Applied Energy 305 (2022) 117846

Contents lists available at ScienceDirect

Applied Energy

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

Machine learning assisted development of IT equipment compact models for data centers energy planning

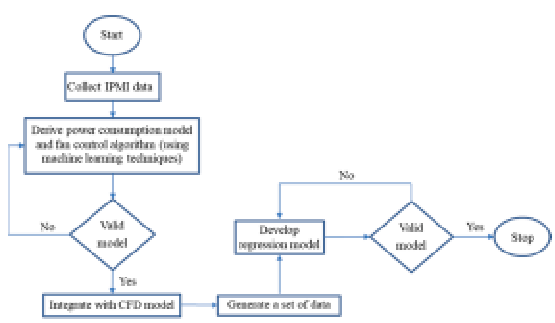
Yaman M. Manaserh [a,*], Mohammad I. Tradat [a], Dana Bani-Hani [b], Aseel Alfallah [b], Bahgat G. Sammakia [a], Kourosh Nemati [c], Mark J. Seymour [c]

[a] Department of Mechanical Engineering, ES2 Center, Binghamton University-SUNY, NY 13902, USA
[b] Department of Systems Science and Industrial Engineering, Binghamton University-SUNY, NY 13902, USA
[c] Future Facilities, London, UK and NY, USA

HIGHLIGHTS

- Full physics-based IT equipment CFD model is built.
- Machine learning is integrated with detailed IT equipment CFD model to improve the CFD model accuracy at off-design conditions.
- Novel approach for developing compact IT equipment models which can predict its power consumption, airflow demand, and exhaust air temperature is proposed.
- The derived IT equipment compact model is validated with the experimental results.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Data center
Energy consumption modeling
Machine learning
CFD modeling
Energy efficiency
Thermal management

ABSTRACT

In most data centers, performance reliability is often ensured by setting the amount of airflow provided by the cooling units to substantially exceed that which is needed by the IT equipment. This overly conservative strategy requires additional energy expenditure, which inevitably results in a huge amount of energy being wasted by the cooling system. To eliminate adopting such wasteful policies, conducting proper management of airflow, temperature, and energy is critical. To that end, this work proposes a novel approach to developing a compact IT equipment model at off-design conditions. This model is designed to support thermal and energy management functions in data centers. The benefit of this model is that it can accurately predict not only the IT equipment power consumption, but also the amount of flowrate required for the equipment and the air temperature leaving the equipment. While the compact model's power consumption was derived as a function of CPU utilization, its flowrate demand and exhaust temperature were obtained from a dynamic detailed CFD model. Results from the compact model were validated with experiments where the maximum mismatch was found to be 5.7% in the outlet temperature field and 11.4% in flowrate. Compared to a state-of-the-art IT equipment compact model, the developed model was found to reduce the prediction error of the equipment's flowrate and outlet air temperature by up to 5.2% and 9.3 % that of the state-of-the-art IT equipment compact model, respectively.

* Corresponding author.
E-mail address: yyaseen1@binghamton.edu (Y.M. Manaserh).

https://doi.org/10.1016/j.apenergy.2021.117846
Received 2 May 2021; Received in revised form 14 August 2021; Accepted 7 September 2021
Available online 17 September 2021

Nomenclature

C_1 , C_2 ,C	Coefficients in eqns. 5 and 6
	Body forces
G	Generation of turbulent kinetic energy
h	Heat transfer coefficient
k	Kinetic energy
k^*	Thermal conductivity
	Mass flow rate
	Pressure
q	Heat flow rate
Q	Volume flow rate
R^2	Coefficient of determination
S	Volumetric heat generation
S_k, S	Source terms
	Temperature
	Velocity
Y_M	Fluctuating dilation in compressible turbulence
	Rate of turbulent kinetic energy dissipation
	Dynamic viscosity
$_t$	Turbulent viscosity
	Fluid density
$_k$,	Turbulent Prandtl numbers
	Stress tensor
	Dissipation function

Abbreviations

CAC	Cold aisle containment
CFD	Computational fluid dynamics
CRAH	Computer room air handler
DT	Decision tree
IPMI	Intelligent platform management interface
IT	Information technology
MAE	Mean absolute error
OAT	Outlet air temperature
PDU	Power distribution unit
RANS	Reynolds averaged Navier-Stokes
RF	Random forest
SAT	Supply air temperature
UPS	Uninterrupted power supply
VRM	Voltage regulator modules

Subscripts

B	Buoyancy force
Eff	Effective

1. Introduction

Over the last several decades, the world has witnessed a substantial increase in the demand for electricity [1]. Consequently, there has been a significant rise in carbon emissions from power generation systems [2], which has raised concerns over their environmental impact and contribution to the climate change crisis [3]. These facts encouraged the research to investigate the potentials for improving systems energy efficiency [4]. Data centers, which support information-driven societies, account for nearly 1.3% of the world s electricity consumption [5]. The energy consumption of these bulk energy consumers was forecasted to increase by approximately 140 billion kWh by 2020, in the US only [6]. It is reported that one to two thirds of a data center s energy consumption is allocated to cooling infrastructure [7,8]. In the electronics industry, increasing heat flux dissipation density in chips has become a bottleneck challenge that restricts further technological development. Thus, data center cooling systems have become the focal point of many researchers who want to investigate energy-saving opportunities while considering reliability concerns.

Increasingly, many emerging cooling technologies like liquid cooling [9] and two-phase cooling [10] are being widely adopted. However, air cooling remains the most reliable and widely used cooling approach to date [11]. R. Gupta, S. Asgari, H. Moazamigoodarzi, S. Pal, and I. Puri [12] compared four cooling architectures for air-cooled data centers from exergy point of view. H. Moazamigoodarzi, P. Tsai, S. Pal, S. Ghosh, and I. Puri [13] compared room, row, and rack cooling architectures in terms of data center power consumption. Their results revealed that adopting cooling architectures at the row and rack levels decreased the overall cooling system power consumption by up to 29% when compared with the room level cooling architecture. Z. Song [14] examined using fans to assess the cooling system performance in data centers with and without Cold Aisle Containment (CAC). H. Lu and Z. Zhang [15] analyzed the data center thermal performance variation while changing its geometrical configurations and the computer room s air conditioning arrangement. X. Meng, J. Zhou, X. Zhang, Z. Luo, H. Gong, and T. Gan [16] optimized the thermal performance of a relatively small data center in China. Y. Zhang, K. Zhang, J. Liu, R. Kosonen, and X. Yuan [17] constructed T-shaped underfloor air ducts in a modular data center to improve the airflow uniformity. A. Almoli, A. Thompson, N. Kapur, J. Summers, H. Thompson, and G. Hannah [18] developed a new computational fluid dynamics (CFD) methodology for modeling data centers that adopted rear door heat exchangers. M. T-Evans, N. Kapur, J. Summers, H. Thompson, and D. Oldham [19] investigated the effect of air bypass and recirculation in data centers on the cooling infrastructure power consumption. J. Cho and J. Woo [20] introduced an in-row coolers for improving the thermal environment of a data center by providing a lesser air distribution path. L. S-Llanca, A. Ortega, K. Fouladi, M. Valle, and V. Sundaralingam [21] tested direct and indirect exergy destruction approaches against simplified and actual data center flow. They found that using the direct method results in a better when calculating the exergy destruction for the airside in data center. C. Zhou, C. Yang, C. Wang, and X Zhang [22] conducted a numerical analyses on a cooling system designed for small-scale data centers. J. Athavale, M. Yoda, and Y. Joshi [23] compared different data-driven modeling approaches according to their ability to predict temperatures in data centers. A. Khalaj, T. Scherer, J. Siriwardana, and S. Halgamuge [24] tested the impact of distributing the workload on a data center s cooling system s efficiency. It is important to note that all these studies employed CFD models whilst conducting their analyses at the data center facility level.

Over the years, CFD techniques have been recognized as the most convenient tool for designing, characterizing, and diagnosing data centers [17,25]. Using a CFD digital twin model of a data center can accurately replicate the data center s performance at different operating conditions. compared to other data center modeling techniques, such as proper orthogonal decomposition [26 29], console-based simulation tools [30], thermodynamic flow network modeling [31], and machine learning models [32 35], CFD simulations are the most adoptable [36]. However, CFD models are also the most computationally intensive among all modeling approaches. As a result, compact models have been employed in CFD tools as a method for avoiding excess computational cost. A compact model assumes that the complexities of physics inside an object can be represented by an approximate model of the object s impact on its surroundings at the locations where the object and the surroundings interact. This assumption is less computationally intensive than those made by traditional CFD simulations because it requires less level of detail. In facility level analyses of white space, compact models are considered necessary for modeling the complexities of IT equipment in a computationally efficient manner.

Additionally, server compact models are not limited to CFD analyses, but can also be used to conduct thermal and energy management functions in a data center, identify the potential for power management, predict power consumption, and balance the needs of the data center, and so on [37]. Other approaches to developing compact models for IT equipment and power consumption have been investigated by researchers for use in CFD modeling at the data center facility level. H. Cheung, S. Wang, C. Zhuang, J. Gu [38] presented IT equipment simplified model for calculating power consumption in a data centers. This model was developed to be utilized in simulating the dynamic IT equipment energy performance in data centers under various operation conditions. J. VanGilder, Z. Pardey, X. Zhang, and C. Healey [39] introduced a compact server model that captured the effects of the server s thermal mass. H. Moazamigoodarzi, R. Gupta, S. Pal, P. Tsai, S. Ghosh, and I. Puri [40] developed IT server enclosures compact model which can capture the temperature distribution and power consumption inside the enclosures. S. Erden, E. Khalifa, and R. Schmidt [41] proposed a new method to derive some server s properties such as thermal capacitance and conductance from external measurements. A review of existing power consumption models for IT equipment was conducted by C. Jin, X. Bai, C. Yang, W. Mao, and X. Xu [37]. All these studies focused mainly on predicting the power consumed by the IT equipment or on developing a simplified equipment model without considering the IT equipment s different operating conditions. Furthermore, these studies ignored the physics-based variations between different IT equipment, which significantly affected their model s accuracy, especially at the off-design conditions. Lastly, none of the previous studies were able to provide an accurate model for calculating the IT equipment s airflow requirements.

In a real-world data center, IT equipment operating conditions, such as the data center s architecture, spatial location, inlet pressure, inlet air temperature, and CPU utilization, can vary significantly. This requires a physics-based compact model that can predict the IT equipment s behavior under off-design conditions. Since error can be significantly magnified when implementing these compact models in data center facility level CFD models, the compact model s degree of accuracy at different operating conditions is a mandatory factor that must be considered. Moreover, capturing the amount of flow required by each IT equipment is vital for power and capacity planning in data centers. Generally, data center operators tend to overprovide IT equipment with cool air to ensure their reliability. Adopting this practice results in a huge amount of wasted energy by the Computer Room Air Handler (CRAH) unit, which is responsible for a considerable share of the cooling system s total power consumption. On the other hand, under-provisioning leads to an increase in the amount of hot air recirculation and the formation of hotspots, consequently decreasing the cooling system s efficiency and exposing the equipment to reliability issues. Therefore, accurately predicting the amount of air flow required by IT equipment is as critical as predicting its heat dissipation, since it is used to calculate the cooling units air delivery.

This study presents a novel extended method of building an IT equipment compact model that employs a combination of machine learning (regression analysis) and CFD simulation. By integrating machine learning with the detailed CFD IT equipment model, the CFD model transforms from a rigid model that can predict the IT equipment performance at the design point conditions to a dynamic model which is capable of predicting the overall performance at various environmental conditions. As this model simulates the overall performance of the IT equipment, it can be used to derive a compact model that predicts the required amount of flowrate and the air temperature leaving the IT equipment, besides the power dissipated by the equipment. The key novelty of this approach that distinguish it from existing methods for building compact models of IT equipment, and they can be summarized as follows:

- This approach captures the IT equipment s behavior under different environmental conditions in the data center, which is essential for predicting the data center s energy load demand, hence regulating the data center s power consumption [42].
- This model adapts a full physics-based model, hence it is expected to provide very accurate predictions at both design and off-design conditions.
- All the data and tools used in this approach are readily available to almost everyone involved in the thermal management of a DC (CFD software, technical manuals, Intelligent Platform Management Interface (IPMI) data, etc.).
- This approach can also compute the amount of airflow required by the IT equipment at various operating conditions with a high degree of accuracy.

Combined, these distinguishing characteristics provide the foundation for a novel method of building an IT equipment compact model. Beside introducing the novel approach, the integration of machine learning algorithms with the IT equipment CFD model, where the machine learning model was used to generate data that is used as input for IT equipment CFD model, is completely new.

For many years, specifying the accurate flowrate demand in data centers at different operating conditions is identified as a key restriction for further improvement in the data center s cooling system s energy efficiency. Without providing an accurate practical solution, the incapability of calculating the airflow demand leaves a remarkable knowledge gap in the available literature. Hence, this study is to provide a significant contribution to the new body of knowledge as it introduces a reliable and fast tool for calculating the flowrate demand in data centers. The following are the major advantages of utilizing this tool:

- By integrating this tool into the thermal and energy management of a DC, its cooling system capacity planning can potentially be expanded to cover calculating the amount of airflow required by the equipment, along with the IT equipment power consumption and heat dissipation. Since this extended method considers the amount of airflow required by the IT equipment, with the appropriate adjustments to the airflow supply from the cooling unit it introduces the potential for substantial energy savings in currently operating data centers. It was proven that adopting effective airflow management in data centers could result in a substantial amount of energy savings [43]. According to a previous study, adopting proper flow management showed a possibility for significant energy saving opportunities on both the CRAH unit and the chiller, where the CRAH power consumption was reduced by 75% and the chiller power consumption was reduced by 16% [44].
- Using this tool can fill in the lack of information on IT equipment s flow requirements, which can potentially prevent cooling system oversizing or overprovisioning in future data centers.
- When it comes to CFD models, implementing this approach could remarkably improve their accuracy and reduce their computational costs, especially during transient simulations in which a data center experiences various operating conditions.
- This tool eliminates the need to calibrate CFD models because the IT equipment compact model can replicate accurate performances at different environmental conditions.

The remainder of this paper is organized as follows: the method used to develop the IT equipment compact model is described, including the details of data collection, deriving the power consumption model, the fan control algorithm, the numerical model, and the regression function. Then, the experimental setup that was used to validate the results is introduced. Thereafter, the results from each stage of development for the IT equipment compact model are discussed and compared with the experimental results. Finally, further conclusions are drawn in the last section.

2. Methodology

In many CFD tools, traditional compact models for IT equipment are overly simplistic. Oftentimes, they treat the airflow and heat dissipation as fixed values or are lumped together as a single rack compact model. Some tools go further, and the flow can be controlled as well as responding to differential pressure. The method presented in this work improves upon the accuracy of the IT equipment compact model. It proposes utilizing data that can be collected in almost all data centers and developing a series of regression models to build the IT equipment compact model. The methodological approach starts with collecting the IPMI data for the IT equipment. Then, this data is used to derive the IT equipment's power consumption model and the fan speed prediction models. Thereafter, these models are integrated with a detailed physics-based IT equipment CFD model to predict the IT equipment's performance at off-design conditions. This model is utilized to produce a new set of data that correlates the operating conditions with the response parameters. Finally, a proper regression model is applied to yield simple, yet accurate, regression functions that represent the IT equipment compact model. Fig. 1 shows the flow diagram for the proposed methodological approach adopted in this study.

It should be noted that there may be considerable differences between various types of IT equipment. The differences between different servers' types and models include size, layout, geometrical parameters, rated power, number of fans, etc. Even for some modern servers, the BIOS settings allow the user to set the server in a high-performance mode where the fan power is reduced. Hence, these differences result in different airflow patterns, flowrate, temperature variations across the server, and dissipated power. The objective of this paper is to introduce a general methodology of creating various compact models for various IT equipment based on a combined CFD and operational data approach. After thoroughly explaining the compact model development approach, the results and discussion section examined constructing an example compact model for specific IT equipment to assess the suggested approach. Thus, all the results presented in that section are for a single specific server.

2.1. Data collection

The methodological approach presented in Fig. 1 starts with collecting IT equipment monitoring data through the IPMI. IPMI is defined as set of regulations for hardware-based platform management systems, in another words it is a kind of security guard for IT equipment. One function of the IPMI is to monitor the IT equipment status by reporting the temperature readings, power consumption, fan speeds, etc. Note, different types of data can be collected from different servers. The most common sets of collected data are inlet air temperature; individual CPU utilization; motherboard temperature; individual CPU max temperature, average temperature, and CPU power consumption; and total power entering the IT equipment. To ensure that a rich dataset is generated for machine learning techniques, the data should be collected at various environmental operating conditions. The data dataset considered in this study was generated in the ES2 data center laboratory at the State University of New York at Binghamton, wherein the chosen IT equipment was tested at various operating conditions. ES2 data center laboratory was built for research purposes, and a special in-house C program was developed for gathering the data from the different components within the data center through Linux operating system. Basically, this program reads the values reported by a sensor or a component and gathers them in a comma-separated values (CSV) file. The program was developed to collect a wide range of data including data that is not often gathered in data centers, such as CPU utilization.

2.2. Deriving power consumption model and fan control algorithm (using machine learning techniques)

- **Power consumption model**

Data center components can be classified into four main categories: power components, cooling equipment, IT equipment, and miscellaneous components (sensors, monitoring systems, etc.) [45]. Together, IT and cooling equipment are responsible for around 90% of overall data center power consumption [46]. Additionally, the IT and cooling equipment are highly coupled components because of their interactions in the data center's thermal environment. For the cooling system, its load is defined upon the total energy (heat) dissipated by the IT equipment. For the IT equipment, flowrate increases when the cooling unit SAT is increased, subsequently increasing both their power consumptions [37]. A schematic for the power and heat flow inside a data center is provided in Fig. 2.

Based on the complexity of these interactions, to understand energy and heat flow in each data center, a more comprehensive compact model that can accurately describe the overall power consumption and thermal performance of the IT equipment is needed.

When it comes to IT equipment power consumption, several studies concluded that the CPU utilization rate and the IT equipment status (on/

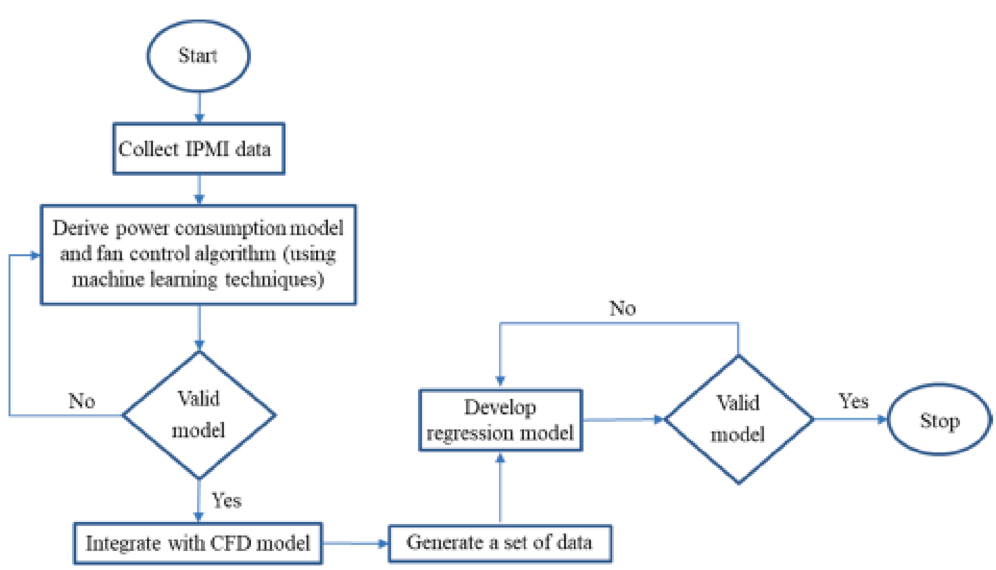


Fig. 1. Methodology for building IT equipment compact model.

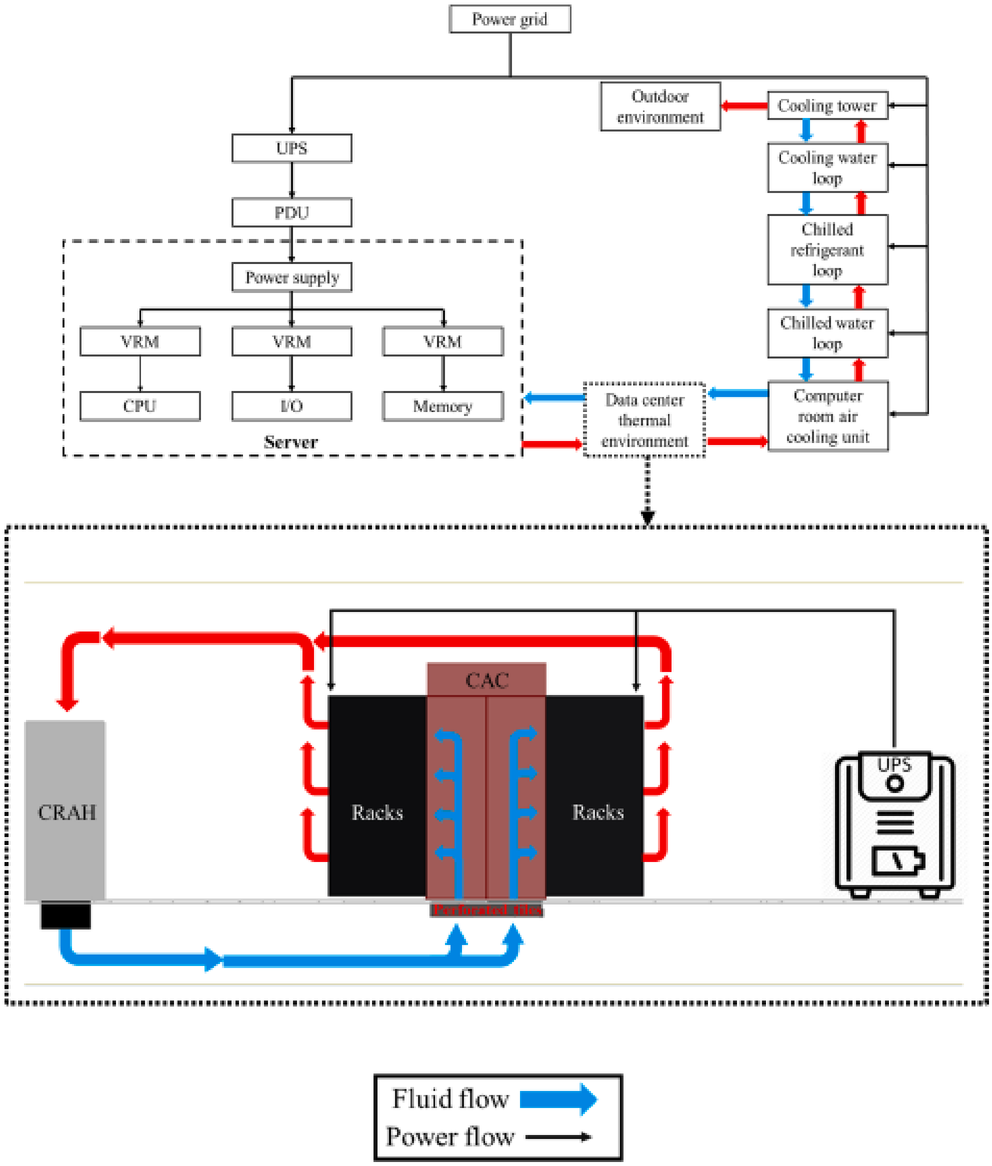


Fig. 2. Schematic of power and heat flow in legacy raised floor data centers [37].

off) significantly affected the power consumed by the IT equipment and consequently, the thermal load in the data center [47–49]. For that reason, a power usage model was obtained from the IPMI data to correlate the CPU power consumption with its utilization factor. The power consumption model for each IT equipment can be easily extracted from the IPMI data as a function of CPU utilization.

- **Fan speed prediction**

As previously mentioned, overprovisioning and under provisioning airflow can have detrimental effects on power consumption and reliability, respectively. Therefore, it is important to establish the exact amount of airflow required by each aisle by estimating the airflow required by the IT equipment. A practical way to calculate the IT equipment airflow demand would be to calculate the flowrate driven by its fan at certain operating conditions, which can be done using the fan speed and the fan performance curve. However, unlike the fan performance curve, the fan control algorithm is never reported by vendors. Thus, in this study, we attempted to utilize the IPMI data to extract the fan control algorithm. However, the interaction between the different features (collected from the IT equipment sensors through IPMI) is preventing capturing the exact fan control algorithm. For instance, increasing the SAT affects all other temperatures reported by the different sensors inside the IT equipment.

Therefore, instead of extracting the actual fan control algorithm, fan speed prediction models were developed. Building speed prediction models for the IT equipment fans is an extremely complex task to be done. Whereas a single fan could be controlled by different sensors inside the IT equipment, also each fan within the same IT equipment could have its own control algorithm. According to the level of task complexity and the huge amount of collected data, machine learning techniques, specifically regression and feature selection, were utilized in building fan speed prediction models. The models were built for each fan in the IT equipment at different operating conditions.

- **Machine learning application**

Machine learning helps extract valuable information and associations from massively large and complex data (big data), which can provide meaningful insights and help make critical decisions. Processing big data manually would be impractical. Therefore, the use of machine learning techniques to automate analytical processes is extremely important and beneficial [50] as it saves time and reduces costs. Machine learning helps analyze big data by designing algorithms that can learn patterns and relationships in the data to make predictions on newly obtained data based on what the algorithm has learnt. Fig. 3 depicts the overall application of machine learning in this study.

In supervised machine learning, the output variable (dependent variable) of recorded data is provided and machine learning algorithms learn how to map inputs (independent variables) to outputs. This learnt relationship between the inputs and outputs is called a prediction model. Therein, an observation with no known output may be fed into the model and produce an output based on the relationships learnt by the algorithm. Supervised machine learning can be broken down into classification tasks and regression tasks. The former has a categorical output and the latter has a continuous output. An example of a classification task in machine learning is building a model to predict if the amount of air delivered to a server is sufficient, where the predicted variable would be 0 or 1, representing no or yes, respectively. This is based on previously recorded data where the algorithm can learn to differentiate between sufficient airflow or insufficient airflow based on features such as air temperature, CPU temperature, CPU utilization, etc. An example of a regression task is building a model to predict the speed of IT equipment fans, as the case of this study, using given features such as those described in the classification task. The output for the classification task is presented as a categorical output (i.e., 0 or 1), whereas the regression task is presented as continuous values (e.g., 3000 RPM).

- **Adopted machine learning ensemble**

Random forest (RF) is a machine learning ensemble created by Ho TK [51] for classification and regression tasks. It is the machine learning algorithm used in this study to develop prediction models for the fans' speeds. Ensemble models are multiple algorithmic systems whose decisions are combined to improve the performance of the overall system [52]. To obtain a single prediction from an ensemble making up of several algorithms, two methods are commonly used: voting and averaging. Voting is used in classification where the predicted class of an observation is chosen from the majority of the classifiers' votes that the ensemble is composed of. However, in regression, the average over all the regressor algorithms' predictions is taken as the final prediction of the ensemble. In the case of RF, it consists of many decision trees (DTs) and uses the bagging technique to overcome the DT's overfitting drawback by decreasing the model's variance without increasing the bias. DT is a machine learning algorithm made up of nodes and leaves and can predict an output by learning decision rules inferred from the data. The DTs in RF are created by selecting subsets of the training data with replacement, and each tree consists of a random subsample of features.

RFs are very flexible and powerful ensembles; they have been adopted in many research fields including biological and biomedical research. Researchers tend to use ensembles in their studies because they combine the strengths of several algorithms together to produce superior results. RF was chosen based on the unique way DTs work. DT generates a prediction based on if-else rules that are built in a hierarchal order of feature importance, unlike many regressors, such as linear regression, where a line is fitted to the model and the prediction is generated from the line's equation. RF provides a way to rank features based on their importance. It consists of a number of DTs, where each DT is built on a condition made on a single feature and designed to split the dataset into two separate sets with similar responses within. The features that are selected for the internal nodes are selected based on a measure called impurity. In the case of regression, the impurity measure is variance reduction. During the training phase, each feature is measured by how much it can decrease the impurity in a tree, wherein the feature with the highest decrease is selected for the internal node. The decrease in impurity for each feature is averaged for all the trees, and the averages are then ranked according to this measure.

The RF algorithm helps solve complex problems due to its capability of discovering nonlinear relationships in the data without statistical assumptions [53]. Numerous studies have attested to its power such as [54–56], as well as a comparative study by J. Zhou, X. Li, and H.S. Mitri [57] of ten machine learning algorithms to classify rock burst events in which RF achieved the highest performance. Since RF tackles the problem of overfitting, as well as being an ensemble and predicting based on learnt decision rules, it was adopted as the machine learning algorithm in this study. Fig. 4 shows an example of a RF algorithm with K DTs in which the ensemble generates an overall prediction for an unknown observation. The RF regressor's final prediction is given in the following equation:

$$\hat{y} = \frac{1}{K}\sum_{k=1}^{K}\hat{y}_k \tag{1}$$

where $\hat{y}$ is the overall prediction of the RF, K is the number of DTs that make up the RF, and $\hat{y}_k$ is the prediction for tree k. Since the RF regressor

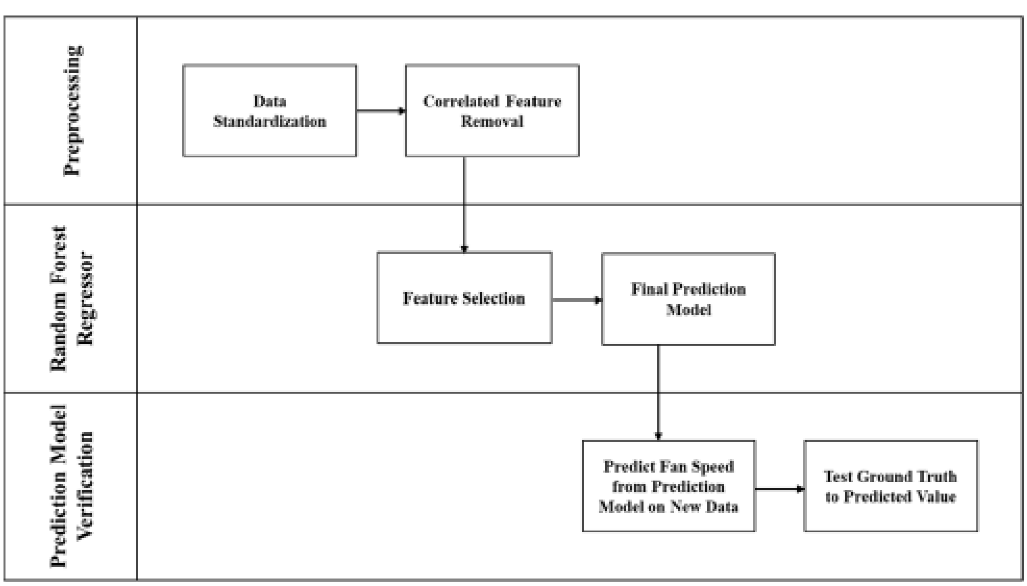


Fig. 3. An illustration of the overall application of machine learning.

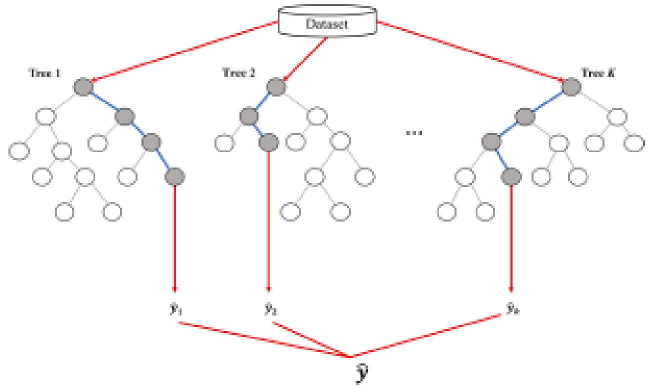


Fig. 4. A RF algorithm with K DTs producing one overall prediction $\hat{y}$.

is an ensemble, the final prediction is given as the average of all predictions generated from the DTs.

- **Training and testing machine learning algorithm**

In machine learning, to evaluate a model, the algorithm must be trained on a subset of the data, called the training set, and then tested on another subset, called the testing set. The training phase of the algorithm is when it learns how to map the inputs to the outputs and learns the relationships and patterns in the data. After the trained model is created, it needs to be tested and validated on a part of the data that was not involved in training to prevent overfitting and improve model generalization. The model needs to be generalized to predict observations not seen during training to allow real-world applications. Therefore, after the machine learning model is created, it is validated on a testing set and the model's prediction capabilities are measured using performance metrics. Commonly used performance metrics for classification tasks include accuracy, area under the receiver operating characteristic (ROC) curve (AUC), sensitivity, and specificity; and for regression tasks, metrics such as the R^2 coefficient, root mean squared error (RMSE), and mean absolute error (MAE) are commonly used.

Splitting the dataset into training and testing subsets is one of the most widely used methods to avoid overfitting and improve model generalization. This method is called the holdout method. Therein, the training set consists of 70–80% of the dataset while the testing set holds the rest. A training set with 80% of the data and a testing set with the remaining 20% is called an 80–20 split. The training set should be significantly larger than the testing set because the model needs more data to learn the relationships between the inputs and outputs. However, in cases where the amount of data provided is insufficient, leave-one-out cross-validation can be applied. It is where each observation in the dataset with N observations is tested separately. The training set consists of $N-1$ observations, and the model is then tested on the one left-out observation by generating a prediction. This procedure is repeated N times, in which each observation is tested once and trained $N-1$ times, thus, a total of N models is trained and tested. The final accuracy is computed as the average accuracy for all N models. In machine learning, a more robust approach is favored over the holdout method, which is called the k-fold cross-validation method. To ensure robustness and reduce bias, the dataset is partitioned into k folds where k models are created, and then each model is trained on $k-1$ partitions and tested on one. Fig. 5 demonstrates how k-fold cross-validation works when k = 10. The overall performance is taken as the average across all folds and is shown in equation. 2, where R^2 is the overall coefficient of determination, K is the number of folds, and R_k^2 is the coefficient of determination for fold k.

$$R^2 = \frac{1}{K}\sum_{k=1}^{K} R_k^2 \tag{2}$$

2.3. Numerical model

- **IT equipment description**

In this study, the compact model is developed for only a piece of IT equipment as a case study. Thus, a detailed model of the considered IT equipment is built and simulated. Fig. 6 depicts the IT equipment studied in this work. The dimensions showed in Fig. 6 (c), are provided in Table 1 along with other geometrical parameters. The selected IT equipment consists mainly of two CPUs, six fans, twelve RAMs, a power supply, and four hard disk drives. The hard disk drives are located at the IT equipment inlet, while the fans are installed in the middle of the IT equipment chassis. As shown in Fig. 6, five of the fans are allocated to cool the CPUs and the RAMs, while the last fan is assigned to cool the power supply board and the power supply itself.

- **Governing equations and turbulence model**

Simulations are carried out using the commercially available CFD package of 6SigmaET®. The solution governing equations are given as follows:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u_i)}{\partial x_i} = 0 \tag{3}$$

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial(\rho u_i u_j)}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + F_i \tag{4}$$

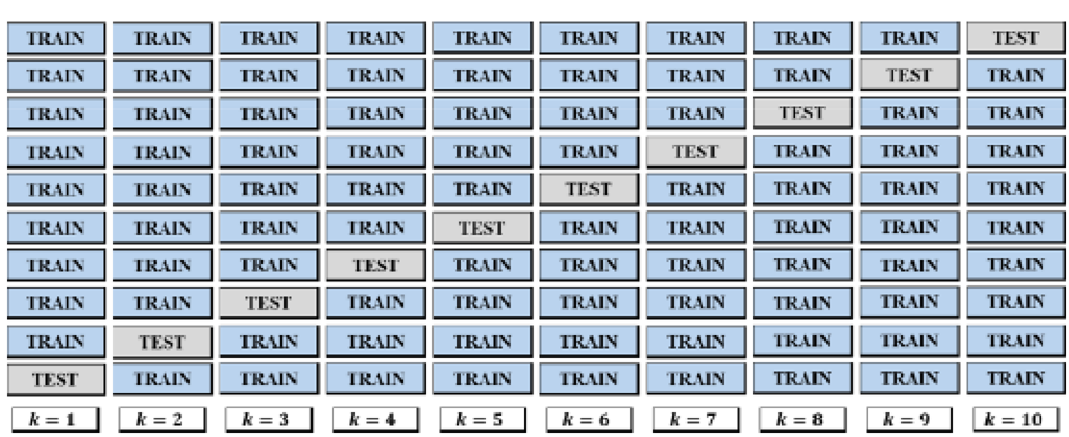
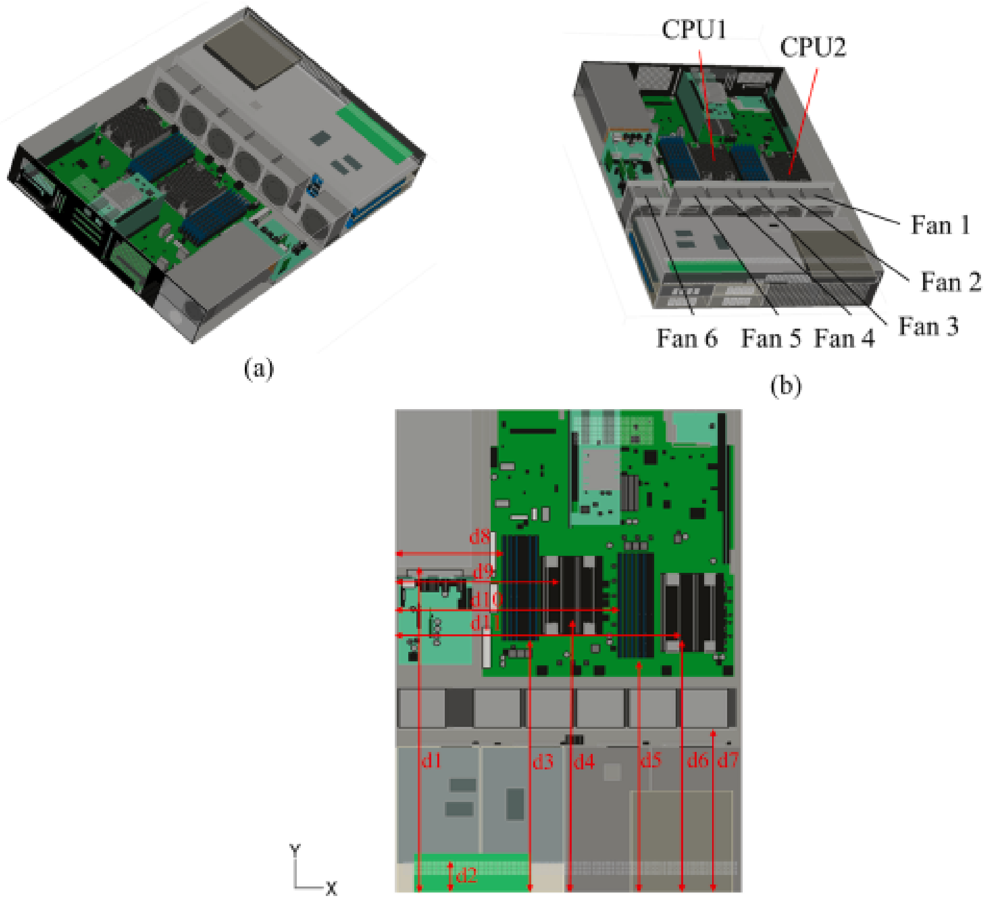


Fig. 5. An illustration of 10-fold cross-validation.



(a)
(b)
(c)

Fig. 6. CAD drawing of the IT equipment investigated in this study.

Table 1
IT equipment geometrical parameters.

Parameter	Value	Parameter	Value
Chassis height, width, length (mm)	89, 440, 615	Distance to HDDs d2 (mm)	38
Fans height, width, length (mm)	60, 60, 50	Distance to RAMs d3 (mm)	322
Fans hub diameter (mm)	36	Distance to CPU1 d4 (mm)	361
HDDs height, width, length (mm)	24, 101, 147	Distance to RAMs d5 (mm)	300
CPUs height, width, length (mm)	4, 37, 37	Distance to CPU2 d6 (mm)	338
Heat sink base height, width, length (mm)	3.5, 79.3, 99.5	Distance to fans d7 (mm)	160
Fins height, width, length (mm)	19, 0.5, 99.5	Distance to RAMs d8 (mm)	134
Fins number per heat sink	40	Distance to CPU1 d9 (mm)	224
PSU height, width, length (mm)	40, 88, 195	Distance to RAMs d10 (mm)	282
RAMs height, width, length (mm)	10, 6, 133	Distance to CPU2 d11 (mm)	358
Distance to power supply d1 (mm)	420		

$$\frac{\partial(\rho E)}{\partial t}+\frac{\partial(u_m(\rho E+P))}{\partial x_m}=\frac{\partial}{\partial x_j}\left(k_{eff}\frac{\partial T}{\partial x_j}\right)+\Phi+S \tag{5}$$

k-ε turbulence model is adopted for simulating the flow field through coupling the conservation of mass, momentum, and energy governing equations with RANS equations. The selection of k-ε RANS model was according to the capabilities and reliability these models have shown in modeling different components in data centers [18,58–62]. Even for fans, which exhibits complex internal flow dynamics including multiple vortices, k-ε RANS showed significantly high accuracy when applied [63–67]. The standard k-ε turbulence model governing equations which were used for calculating the turbulent kinetic energy k, the rate of dissipation of the turbulent kinetic energy ε, and turbulent viscosity are [68–70]:

$$\frac{\partial(\rho k)}{\partial t}+\frac{\partial(\rho k u_i)}{\partial x_i}=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right]+G_k+G_b-\rho\varepsilon-Y_M+S_K \tag{6}$$

$$\frac{\partial(\rho\varepsilon)}{\partial t}+\frac{\partial(\rho\varepsilon u_i)}{\partial x_i}=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial\varepsilon}{\partial x_j}\right]+C_{1\varepsilon}\frac{\varepsilon}{k}(G_k+C_{3\varepsilon}G_b)-\rho C_{2\varepsilon}\frac{\varepsilon^2}{k}+S_\varepsilon \tag{7}$$

$$\mu_t=\rho C_\mu\frac{k^2}{\varepsilon} \tag{8}$$

For the constants $C_{1\varepsilon}$, $C_{2\varepsilon}$, and C_μ, the standard values were adopted while conducting the simulation. The corresponding values for these constants are 1.44, 1.92, and 0.09, respectively. With regards to wall treatment, the standard k-ε is used with the standard wall treatment, which implies that the near-wall grid lies within the logarithmic region. This is feasible since y^+ lies within the range 30 $<y^+$ less than 300,

where y^+ is a dimensionless parameter indicating the wall coordinate, $y^+ = \frac{y u_\tau}{\nu}$, $u_\tau = \sqrt{\tau_w/\rho}$. y is the distance between the near-wall grid and the wall [71], u_τ represents the friction velocity, and τ_w donates the wall shear stress.

The following assumptions are made while conducting the simulations: air is incompressible and has constant properties, steady state conditions are maintained, and the impacts of wall roughness and gravity are negligible. The detailed CFD model and the boundary conditions that are implemented to replicate the test environment are shown in Fig. 7.

- **Grid generation**

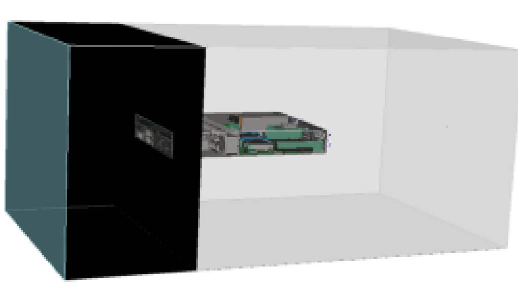
In any CFD model, grid generation plays an integral role in governing the results accuracy. In this work, a hex structured mesh has been adopted while running CFD simulations, since using flow aligned cubic hex cells guarantees a high grid quality and eliminates the aspect ratio and skewness issues of tetrahedral mesh [72]. The grid is built in a way that more cells are considered inside the IT equipment while keeping a lower number of grids outside the IT equipment. After that, a grid independence study is executed to identify the appropriate number of grids at which further increasing in cells numbers will cause minimal change in the results. Fig. 8 shows the grid independence study along with the generated grids. Accordingly, the number of grids adopted to conduct all analysis in this work is 15.4 million.

2.4. Develop regression functions.

As a final step in developing the compact model, several functions are derived to describe the overall IT equipment performance. These functions, which represent the IT equipment compact model, are used to describe the relationship between the dependent variables (response parameters: IT equipment power, flowrate, and OAT) and the independent variables (operating conditions: CPU utilization and SAT). Choosing the correct regression model can be difficult, seeing as there are many regression models that can be used. Therefore, multiple regression functions can be developed with different statistical models depending on the relevant data trends. Thereafter, the data can be validated using statistical measurements and the best model can be selected accordingly.

3. Experimental setup.

The experimental setup was designed to control the input parameters, namely inlet air temperature and CPU utilization, and to measure the response parameters, which were average outlet air temperature and velocity The room temperature was adjusted by controlling the supply air temperature of the cooling units. Three temperature sensors were



Legend
- Constant inlet pressure, constant ambient temperature.
- Adiabatic walls.
- Constant outlet pressure, constant ambient temperature.

Fig. 7. Detailed CFD model and the boundary conditions used in simulations.

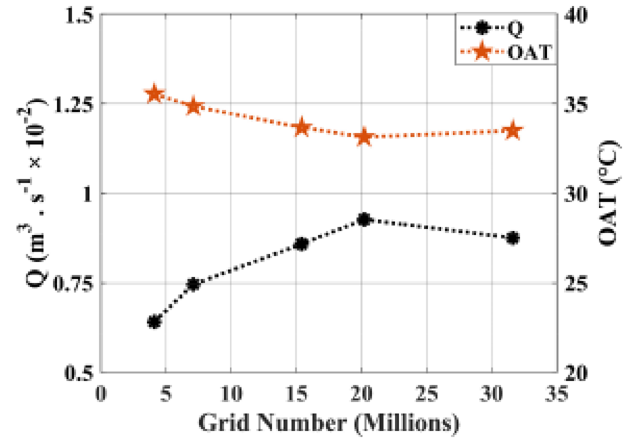


Fig. 8. Grid sensitivity analysis.

installed at the IT equipment inlet to measure the IT equipment inlet temperature, while 30 velocity and temperature sensors were installed at the IT equipment exhaust. Further details of the instrumentation are provided in Table 2.

First, the experimental setup was used to validate the CFD model (Fig. 9). Then, it was used to test the compact model results. To obtain an objective comparison of the numerical and experimental results, the precise X-Y coordinates of the sensors were measured in the physical setup and then replicated in the CFD model. Yet it is still hard to capture the precise spatial location of the sensor due to the geometry of the sensor as illustrated in Fig. 9 (c). Even though the differences in spatial locations were relatively small, they formed a considerable source of error in the results. Therefore, to reduce the impact of spatial location error, a duct was attached to the IT exhaust that took measurements at various distances from the IT equipment (Fig. 9 (d)). An identical size duct was built and attached to IT equipment exhaust in the CFD model. Primary results showed that 0.2 m was enough distance to minimize the error that arose from the location discrepancy. After that, experiments were performed at a wide range of CPU utilizations (idling-100) % and inlet air temperatures (18–33) °C.

The main objective for conducting the experiment is to generate a set of experimental results to be utilized in evaluating the CFD model and the regression model. The experimental measurements were recorded for the IT equipment in the DC at different operating conditions. After that, the CFD model and the compact model were tested at the exact same operating conditions as in the experiment to assess the models. However, it is not mandatory for the process of developing the CFD model or the compact model in this presented approach. Owing to the fact that numerical techniques have manifested its viability in a numerous number of applications for example [73–77]. Other than this, cost, equipment availability, time and efforts can be considered as vital constraints for performing experimental verifications.

4. Results and discussion

4.1. Power consumption and fan speed prediction models

- **Power consumption**

Table 2
Details of instrumentation used in the experimental setup.

Instrument	Specifications
Airflow test chamber	Designed in accordance with the air movement and control association (AMCA) 210–99 standards.
Velocity and temperature sensors	Degree C UAS1200LP (velocity sensor accuracy: ±3% of reading, thermistor accuracy ± 1 °C).
DAQ	Degree C ATM2400

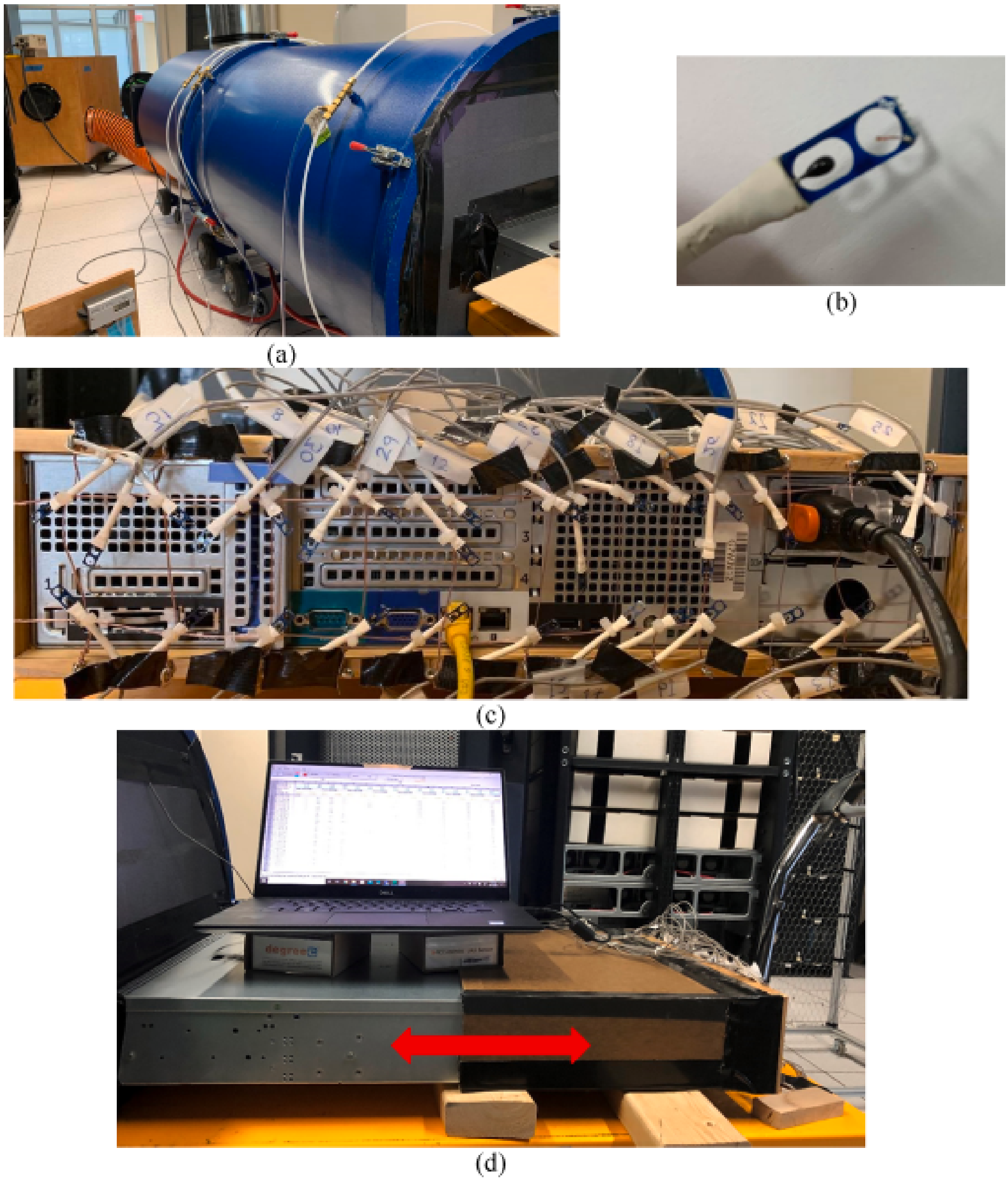


Fig. 9. Experimental setup used for verifying results. (a) flow chamber (b) temperature and velocity sensor (c) sensors installment (d) flow duct.

In the literature, there are two main power consumption models introduced, which are the additive model and the baseline/active model. The additive model considers the power contributions of all components, wherein each component has its own power consumption model [78,79]. On the other hand, the baseline/active model assumes that all components are independent of the IT equipment utilization except the CPUs [80,81].

The power consumption model developed in this work represents a combination of these two models, since the overall IT equipment power consumption was estimated upon the CPU utilization. The correlation between the CPU utilization and the CPU power consumption, as well as the total IT equipment power consumption is shown in Fig. 10. To operate the IT equipment at different power levels, Prime95 software that calculates new Mersenne prime numbers was adopted. The software uses the equation $M_p = 2^p - 1$ to calculate the Mersenne prime number, where p here denotes prime number. To run the IT equipment at a steady load, a feature called "Torture Test" in this software was used, while changing the IT equipment utilization is done by changing the workers'

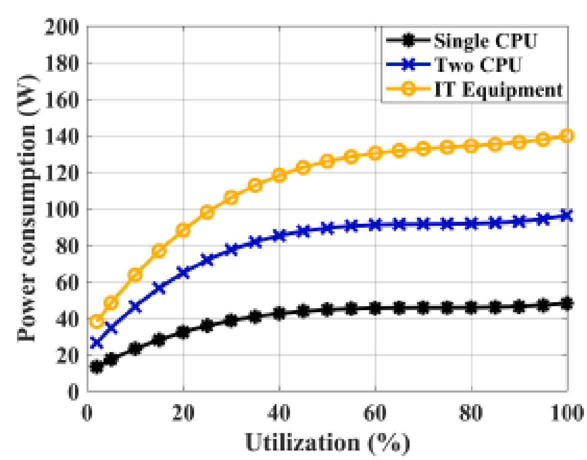


Fig. 10. Relationship between the CPU utilization and the power consumption.

number, as every worker ensure full utilization of a single core. Fig. 10 shows that the difference between the single CPU power consumption and the total IT equipment power is increasing with increasing the CPU utilization, which means that the CPU utilization is also affecting other IT equipment components' power consumption.

It can be also noted from Fig. 10 that the correlation between CPU utilization and power consumption is nonlinear. However, such correlations can be directly extracted and fitted with a third order polynomial from the IPMI data as follows:

$$CPUPower = 9.5\times10^{-5}Utilization^{3} - 0.0204Utilization^{2} + 1.4662Utilization + 10.554 \quad (9)$$

$$ITequipmentPower = 0.0002Utilization^{3} - 0.0462Utilization^{2} + 3.7091Utilization + 31.013 \quad (10)$$

Finally, besides using this correlation for estimating the IT equipment power consumption, it was also used for developing the other features in the compact model. This correlation was used to predict the total heat dissipated by the IT equipment and the consequential IT equipment flowrate requirement.

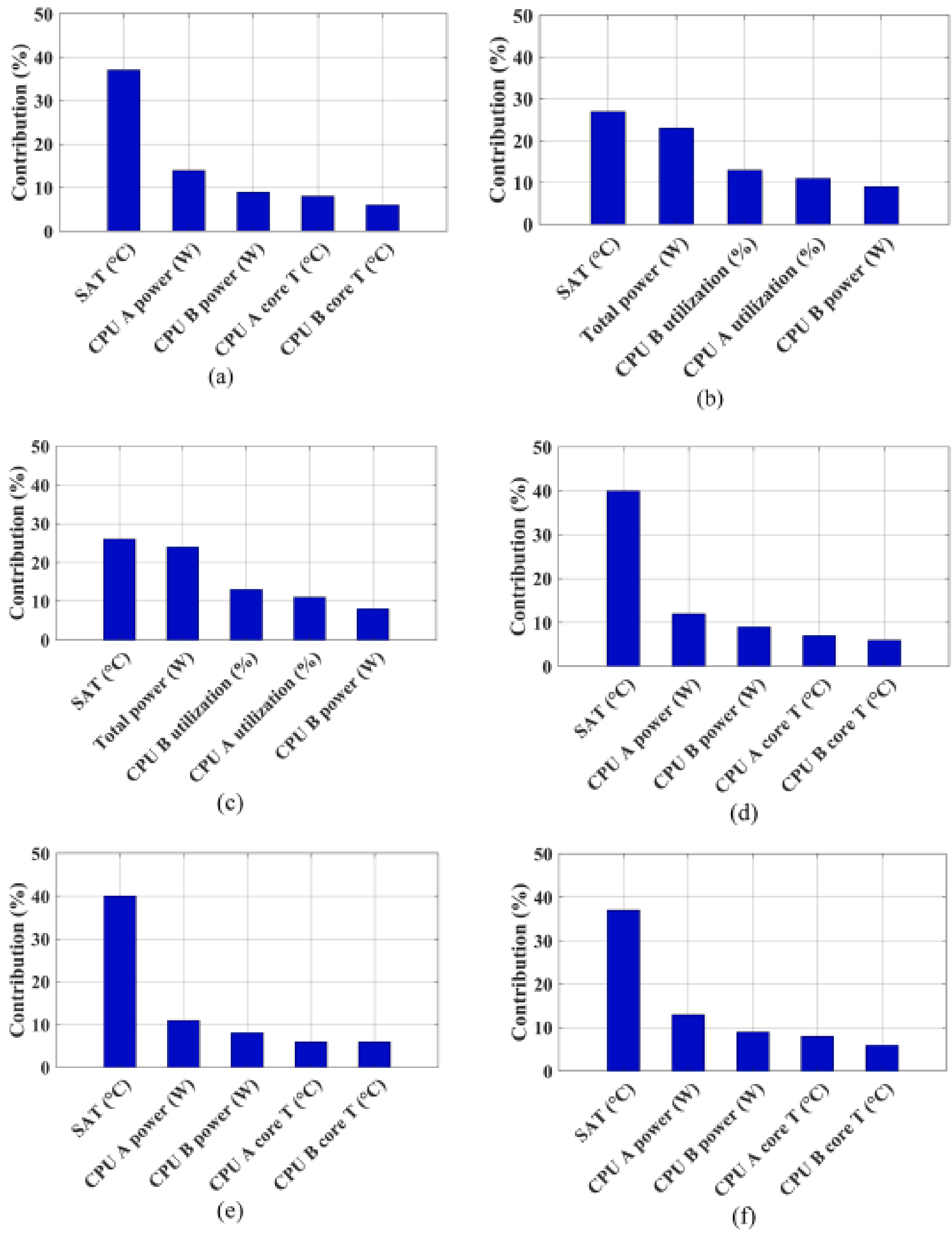


Fig. 11. Feature impact weight on fans rotational speed. (a) Fan 1 (b) Fan 2 (c) Fan 3 (d) Fan 4 (e) Fan 5 (f) Fan 6.

Fan speed prediction models

Notably, fan speed is critical for calculating the IT equipment airflow demand. Yet, fans themselves can have different control algorithms. Therefore, machine learning techniques were used to develop models that can predict the speed of fans individually. The machine learning algorithm is integrated with the CFD model, where the machine learning prediction is used as an input to the CFD model to develop a dynamic CFD model that can predict the performance of the IT equipment at different operating conditions. After that, a compact model (regression model) was extracted from this dynamic CFD model.

As mentioned earlier, RF provides a way to rank features based on their importance, because of this, RF s are also used as a feature selection method. The features ranked by the RF algorithm that impact the fan speed are illustrated in Fig. 11. The RF hyperparameters used throughout this study are as following: each RF consisted of 50 DTs, the maximum number of features considered when looking for the best split was six, the minimum number of samples required to split an internal node was set to two, and the maximum depth of each tree was set to none so that nodes are expanded until all leaves contain less than the minimum number of samples set, which is two.

Based on Fig. 11, the feature selection results showed that the most relevant contributor to all fans was the SAT. According to the similarity in the feature weight order, it was observed that there were two main fan control algorithms within the IT equipment. One was for fans 1, 4, 5, and 6; the second was for fans 2 and 3. The similarity was not only in the feature weight order, but also in the weight contribution. Fans 1, 4, 5, and 6 were mainly controlled by the SAT, while the rest of the features had minimal contribution. Meanwhile, fans 2 and 3 have a second major contributor beside the SAT, which was the total IT equipment power. This conclusion that multiple fans within the same IT equipment could have multiple fan control algorithm was confirmed by a server vendor company.

Some of the features used for developing the prediction models are not reported by all IT equipment. Additionally, some of the highly correlated or redundant features must be excluded from the model. For highly correlated features, having two or more independent variables (features) that are highly correlated with each other and share almost the same information, this is a phenomenon in statistics called multicollinearity. When present, multicollinearity can cause the estimate of one variable s impact on the dependent variable while controlling for the others tends to be less precise than if the independent variables were uncorrelated with one another. A commonly used way to address this issue is that only one of the highly correlated features needs to remain, while the rest are removed. Regarding redundant features, some datasets may include features that have little to no significance impact on the output and removing them may decrease the model s complexity and increase its performance and accuracy.

For these reasons, the number of features used in the fan speed prediction models were reduced. Feature reduction was carried out while considering the CFD boundary conditions as these data should be available to run the model. Initially, the prediction model was modified to predict the fan speed using the SAT, CPU utilization, and CPU power. After that, the model was customized to predict the fan speed based on just the SAT and CPU utilization. A total of three different models were developed, Model 1 considered all the features that remained after removing highly correlated and redundant features; Model 2 used the features SAT, CPU utilization, and CPU power; and Model 3 used the features SAT and CPU utilization. To evaluate their reliabilities, the R^2 performance measure was used, and 10-fold cross validation was done across all runs, where, as previously mentioned, training is carried out on nine folds and tested on the remaining fold, and this is repeated ten times, so that each fold is trained nine times and tested once. The average performance across all folds is taken. In addition, a validation set that was not used in the training phase of the models was used to validate the performances of feature selection and predictions. In regression problems, the most common performance metric for evaluating a machine learning model is the R^2 coefficient. The R^2 coefficient, also called the coefficient of determination, indicates the proportion of variance in the dependent variable that is predictable from the independent variables. It is a statistical measure of the goodness of fit of a model and is calculated using the following equation:

(11)

Where y_i is the true output value of observation i, y_i is the predicted value of observation i, and $\bar{y}$ is the average value of the dependent variable y. Generally, the closer R^2 value to unity, the better the model accuracy, and when R^2 1, it indicates that the regression model fits perfectly and explains all variance. For model 1, the calculated R^2 coefficients revealed that the fan speed predictions fit almost perfectly to the experimental data. The lowest R^2 was found to be 0.957, belonging to fan 4, while the highest was 0.985, belonging to fans 2 and 3. The lowest calculated R^2 value drops down to 0.77 when only two features were considered. R^2 values for the six fans that were obtained using the three prediction models are summarized in Table 3.

Table 3 shows that reducing the number of inputs for the fan speed prediction model decreased its prediction ability. Nonetheless, Model 2 showed an impressive performance since its lowest calculated R^2 value of all the fans was 0.84. In Model 3, when the CPU power was disregarded, the model s predictive ability deteriorated considerably and the lowest calculated R^2 value dropped to 0.77. Table 3 also shows that fans reacted differently when the number of inputs was reduced. This agrees with what was concluded from Fig. 11 that the fans have various control algorithms.

To further demonstrate RF s capability, its performance was compared with linear regression, lasso regression, support vector machine (SVM) with a linear kernel, and a SVM with a radial basis function (RBF) kernel. The R^2 for the different algorithms on Model 1 is shown in Table 4, where the RF achieved the highest performance for all fans.

As mentioned earlier, not all IT equipment reports the same features. Hence, in order to generalize the proposed approach and to facilitate adopting in building compact models for various IT equipment, a machine learning prediction model that predicts the fan speed using common features which are reported by different IT equipment should be considered. Accordingly, and since that Model 2 provided high R^2 values while considering only three features (SAT, CPU utilization, and CPU power), it was chosen to conduct investigations in the remainder of this study.

4.2. CFD results

After the power dissipation and fan speed prediction models were developed, their results were incorporated into the CFD model. By doing so, it was expected that the CFD model would replicate the IT equipment s performance at off-design conditions. For validation the CFD model was tested against the experimental results considering different cases. These cases were selected to cover a wide range of the IT equipment s operating conditions, as shown in Table 5. The corresponding predicted and experimental values of the fans speeds for these cases are

Table 3
Calculated R^2 for different fans speed prediction models.

Model	R^2 Fan 1	R^2 Fan 2	R^2 Fan 3	R^2 Fan 4	R^2 Fan 5	R^2 Fan 6
Model 1	0.959	0.985	0.985	0.957	0.960	0.958
Model 2	0.842	0.947	0.947	0.843	0.851	0.840
Model 3	0.772	0.924	0.923	0.775	0.785	0.770

Table 4
A comparison of the R^2 measure for different algorithms on Model 1.

Algorithm	R^2 Fan 1	R^2 Fan 2	R^2 Fan 3	R^2 Fan 4	R^2 Fan 5	R^2 Fan 6
RF	0.959	0.985	0.985	0.957	0.960	0.958
Linear Regression	0.620	0.864	0.862	0.630	0.640	0.622
Lasso Regression	0.620	0.864	0.862	0.630	0.640	0.622
Linear SVM	0.554	0.820	0.817	0.582	0.593	0.593
RBF SVM	0.778	0.926	0.926	0.783	0.797	0.778

Table 5
CPU utilization and SAT values for different cases used in testing the results.

Parameter	Case 1	Case 2	Case 3	Case 4
CPU Utilization (%)	Idling	25	50	100
SAT (C)	23	29	18	30

Table 6
Comparison between the experimental and the predicted fans speeds at the different cases.

Case	Fan	Experimental fan speed (RPM)	Predicted fan speed (RPM)	Relative error (%)
Case 1	Fan 1	2520	2520	0
	Fan 2	2280	2229.6	2.21
	Fan 3	2280	2289.6	0.42
	Fan 4	2160	2275.2	5.33
	Fan 5	2280	2364	3.68
	Fan 6	2400	2400	0
Case 2	Fan 1	3360	3453.6	2.8
	Fan 2	3360	3288	2.1
	Fan 3	3360	3110.4	7.4
	Fan 4	3000	3007.2	0.2
	Fan 5	2880	2990.4	3.8
	Fan 6	3240	3261.6	0.7
Case 3	Fan 1	2040	2049.6	0.5
	Fan 2	2640	2685.6	1.7
	Fan 3	2640	2954.4	11.9
	Fan 4	1920	1876.8	2.3
	Fan 5	1920	1932	0.6
	Fan 6	1920	1972.8	2.8
Case 4	Fan 1	4680	4588.8	1.9
	Fan 2	5880	5584.8	5
	Fan 3	5880	5532	5.9
	Fan 4	4080	4027.2	1.3
	Fan 5	4080	4051.2	0.7
	Fan 6	4560	4560	0

shown in Table 6.

The fan speed model showed an impressive prediction ability, seeing as the maximum absolute error was found to be less than 12%, as provided in Table 6. The impact of the 11.9% discrepancy between the experimental and predicted fan speeds on the overall mode accuracy is expected to be minimal. This is attributed to the fact that each fan contributed with almost $\frac{1}{Totalnumberoffans}$ when calculating the total IT equipment flowrate. Thus, even if one or two of the fans has a relatively large error, the overall system will continue to show great prediction ability.

A comparison of the numerical results and the field measurements, in terms of outlet air velocity and temperature, is provided in Table 7. It can be inferred from this Table 7 that there was great consistency between the results in all cases. For the outlet air velocity, the maximum error was found to be 5.4% and the minimum was 1.8%. For the outlet air temperature, the maximum error was found to be 4.6% and the minimum was 1.1%. This confirms what were discussed earlier in this section, where case 3 exhibited 11.9% absolute relative error in predicting one of the fans speeds, however, CFD results showed only 2.4% and 1.6% error in the overall IT equipment outlet air velocity and temperature, respectively. Based on this, it was found that the reliability and accuracy of the CFD model applied in this work is comprehensively validated.

Thereafter, the model was utilized to calculate the IT equipment s flowrate and outlet air temperature at various operating conditions as shown in Fig. 12. As expected, Fig. 12 (a) reveals that the air flowrate increased when the SAT and the CPU utilization increased, which was due to the increase in fan speed. At some points when the flowrate increased the OAT exhibited a slight variation or even stayed constant, even when the SAT increased as more flowrate was delivered. Furthermore, it can be inferred from Fig. 12 (b) that the OAT was almost identical for different CPU utilizations, which meant that the fans speeds were being adjusted to maintain a certain thermal field inside the IT equipment at different SAT regardless of CPU utilization.

It can be also noted from Fig. 12 (b) that the air temperature difference between the server inlet and outlet is rather low. This highlights that this specific IT equipment is adopting excessively conservative fan control algorithms by running the fans at unnecessary high speed. Adopting such algorithms result in wasting a significant amount of energy by the IT equipment fans, especially when hundreds of such IT equipment are operating in different data centers. This leaves a huge spot for improving and optimizing IT equipment fan control algorithms. For example, the work conducted by J. Sarkinen, R. Brannvall, J. Gustafsson, and J. Summers [82] in which they specified an optimal operating condition considering the components power losses and the fan power consumption. In addition, this relatively low T across the IT equipment could cause energy losses by the data centers cooling equipment in most cases. This attributed to fact that more air is required by the IT equipment to maintain this T across the IT equipment, and hence more air consumed by the CRAH units blowers. More in-depth discussion can be found in [82].

Table 7
Experimental and CFD results for the outlet air velocity and OAT.

Case	Speed			OAT		
	Experimental (m/ s)	CFD (m/ s)	Error (%)	Experimental (C)	CFD (C)	Error (%)
Case 1	0.37 0.01	0.39	5.4	27.1 1	26.8	1.1
Case 2	0.55 0.02	0.54	1.8	33.56 1	32	4.6
Case 3	0.41 0.01	0.4	2.4	24.8 1	24.4	1.6
Case 4	0.9 0.03	0.86	4.4	34.5 1	33.5	2.9

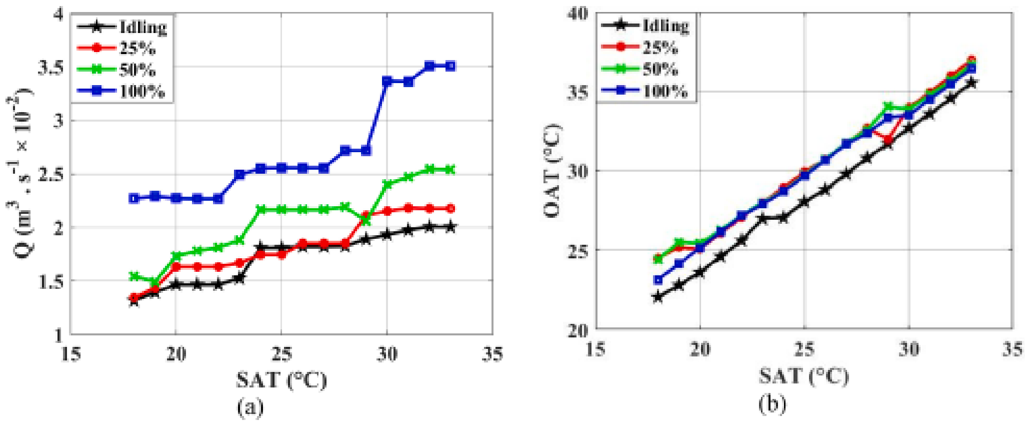


Fig. 12. Effect of varying CPU utilization and SAT on the IT equipment's (a) flowrate (b) OAT.

Finally, it is essential to elaborate on adopting the CFD simulation instead of using the fan and the conservation of energy laws for calculating the IT equipment flow rate and the outlet temperature. Even if the power consumption model and fan speed prediction models can be used to calculate the IT equipment flow rate and the outlet temperature using these laws, they fail to capture the full physics inside the server, which will affect the overall model accuracy. For example, a significant amount of flow was found to recirculate within the tested IT equipment's chassis, as shown in Fig. 13. As a result, the air flow passing through the IT equipment was less than the total flow driven by the fans, which meant that the air flow leaving the IT equipment was at a higher temperature.

4.3. Developing IT equipment compact model

After developing the IT equipment power and fan speed prediction models and then conducting the CFD study, this work was extended to use regression tools to derive correlations between the response and input parameters. To that end, a prediction model that described the overall IT equipment performance at different operating conditions was obtained. This model, which represents the IT equipment compact model, was used to predict the IT equipment flowrate and OAT based on the SAT and CPU utilization. Initially, multiple linear regression, which is a more general form of simple linear regression, was adopted. The regression function was developed from 64 CFD data points, as shown in Fig. 12. The general mathematical model for multiple linear regression is given by:

$$Y = \beta_0 + \sum_i (\beta_i {}^* x_i) + \epsilon \tag{12}$$

By using this model, two multiple linear regression functions were developed to predict the values of the IT equipment's OAT and flowrate. To evaluate these regression functions, the R^2 values were calculated with respect to the CFD results. High R^2 values for the OAT and flowrate were obtained from the multiple linear regression function. To further improve the regression function's accuracy, nonlinear multivariate regression models were developed to consider the non-linearity in the IT equipment flowrate data generated from the CFD. The basic form of this model is given by:

$$Y = f(\beta_i, x_i) + \epsilon \tag{13}$$

This model's performance was superior to the multiple linear regression model, hence it was adopted later on in the analysis. However, the difference was not significant. The R^2 improved by 2% for the IT equipment's flowrate, while the R^2 for the OAT only increased by 0.9%. Table 8 shows the regression functions derived from both models and a summary of the statistical analysis.

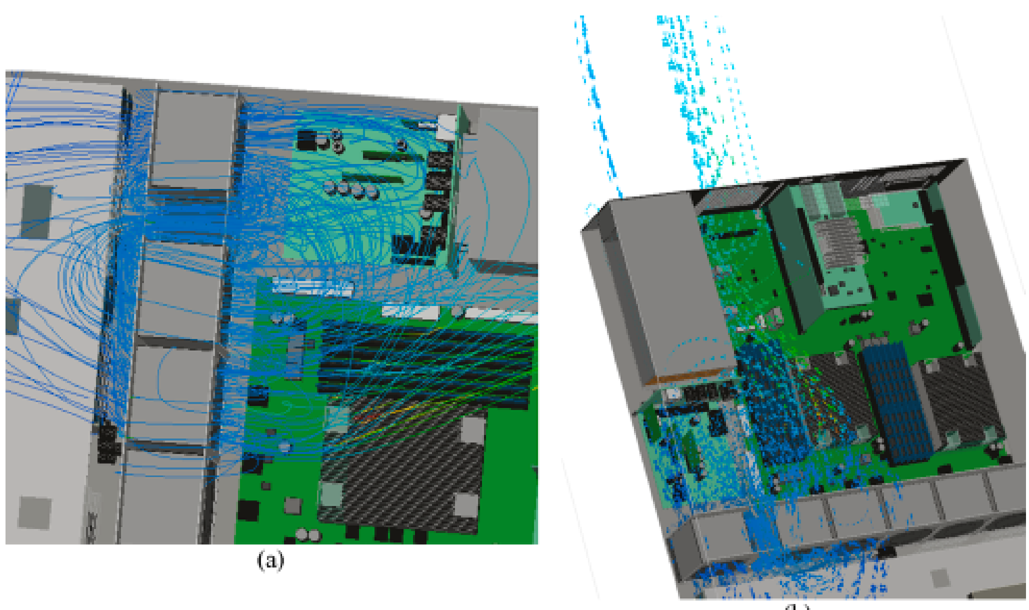


Fig. 13. Streamlines of airflow recirculating inside the IT equipment.

Table 8
Regression function developed for the IT equipment flowrate and OAT prediction.

	Parameter	Equation	R^2	Adjusted R^2	MAE
Multiple linear regression	Q	Q 0.01023*$utilization$ 0.06469*SAT 0.02003	0.912	0.909	5.4
	OAT	OAT 0.01038*$utilization$ 0.86152*SAT 7.3844	0.974	0.973	2
Multivariate nonlinear regression	Q	Q 0.0006694*$Utilization^{1.59914}$ 0.000341695*$SAT^{2.37342}$ 0.932246	0.93	0.929	5.1
	OAT	OAT 124.109*$Utilization^{0.00142692}$ 0.749164*$SAT^{1.03308}$ 116.357	0.983	0.983	1.7

For the latter model, both the R^2 and adjusted R^2 values were very close to unity, which verified the quality of fit and the model s ability to predict new operating points. Moreover, the R^2 and adjusted R^2 values were high and similar, which indicated that the model was not overfitted. Finally, the function accurately predicted the IT equipment s flowrate and OAT, with an MAE of 5.1 and 1.7, respectively.

After conducting a diagnostic of the regression model and validating it statistically, its accuracy was assessed comparatively using the experimental data. The experimental results used for comparison were those obtained in the four cases discussed in the previous section. Table 9 and Table 10 present a comparison of the results obtained through the experiment and those predicted by the regression model for the IT equipment s flowrate and the OAT, respectively. Unfortunately, to the best of the authors knowledge, no article has created a compact model which predicts the flowrate required for the IT equipment and the air temperature leaving the equipment. Therefore, to further assessing the developed compact model, these two Tables compared the developed regression function with state-of-the-art compact for the same IT equipment. This state-of-the-art model is currently used in some CFD software, which adopts generalized and linear correlation for the various IT equipment.

For the regression model, the maximum error was found to be 11.4% and 5.7% for the IT equipment s flowrate and OAT, respectively. This confirms its accuracy and ability to predict the IT equipment s thermal performance at off-design operating conditions. For the state-of-the-art compact model, the discrepancy between its results and the experimental results was substantial. In some cases, the flowrate calculated by this model was more than double that of the experimental results. For the OAT, slightly higher than an 8 C difference was observed. The regression model developed in this work reduced the error for the IT equipment s flowrate to 5.2% that of the state-of-the-art IT equipment compact model. For the OAT, it reduced error to 9.3 % that of the state-of-the-art compact model.

With hundreds of IT equipment components installed at the data center facility level, the impact of any error introduced by the state- of the art compact model is certain to grow considerably. As a result, calibration experiments are seen as necessary for improving a given model s reliability. However, the method used for developing an IT equipment compact model in this study can potentially eliminate the need for calibration. It has been shown that this methodology can tremendously improve the CFD model s accuracy at the data center level and its ability to predict the IT equipment s thermal performance even at off-design conditions.

Table 9
Experimental and predicted flowrate for the developed model and state of the art model.

Case	Experimental Q (m^3/ s × 10^{-2})	Developed model Q (m^3/ s × 10^{-2})	Developed model Error (%)	State-of-the-art compact model Q (m^3/ s × 10^{-2})	State-of-the-art compact model Error (%)
Case 1	1.4 0.04	1.5	7.1	1.96	40
Case 2	2.2 0.07	2	9.1	2.13	3.2
Case 3	1.6 0.05	1.7	6.2	3.52	120
Case 4	3.5 0.11	3.1	11.4	8.94	155.4

Table 10
Experimental and predicted OAT for the developed model and state of the art model.

Case	Experimental OAT (C)	Developed model OAT (C)	Developed model Error (%)	State-of-the-art compact model OAT (C)	State-of-the-art compact model Error (%)
Case 1	27.1 1	26.9	0.7	24	11.4
Case 2	33.6 1	31.7	5.7	40.8	21.4
Case 3	24.8 1	24.1	2.8	32.3	30.2
Case 4	34.5 1	33.7	2.3	41.3	19.7

5. Conclusions

Proper thermal and energy management of data centers requires a thorough understanding of the basic unit for power and heat flow in data centers, which is the IT equipment. This study establishes a comprehensive methodology that builds an accurate compact model for the IT equipment. Initially, a power consumption model and fan speed prediction models were derived from the IT equipment data. Merging these models with the CFD code results in an IT equipment CFD model that can predict the IT equipment at off-design conditions. To verify the reliability of this model it was tested against the experimental measurements under four different operating conditions. The maximum mismatch between the CFD model and the experiments was found to be 5.4% and 4.6% in terms of air velocity and air temperature, respectively. Finally, a compact model for the IT equipment was created using the data generated by the CFD model. Compared to the experiment, the developed model showed a maximum inconsistency in the IT equipment flowrate and outlet air temperature of 11.4% and 5.7%, respectively. This model is superior to existing IT equipment compact models for the following reasons:

- It can be easily embedded with any CFD code or any other simulation program used in designing energy systems for data centers.
- This approach can be adapted to build a compact model for any IT equipment.
- The model can simulate the IT equipment performance at different supply air temperature and CPU utilization.
- This model does not only predict the power consumption of the IT equipment, but also the amount of flow required by this equipment and its OAT.
- This model can be derived using tools and data that are easily accessible.
- As this compact model is developed from a full physics based CFD model, it is expected to be the most accurate IT equipment compact model. Compared with a state-of-the-art IT equipment compact model, the maximum error in the flow rate was reduced from 155.4%

to 11.4%. For the outlet air temperature, the maximum error was reduced from 30.2% to 5.7%.

CRediT authorship contribution statement

Yaman M. Manaserh: Conceptualization, Methodology, Formal analysis, Investigation, Supervision, Resources, Data Curation, Visualization, Validation, Software, Writing original draft, Writing - Review & Editing. **Mohammad I. Tradat:** Methodology, Formal analysus, Investigation, Resources, Data Curation, Validation, Software, Supervision, Writing original draft, Writing - Review & Editing. **Dana Bani-Hani:** Methodology, Formal analysis, Investigation, Software, Validation, Visualization, Writing original draft, Writing - Review & Editing. **Aseel Alfallah:** Methodology, Formal analysis, Investigation, Validation, Visualization, Software, Writing original draft, Writing - Review & Editing. **Bahgat G. Sammakia:** Supervision, Project administration. **Kourosh Nemati:** Supervision, Project administration. **Mark J. Seymour:** Supervision, Project administration.

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.

Acknowledgments

We would like to acknowledge Future Facilities Ltd. We would also like to thank the ES2 Partner Universities for their support and advice. This work is supported by NSF IUCRC Award No. IIP- 1738793 and MRI Award No. CNS1040666.

References

[1] Abubaker AM, Ahmad AD, Singh BB, Akafuah NK, Saito K. Multi-objective linear-regression-based optimization of a hybrid solar-gas turbine combined cycle with absorption inlet-air cooling unit. Energy Convers Manage 2021;240:114266. https://doi.org/10.1016/j.enconman.2021.114266.
[2] Ahmad AD, Abubaker AM, Najjar YS, Manaserh YMA. Power boosting of a combined cycle power plant in Jordan: An integration of hybrid inlet cooling & solar systems. Energy Convers Manage 2020;214:112894.
[3] Al-Ghussain L, Abubaker AM, Darwish Ahmad A. Superposition of Renewable-Energy Supply from Multiple Sites Maximizes Demand-Matching: Towards 100% Renewable Grids in 2050. Appl Energy 2021;284:116402. https://doi.org/10.1016/j.apenergy.2020.116402.
[4] Manaserh YMA, Abubaker AM, Ahmad AD, Ata AB, Najjar YSH, Akafuah NK. Assessment of integrating hybrid solar-combined cycle with thermal energy storage for shaving summer peak load and improving sustainability. Sustainable Energy Technol Assess 2021;47:101505. https://doi.org/10.1016/j.seta.2021.101505.
[5] Sharma CS, Tiwari MK, Zimmermann S, Brunschwiler T, Schlottig G, Michel B, et al. Energy efficient hotspot-targeted embedded liquid cooling of electronics. Appl Energy 2015;138:414 22.
[6] Ni J, Bai X. A review of air conditioning energy performance in data centers. Renew Sustain Energy Rev 2017;67:625 40.
[7] Habibi Khalaj A, Halgamuge SK. A Review on efficient thermal management of air- and liquid-cooled data centers: From chip to the cooling system. Appl Energy 2017; 205:1165 88.
[8] Manaserh YM, Tradat MI, Mohsenian G, Sammakia BG, Seymour MJ. General guidelines for commercialization a small-scale in-row cooled data center: a case study. In: 2020 36th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM): IEEE; 2020. p. 48 55.
[9] Radmard V, Hadad Y, Azizi A, Rangarajan S, Hoang CH, Arvin C, et al. Direct Micro-Pin Jet Impingement Cooling for High Heat Flux Applications. In: 2020 36th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM): IEEE; 2020. p. 1 9.
[10] Hoang CH, Rangarajan S, Khalili S, Ramakrisnan B, Radmard V, Hadad Y, et al. Hybrid microchannel/multi-jet two-phase heat sink: A benchmark and geometry optimization study of commercial product. Int J Heat Mass Transf 2021;169: 120920. https://doi.org/10.1016/j.ijheatmasstransfer.2021.120920.
[11] Manaserh YM, Tradat MI, Mohsenian G, Sammakia BG, Ortega A, Seymour MJ, et al. Novel Experimental Methodology for Characterizing Fan Performance in Highly Resistive Environments. In: 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm): IEEE; 2020. p. 1 7.
[12] Gupta R, Asgari S, Moazamigoodarzi H, Pal S, Puri IK. Cooling architecture selection for air-cooled Data Centers by minimizing exergy destruction. Energy 2020;201:117625. https://doi.org/10.1016/j.energy.2020.117625.
[13] Moazamigoodarzi H, Tsai PJ, Pal S, Ghosh S, Puri IK. Influence of cooling architecture on data center power consumption. Energy 2019;183:525 35.
[14] Song Z. Studying the fan-assisted cooling using the Taguchi approach in open and closed data centers. Int J Heat Mass Transf 2017;111:593 601.
[15] Lu H, Zhang Z. Numerical and experimental investigations on the thermal performance of a data center. Appl Therm Eng 2020;180:115759. https://doi.org/10.1016/j.applthermaleng.2020.115759.
[16] Meng X, Zhou J, Zhang X, Luo Z, Gong H, Gan T. Optimization of the thermal environment of a small-scale data center in China. Energy 2020;196:117080. https://doi.org/10.1016/j.energy.2020.117080.
[17] Zhang Y, Zhang K, Liu J, Kosonen R, Yuan X. Airflow uniformity optimization for modular data center based on the constructal T-shaped underfloor air ducts. Appl Therm Eng 2019;155:489 500.
[18] Almoli A, Thompson A, Kapur N, Summers J, Thompson H, Hannah G. Computational fluid dynamic investigation of liquid rack cooling in data centres. Appl Energy 2012;89(1):150 5.
[19] Tatchell-Evans M, Kapur N, Summers J, Thompson H, Oldham D. An experimental and theoretical investigation of the extent of bypass air within data centres employing aisle containment, and its impact on power consumption. Appl Energy 2017;186:457 69.
[20] Cho J, Woo J. Development and experimental study of an independent row-based cooling system for improving thermal performance of a data center. Appl Therm Eng 2020;169:114857. https://doi.org/10.1016/j.applthermaleng.2019.114857.
[21] Silva-Llanca L, Ortega A, Fouladi K, del Valle M, Sundaralingam V. Determining wasted energy in the airside of a perimeter-cooled data center via direct computation of the Exergy Destruction. Appl Energy 2018;213:235 46.
[22] Zhou C, Yang C, Wang C, Zhang X. Numerical simulation on a thermal management system for a small data center. Int J Heat Mass Transf 2018;124: 677 92.
[23] Athavale J, Yoda M, Joshi Y. Comparison of data driven modeling approaches for temperature prediction in data centers. Int J Heat Mass Transf 2019;135:1039 52.
[24] Habibi Khalaj A, Scherer T, Siriwardana J, Halgamuge SK. Multi-objective efficiency enhancement using workload spreading in an operational data center. Appl Energy 2015;138:432 44.
[25] Ye X, Li P, Li C, Ding X. Numerical investigation of blade tip grooving effect on performance and dynamics of an axial flow fan. Energy 2015;82:556 69.
[26] Ghosh R, Joshi Y. Error estimation in POD-based dynamic reduced-order thermal modeling of data centers. Int J Heat Mass Transf 2013;57(2):698 707.
[27] Fouladi K, Wemhoff AP, Silva-Llanca L, Abbasi K, Ortega A. Optimization of data center cooling efficiency using reduced order flow modeling within a flow network modeling approach. Appl Therm Eng 2017;124:929 39.
[28] Samadiani E, Joshi Y. Proper orthogonal decomposition for reduced order thermal modeling of air cooled data centers. J Heat Transfer 2010;132.
[29] Samadiani E, Joshi Y, Hamann H, Iyengar MK, Kamalsy S, Lacey J. Reduced order thermal modeling of data centers via distributed sensor data. J Heat Transfer 2012; 134.
[30] Khalid R, Joshi Y, Wemhoff A. Rapid modeling tools for energy analysis of modular data centers. In: 2016 15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm): IEEE; 2016. p. 1444 52.
[31] Khalid R, Wemhoff AP. Thermal Control Strategies for Reliable and Energy-Efficient Data Centers. J Electron Packag 2019;141.
[32] Zapater M, Risco-Martín JL, Arroba P, Ayala JL, Moya JM, Hermida R. Runtime data center temperature prediction using Grammatical Evolution techniques. Appl Soft Comput 2016;49:94 107.
[33] Fulpagare Y, Joshi Y, Bhargav A. Rack level forecasting model of data center. In: 2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm): IEEE; 2017. p. 824 9.
[34] Wang L, von Laszewski G, Huang F, Dayal J, Frulani T, Fox G. Task scheduling with ANN-based temperature prediction in a data center: a simulation-based study. Engineering with Computers 2011;27(4):381 91.
[35] Chen J, Tan R, Wang Y, Xing G, Wang X, Wang X, et al. A high-fidelity temperature distribution forecasting system for data centers. 2012 IEEE 33rd Real-Time Systems Symposium: IEEE; 2012. p. 215-24.
[36] Moazamigoodarzi H, Pal S, Ghosh S, Puri IK. Real-time temperature predictions in it server enclosures. Int J Heat Mass Transf 2018;127:890 900.
[37] Jin C, Bai X, Yang C, Mao W, Xu X. A review of power consumption models of servers in data centers. Appl Energy 2020;265:114806. https://doi.org/10.1016/j.apenergy.2020.114806.
[38] Cheung H, Wang S, Zhuang C, Gu J. A simplified power consumption model of information technology (IT) equipment in data centers for energy system real-time dynamic simulation. Appl Energy 2018;222:329 42.
[39] VanGilder JW, Pardey ZM, Zhang X, Healey C. Experimental measurement of server thermal effectiveness for compact transient data center models. International Electronic Packaging Technical Conference and Exhibition. American Society of Mechanical Engineers 2013.
[40] Moazamigoodarzi H, Gupta R, Pal S, Tsai PJ, Ghosh S, Puri IK. Modeling temperature distribution and power consumption in IT server enclosures with row-based cooling architectures. Appl Energy 2020;261:114355. https://doi.org/10.1016/j.apenergy.2019.114355.
[41] Salih Erden H, Ezzat Khalifa H, Schmidt RR. Determination of the lumped-capacitance parameters of air-cooled servers through air temperature measurements. J Electron Packag 2014;136.

[42] Wang W, Abdolrashidi A, Yu N, Wong D. Frequency regulation service provision in data center with computational flexibility. Appl Energy 2019;251:113304. https://doi.org/10.1016/j.apenergy.2019.05.107.

[43] Chu W-X, Wang C-C. A review on airflow management in data centers. Appl Energy 2019;240:84 119.

[44] Ghazal Mohsenian SK, Mohammad Tradat, Yaman Manaserh, Srikanth Rangaranjan, Anuroop Desu, Dushyant Thakur4, Koroush Nemati, Kanad Ghose, Bahgat Sammakia. A Novel Integrated Fuzzy Control System toward Automated Local Airflow Management in Data Centers. Control Engineering Practice. 2021;In press (accepted).

[45] Dai J, Ohadi MM, Das D, Pecht MG. OPTIMUM COOLING OF DATA. CENTERS: Springer; 2016.

[46] Vasques TL, Moura P, de Almeida A. A review on energy efficiency and demand response with focus on small and medium data centers. Energ Effi 2019;12(5): 1399 428.

[47] Tang C-J, Dai M-R, Chuang C-C, Chiu Y-S, Lin WS. A load control method for small data centers participating in demand response programs. Future Generation Computer Systems 2014;32:232 45.

[48] Ham S-W, Kim M-H, Choi B-N, Jeong J-W. Simplified server model to simulate data center cooling energy consumption. Energy Build 2015;86:328 39.

[49] Fang Q, Wang J, Gong Qi. Qos-driven power management of data centers via model predictive control. IEEE Trans Autom Sci Eng 2016;13(4):1557 66.

[50] Bani-Hani D, Khasawneh M. A recursive general regression neural network (R-GRNN) oracle for classification problems. Expert Syst Appl 2019;135:273 86.

[51] Ho TK. Random decision forests. Proceedings of 3rd international conference on document analysis and recognition: IEEE; 1995. p. 278-82.

[52] Valentini G, Masulli F. Ensembles of learning machines. Italian workshop on neural nets. Springer; 2002. p. 3 20.

[53] Rodriguez-Galiano V, Mendes MP, Garcia-Soldado MJ, Chica-Olmo M, Ribeiro L. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). Sci Total Environ 2014;476-477:189 206.

[54] Zhang W, Wu C, Li Y, Wang L, Samui P. Assessment of pile drivability using random forest regression and multivariate adaptive regression splines. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 2021;15(1):27 40.

[55] Iannace G, Ciaburro G, Trematerra A. Wind turbine noise prediction using random forest regression. Machines 2019;7(4):69. https://doi.org/10.3390/machines7040069.

[56] Zhou Y, Li S, Zhou C, Luo H. Intelligent approach based on random forest for safety risk prediction of deep foundation pit in subway stations. J Comput Civil Eng 2019; 33(1):05018004. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000796.

[57] Zhou J, Li X, Mitri HS. Classification of rockburst in underground projects: comparison of ten supervised learning methods. J Comput Civil Eng 2016;30(5): 04016003. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000553.

[58] Priyadumkol J, Kittichaikarn C. Application of the combined air-conditioning systems for energy conservation in data center. Energy Build 2014;68:580 6.

[59] Ling Y-Z, Zhang X-S, Zhang K, Jin X. On the characteristics of airflow through the perforated tiles for raised-floor data centers. Journal of Building Engineering 2017; 10:60 8.

[60] Saini S, Shahi P, Bansode P, Siddarth A, Agonafer D. CFD Investigation of Dispersion of Airborne Particulate Contaminants in a Raised Floor Data Center. In: 2020 36th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM): IEEE; 2020. p. 39 47.

[61] Niazmand A, Chauhan T, Saini S, Shahi P, Bansode PV, Agonafer D. CFD Simulation of Two-Phase Immersion Cooling Using FC-72 Dielectric Fluid. ASME 2020 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems: American Society of Mechanical Engineers Digital Collection. 2020.

[62] Radmard V, Hadad Y, Rangarajan S, Hoang CH, Fallahtafti N, Arvin CL, et al. Multi-objective optimization of a chip-attached micro pin fin liquid cooling system. Appl Therm Eng 2021;195:117187. https://doi.org/10.1016/j.applthermaleng.2021.117187.

[63] Achouri R, Mokni I, Mhiri H, Bournot P. A 3D CFD simulation of a self inducing Pitched Blade Turbine Downflow. Energy Convers Manage 2012;64:633 41.

[64] Zhao H, Kang C, Ding K, Zhang Y, Li B. Transient startup characteristics of a drag-type hydrokinetic turbine rotor. Energy Convers Manage 2020;223:113287. https://doi.org/10.1016/j.enconman.2020.113287.

[65] Xu Z, Feng Y-H, Zhao C-Y, Huo Y-L, Li S, Hu X-J, et al. Experimental and numerical investigation on aerodynamic performance of a novel disc-shaped wind rotor for the small-scale wind turbine. Energy Convers Manage 2018;175:173 91.

[66] Mohamed MH. Performance investigation of H-rotor Darrieus turbine with new airfoil shapes. Energy 2012;47(1):522 30.

[67] Shirazi AT, Nazari MR, Manshadi MD. Numerical and experimental investigation of the fluid flow on a full-scale pump jet thruster. Ocean Eng 2019;182:527 39.

[68] Lee JH, Moshfeghi M, Choi YK, Hur N. A numerical simulation on recirculation phenomena of the plume generated by obstacles around a row of cooling towers. Appl Therm Eng 2014;72(1):10 9.

[69] Huang M-H, Chen L, Lei Le, He P, Cao J-J, He Y-L, et al. Experimental and numerical studies for applying hybrid solar chimney and photovoltaic system to the solar-assisted air cleaning system. Appl Energy 2020;269:115150. https://doi.org/10.1016/j.apenergy.2020.115150.

[70] Calautit JK, Hughes BR, Nasir DSNM. Climatic analysis of a passive cooling technology for the built environment in hot countries. Appl Energy 2017;186: 321 35.

[71] Manaserh YM, Tradat MI, Hoang CH, Sammakia BG, Ortega A, Nemati K, et al. Degradation of Fan Performance in Cooling Electronics: Experimental Investigation and Evaluating Numerical Techniques. Int J Heat Mass Transf 2021;174:121291. https://doi.org/10.1016/j.ijheatmasstransfer.2021.121291.

[72] Conner ME, Baglietto E, Elmahdi AM. CFD methodology and validation for single-phase flow in PWR fuel assemblies. Nucl Eng Des 2010;240(9):2088 95.

[73] Tradat MI, Manaserh Y A, Sammakia BG, Hoang CH, Alissa HA. An experimental and numerical investigation of novel solution for energy management enhancement in data centers using underfloor plenum porous obstructions. Appl Energy 2021;289:116663. https://doi.org/10.1016/j.apenergy.2021.116663.

[74] Lopez V, Hamann HF. Heat transfer modeling in data centers. Int J Heat Mass Transf 2011;54(25-26):5306 18.

[75] Fulpagare Y, Bhargav A, Joshi Y. Dynamic thermal characterization of raised floor plenum data centers: Experiments and CFD. Journal of Building Engineering 2019; 25:100783. https://doi.org/10.1016/j.jobe.2019.100783.

[76] Chu W-X, Tsai M-K, Jan S-Y, Huang H-H, Wang C-C. CFD analysis and experimental verification on a new type of air-cooled heat sink for reducing maximum junction temperature. Int J Heat Mass Transf 2020;148:119094. https://doi.org/10.1016/j.ijheatmasstransfer.2019.119094.

[77] Li C, Li X, Li P, Ye X. Numerical investigation of impeller trimming effect on performance of an axial flow fan. Energy 2014;75:534 48.

[78] Basmadjian R, Ali N, Niedermeier F, De Meer H, Giuliani G. A methodology to predict the power consumption of servers in data centres. Proceedings of the 2nd international conference on energy-efficient computing and networking2011. p. 1-10.

[79] Perumal V, Subbiah S. Power-conservative server consolidation based resource management in cloud. Int J Network Manage 2014;24(6):415 32.

[80] Chen Y, Das A, Qin W, Sivasubramaniam A, Wang Q, Gautam N. Managing server energy and operational costs in hosting centers. Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems2005. p. 303-14.

[81] Elnozahy EM, Kistler M, Rajamony R. Energy-efficient server clusters. International workshop on power-aware computer systems: Springer; 2002. p. 179-97.

[82] Sarkinen J, Brannvall R, Gustafsson J, Summers J. 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): IEEE; 2020. p. 341 9.