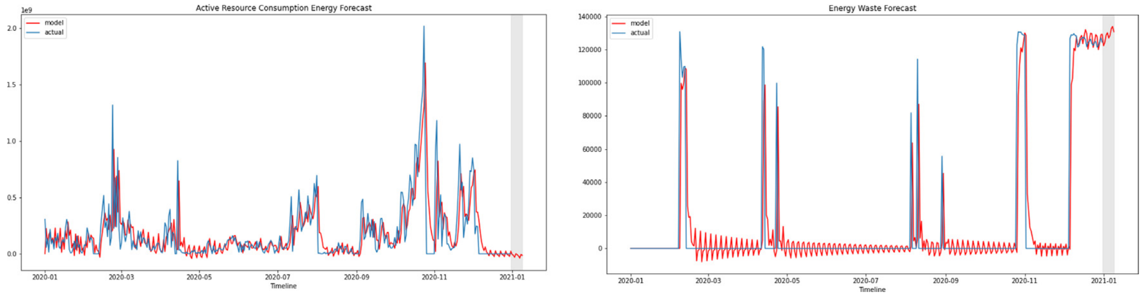

Fig. 4.9. Future forecasting of active resource energy consumption and waster energy by SARIMA.

this research focus on one entire dataset to overcome inconsistency in the results. Furthermore, in comparison with [45] which involves the implementation of several ML algorithms for the prediction of resource utilization results in a higher root mean square error of up to 3.32. This study also focuses on CPU utilization as an observed metric. The system architecture of this research includes window size sensitivity for training prediction models based on time series features of recent window results varying prediction accuracies. However, the design of our research experiment provides a feedback approach to learn the patterns of resource utilization effectively for achieving optimized results.

In conclusion, we can say that this research has provided an efficient deep learning modeling technique (LSTM) for the prediction of resource utilization with the highest accuracy and minimal error rate, thus addressing research objectives RO1.3 and RO1.4.

4.3. Results analysis of phase III

The objective of this phase is to perform future forecasting of DC's active resource energy consumption and waste energy while considering the sustainability aspect. We have merged data from [23] with newly calculated data features to estimate energy consumption by resources during job execution. Our dataset has shown seasonality trends; thus, this has prompted the selection of the SARIMA model which considers the seasonality aspect of data as well as past values during the prediction of future values. It analyzes the seasonal patterns of the data for prediction of active resource energy consumption and energy waste as shown in Fig. 4.9. Therefore, we could conclude that all three phases of this research work have successfully answered our RQ1.

5. Conclusion

Energy efficiency in IT systems ought to be the ultimate approach for a data center (DC) with a sizable high-performance computing facility to achieve sustainable development goals. IT and cooling systems are considered the common areas within a DC which performs the most energy intensive operations. Therefore, the primary emphasis of this study is on: IT system energy efficiency and effective thermal conditions in DC.

To assess the energy consumption of various IT and thermal operations in the ENEA Portici HPC CRESCO6 cluster for efficient energy management, this research work has been divided into three phases. It is known that high and long-term IT equipment utilization under poor operational conditions can degrade the equipment's performance, particularly for DC servers which operate 24/7. For this reason, it is necessary to maintain the operational environment of IT systems for efficient usage of the equipment while maintaining their reliability. Considering this issue, Phase I emphasizes on thermal characteristics analysis of DC as a vital aspect of the operational environment. Data is collected from different sources within the DC (including environmental parameters, equipment usage and cooling supply) where only input features which have shown high correlation with the target feature are selected for experimentation. For thermal class prediction analysis based on ASHRAE guidelines, five supervised machine learning classification models are applied to the processed data. The performances of these models are compared where the Random Forest classification model has outperformed the other models with the highest prediction accuracy of 98% which would help in effectively maintaining the operational environment. Due to inefficient workload management, the resources in DC are sometimes under and overutilized which causes energy waste. The advanced prediction of resources utilization would help in effectively managing the workload and reduces energy waste. Reflecting on this, phase II provides advanced prediction of DC's resource utilization using two modeling techniques i.e., deep learning and autoregression. Based on different patterns and characteristics of the data, LSTM — deep learning model due to its a storage capability of historical data for longer periods and a feedback system has resulted in high prediction accuracy and the least percentage error of 3.05 MAE as compared to autoregressive modeling results. Furthermore, failed jobs execution in DC cause resource and energy wastage. In response to this issue, the forecast of the active energy usage and wastage due to successful and failed jobs is investigated in phase III. Overall, this research work provides state-of-the-art techniques for evidence-based DC energy management and a comparative analysis in terms of reliability, consistency, and prediction accuracy. The applicability of the proposed methods to other DC datasets depends on DC monitoring systems data, accuracy of data collection, and individual characteristics of DCs. Sustainability and energy efficiency goals in DCs can be achieved by integrating the proposed methods and techniques of this study in real-time DC operations.

6. Recommendations and future work

Improvement in job scheduling techniques can help in efficient workload management and resources utilization. Undeniably, inefficient workload placement causes under or overutilization of resources which ultimately result in energy waste. The server's operating power should be monitored and improved based on thermal conditions of the IT room environment. The long-term and high utilization of resources causes performance degradation of IT systems. To maintain the performance of servers and their reliability, servers should be turned off completely after working certain hours instead of transforming them into IDLE mode. Monitoring of server energy consumption based on their technical requirements and power specifications should be considered to determine whether the IT equipment is using normal power or more than it requires. Failed job execution tends to consume more energy as it not only wastes power but also keeps the resources occupied from executing a productive job. Several studies have proposed different methods to deal with such problems but an intelligent solution that can detect the behavior/requirements of the submitted job before allocating the resources can help in saving energy. The operational conditions for an IT room environment should be monitored all the time to examine if the cooling supply is effective. As discussed in this paper, an IT room is over-heated due to insufficient cooling supply. An intelligent system should integrate both the IT systems and cooling systems that can trigger or notify the DC operator in the event the cooling system or IT system is not working, or if the room temperature conditions are not acceptable.

CRediT authorship contribution statement

Wania Khan: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Davide De Chiara:** Data curation, Resources. **Ah-Lian Kor:** Supervision, Validation, Review. **Marta Chinnici:** Project administration, Supervision, Resources, Review.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgment

This research has been supported by EMJMD GENIAL (610619-EPP-1-2019-1-FR-EPPKA1-JMD-MOB).

References

[1] H. Hu, Y. Wen, Y. Gao, T.-S. Chua, X. Li, Toward an SDN-enabled big data platform for social TV analytics, IEEE Netw. 29 (5) (2015) 43–49, http://dx.doi.org/10.1109/mnet.2015.7293304.
[2] Y. Wen, X. Zhu, J.J.P.C. Rodrigues, C.W. Chen, Cloud mobile media: Reflections and outlook, IEEE Trans. Multimed. 16 (4) (2014) 885–902, http://dx.doi.org/10.1109/tmm.2014.2315596 [online].
[3] IEA, Data centres and data transmission networks – Analysis, 2021, [online] Available at: https://www.iea.org/reports/data-centres-and-data-transmission-networks.
[4] Artificial Intelligence, Big Data Analytics, IoT, Intelligent connectivity, 2021, pp. 211–222, http://dx.doi.org/10.1002/9781119685265.ch11.
[5] M. Dayarathna, Y. Wen, R. Fan, Data center energy consumption modeling: A survey, IEEE Commun. Surv. Tutor. 18 (1) (2016) 732–794, http://dx.doi.org/10.1109/comst.2015.2481183.
[6] E. Masanet, A. Shehabi, N. Lei, S. Smith, J. Koomey, Recalibrating global data center energy-use estimates, Science 367 (6481) (2020) 984–986, http://dx.doi.org/10.1126/science.aba3758 [online].
[7] M. Eriksson, Monitoring, modelling and identification of data center servers, 2018, [online] Available at: http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1216563&dswid=1623.
[8] P. Johnson, T. Marker, Data centre energy efficiency product profile, 2009, [online] Available at: https://www.energyrating.gov.au/sites/default/files/documents/200905-data-centre-efficiency_0.pdf.
[9] H. Rong, H. Zhang, S. Xiao, C. Li, C. Hu, Optimizing energy consumption for data centers, Renew. Sustain. Energy Rev. 58 (2016) 674–691, http://dx.doi.org/10.1016/j.rser.2015.12.283.
[10] Cisco, Cisco Annual Internet Report - Cisco AIR Infographic (2018-2023). [online] Available at: https://www.cisco.com/c/en/us/solutions/executive-perspectives/annual-internet-report/infographic-c82-741491.html.
[11] Synergy Research, The total number of large data centers in the world in 2017 exceeded 390, 2018, [online] Available at: http://www.199it.com/archives/670459.html.

[12] Z. Cao, X. Zhou, H. Hu, Z. Wang, Y. Wen, Toward a systematic survey for carbon neutral data centers, IEEE Commun. Surv. Tutor. 24 (2) (2022) 895–936, http://dx.doi.org/10.1109/comst.2022.3161275.

[13] R. Kumar, L. Mieritz (Eds.), Conceptualizing 'Green' IT and data center power and cooling issues, 2007, [online] Gartner. Available at: https://www.gartner.com/en/documents/519717. (Accessed 5 June 2022).

[14] B. Whitehead, D. Andrews, A. Shah, G. Maidment, Assessing the environmental impact of data centres part 1: Background, energy use and metrics, Build. Environ. 82 (2014) 151–159, http://dx.doi.org/10.1016/j.buildenv.2014.08.021.

[15] Z.-J. Zhang, C.-F. Lai, H.-C. Chao, A green data transmission mechanism for wireless multimedia sensor networks using information fusion, IEEE Wirel. Commun. 21 (4) (2014) 14–19, http://dx.doi.org/10.1109/mwc.2014.6882291.

[16] Q. Liu, Y. Ma, M. Alhussein, Y. Zhang, L. Peng, Green data center with IoT sensing and cloud-assisted smart temperature control system, Comput. Netw. 101 (2016) 104–112, http://dx.doi.org/10.1016/j.comnet.2015.11.024.

[17] S.E. Bibri, J. Krogstie, Smart sustainable cities of the future: An extensive interdisciplinary literature review, Sustainable Cities Soc. 31 (2017) 183–212, http://dx.doi.org/10.1016/j.scs.2017.02.016 [online].

[18] A. Capozzoli, G. Serale, L. Liuzzo, M. Chinnici, Thermal metrics for data centers: A critical review, Energy Procedia 62 (2014) 391–400, http://dx.doi.org/10.1016/j.egypro.2014.12.401.

[19] J. Bi, K. Zhang, H. Yuan, Workload and renewable energy prediction in cloud data centers with multi-scale wavelet transformation, in: 2021 29th Mediterranean Conference on Control and Automation, MED, 2021, http://dx.doi.org/10.1109/med51440.2021.9480336.

[20] K. Zhang, Y. Zhang, J. Liu, X. Niu, Recent advancements on thermal management and evaluation for data centers, Appl. Therm. Eng. 142 (2018) 215–231, http://dx.doi.org/10.1016/j.applthermaleng.2018.07.004.

[21] A. Rosa, L.Y. Chen, W. Binder, Predicting and mitigating jobs failures in big data clusters, in: 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 2015, http://dx.doi.org/10.1109/ccgrid.2015.139.

[22] ASHRAE TC 9.9, Equipment thermal guidelines for data processing environments, 2015, [online] ashare.org. Available at: https://www.ashrae.org/File%20Library/Technical%20Resources/Bookstore/datacom1_4th/ReferenceCard_7-25-16.pdf.

[23] M. Dayarathna, Y. Wen, R. Fan, Data center energy consumption modeling: A survey, IEEE Commun. Surv. Tutor. 18 (1) (2016) 732–794, http://dx.doi.org/10.1109/comst.2015.2481183.

[24] K.M.U. Ahmed, J. Sutaria, M.H.J. Bollen, S.K. Ronnberg, Electrical energy consumption model of internal components in data centers, in: 2019 IEEE PES Innovative Smart Grid Technologies Europe, ISGT-Europe, 2019, http://dx.doi.org/10.1109/isgteurope.2019.8905595.

[25] L. Alsbatin, G. Oz, A.H. Ulusoy, An overview of energy-efficient cloud data centres, in: 2017 International Conference on Computer and Applications, ICCA, 2017, http://dx.doi.org/10.1109/comapp.2017.8079789.

[26] I.K.M. User, 2021 Best practice guidelines for the EU code of conduct on data centre energy efficiency, 2021, [online] e3p.jrc.ec.europa.eu. Available at: https://e3p.jrc.ec.europa.eu/publications/2021-best-practice-guidelines-eu-code-conduct-data-centre-energy-efficiency.

[27] European code of conduct on data centre energy efficiency introductory guide for all applicants, 2021, [online] Available at: https://e3p.jrc.ec.europa.eu/sites/default/files/documents/publications/jrc124057_jrcjrc124057_european_code_of_conduct_on_data_centre_energy_efficiency_final_1.pdf. (Accessed 6 June 2022).

[28] eur-lex.europa.eu., Directive 2010/31/EU of the European parliament and of the C. - EUR-lex, 2022, [online] Available at: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:32010L0031. (Accessed 6 June 2022).

[29] L.A. Barroso, U. Hölzle, The case for energy-proportional computing, Computer 40 (12) (2007) 33–37, http://dx.doi.org/10.1109/mc.2007.443.

[30] H.-P. Jiang, D. Chuck, W.-M. Chen, Energy-aware data center networks, J. Netw. Comput. Appl. 68 (2016) 80–89, http://dx.doi.org/10.1016/j.jnca.2016.04.003.

[31] A. Marathe, P.E. Bailey, D.K. Lowenthal, B. Rountree, M. Schulz, B.R. de Supinski, A run-time system for power-constrained HPC applications, Lecture Notes in Comput. Sci. 39 (2015) 4–408, http://dx.doi.org/10.1007/978-3-319-20119-1_28.

[32] Happyware DE - english, Low power servers, 2022, [online] Available at: https://happyware.com/uk-en/servers/solutions/low-power. (Accessed 6 June 2022).

[33] S.A. Nada, K.E. Elfeky, Effects of servers 'Rack location and power loading configurations on the thermal management of data center racks' array, J. Therm. Sci. Eng. Appl. 9 (4) (2017) http://dx.doi.org/10.1115/1.4036009 [online].

[34] N. Tolia, Z. Wang, P. Ranganathan, C. Bash, M. Marwah, X. Zhu, Unified thermal and power management in server enclosures, in: ASME 2009 InterPACK Conference, Volume 2, 2009, http://dx.doi.org/10.1115/interpack2009-89075.

[35] L. Li, C.-J.M. Liang, J. Liu, S. Nath, A. Terzis, C. Faloutsos, ThermoCast, in: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '11, 2011, http://dx.doi.org/10.1145/2020408.2020611.

[36] A.V. Papadopoulos, C. Klein, M. Maggio, J. Dürango, M. Dellkrantz, F. Hernández-Rodriguez, E. Elmroth, K.-E. Årzén, Control-based load-balancing techniques: Analysis and performance evaluation via a randomized optimization approach, Control Eng. Pract. 52 (2016) 24–34, http://dx.doi.org/10.1016/j.conengprac.2016.03.020.

[37] L. Parolini, B. Sinopoli, B.H. Krogh, Z. Wang, A cyber–physical systems approach to data center modeling and control for energy efficiency, Proc. IEEE 100 (1) (2012) 254–268, http://dx.doi.org/10.1109/jproc.2011.2161244.

[38] Y. Chen, D. Gmach, C. Hyser, Z. Wang, C. Bash, C. Hoover, S. Singhal, Integrated management of application performance, power and cooling in data centers, in: 2010 IEEE Network Operations and Management Symposium, NOMS 2010, 2010, http://dx.doi.org/10.1109/noms.2010.5488433.

[39] F. Zhu, Y. Lu, Z. Ding, Batch workloads management for data centers considering nodes efficiency, in: 2018 2nd IEEE Conference on Energy Internet and Energy System Integration, EI2, 2018, http://dx.doi.org/10.1109/ei2.2018.8582052.

[40] F. Cao, Y. Wang, F. Zhu, Y. Cao, Z. Ding, UPS Node based workload management for data centers considering flexible service requirements, in: 2019 IEEE/IAS 55th Industrial and Commercial Power Systems Technical Conference, I & CPS, 2019, http://dx.doi.org/10.1109/icps.2019.8733361.

[41] Y.-F. Hsu, K. Matsuda, M. Matsuoka, Self-aware workload forecasting in data center power prediction, in: 2018 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, CCGRID, 2018, http://dx.doi.org/10.1109/ccgrid.2018.00047.

[42] J. Xue, F. Yan, R. Birke, L.Y. Chen, T. Scherer, E. Smirni, PRACTISE: Robust prediction of data center time series, in: 2015 11th International Conference on Network and Service Management, CNSM, 2015, http://dx.doi.org/10.1109/cnsm.2015.7367348.

[43] S.-U.-R. Baig, W. Iqbal, J.L. Berral, A. Erradi, D. Carrera, Adaptive prediction models for data center resources utilization estimation, IEEE Trans. Netw. Serv. Manag. 16 (4) (2019) 1681–1693, http://dx.doi.org/10.1109/TNSM.2019.2932840 [online].

[44] D. Yi, X. Zhou, Y. Wen, R. Tan, Efficient compute-intensive job allocation in data centers via deep reinforcement learning, IEEE Trans. Parallel Distrib. Syst. 31 (6) (2020) 1474–1485, http://dx.doi.org/10.1109/tpds.2020.2968427.

[45] Jeehoon Choi, Minseok Ha, Yunkeun Lee, S. Graham, Hwankook Kang, Thermal management of high density power servers using a compact two-phase loop cooling system, in: 29th IEEE Semiconductor Thermal Measurement and Management Symposium, 2013, http://dx.doi.org/10.1109/semi-therm.2013.6526801.

[46] C. Song, C. Wang, N. Ahuja, X. Zhou, A. Daniel, Optimize datacenter management with multi-tier thermal-intelligent workload placement, in: 2015 31st Thermal Measurement, Modeling & Management Symposium, SEMI-THERM, 2015, http://dx.doi.org/10.1109/semi-therm.2015.7100134.

[47] J. Sarkinen, R. Brannvall, J. Gustafsson, J. Summers, Experimental analysis of server fan control strategies for improved data center air-based thermal management, in: 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITherm, 2020, http://dx.doi.org/10.1109/itherm45881.2020.9190337.

[48] S. Ilager, K. Ramamohanarao, R. Buyya, Thermal prediction for efficient energy management of clouds using machine learning, IEEE Trans. Parallel Distrib. Syst. (2020) 1, http://dx.doi.org/10.1109/tpds.2020.3040800.

[49] D. De Chiara, M. Chinnici, A.-L. Kor, Data mining for big dataset-related thermal analysis of high performance computing (HPC) data center, Lecture Notes in Comput. Sci. 36 (2020) 7–381, http://dx.doi.org/10.1007/978-3-030-50436-6_27.

[50] W. Zhang, Y. Wen, Y. Wah Wong, K. Chuan Toh, C.-H. Chen, Towards joint optimization over ICT and cooling systems in data centre: A survey, IEEE Commun. Surv. Tutor. 18 (3) (2016) 1596–1616, http://dx.doi.org/10.1109/comst.2016.2545109.

[51] S. MirhoseiniNejad, H. Moazamigoodarzi, G. Badawy, D.G. Down, Joint data center cooling and workload management: A thermal-aware approach, Future Gener. Comput. Syst. 104 (2020) 174–186, http://dx.doi.org/10.1016/j.future.2019.10.040.

[52] Y. Ran, H. Hu, X. Zhou, Y. Wen, Deepee: Joint optimization of job scheduling and cooling control for data center energy efficiency using deep reinforcement learning, in: 2019 IEEE 39th International Conference on Distributed Computing Systems, ICDCS, 2019, http://dx.doi.org/10.1109/icdcs.2019.00070.

[53] W. Khan, D. De Chiara, A.-L. Kor, M. Chinnici, Exploratory data analysis for data center energy management, in: Proceedings of the Thirteenth ACM International Conference on Future Energy Systems, 2022, http://dx.doi.org/10.1145/3538637.3539654.

[54] Y. Gebreyesus, D. Dalton, S. Nixon, D. De Chiara, M. Chinnici, Machine learning for data center optimizations: feature selection using shapley additive explanation (SHAP), Future Internet 15 (2023) 88, http://dx.doi.org/10.3390/fi15030088.

[55] A. Grishina, M. Chinnici, A.-L. Kor, D. De Chiara, E. Guarnieri, J.-P. Rondeau, G. Georges, Thermal awareness to enhance data center energy efficiency, Clean. Eng. Technopoly (ISSN: 26667908) (2022).

[56] Tech Guided, Safe CPU temps: How hot should my CPU be? (Idle & Under Load), 2019, [online] Available at: https://techguided.com/safe-cpu-temp/.

[57] J. Brownlee, How to decompose time series data into trend and seasonality, 2017, [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/decompose-time-series-data-trend-seasonality/#:~:text=Time%20series%20decomposition%20involves%20thinking.

[58] Wikipedia Contributors, Augmented Dickey–Fuller test, 2019, [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/Augmented_Dickey%E2%80%93Fuller_test.

[59] G. Athanasopoulos, D.S. Poskitt, F. Vahid, Two canonical VARMA forms: Scalar component models Vis-à-Vis the Echelon form, Econometric Rev. 31 (1) (2012) 60–83, http://dx.doi.org/10.1080/07474938.2011.607088.

[60] Analytics Vidhya, LSTM architecture | understanding the LSTM architecture, 2021, [online] Available at: https://www.analyticsvidhya.com/blog/2021/01/understanding-architecture-of-lstm/#:~:text=The%20architecture%20of%20LSTM%3A.

[61] Jayesh Salvi, Significance of ACF and PACF plots in time series analysis, 2019, [online] Medium. Available at: https://towardsdatascience.com/significance-of-acf-and-pacf-plots-in-time-series-analysis-2fa11a5d10a8.

[62] neptune.ai., ARIMA & SARIMA: Real-world time series forecasting, 2021, [online] Available at: https://neptune.ai/blog/arima-sarima-real-world-time-series-forecasting-guide.