



OPTIMIZATION MODEL FOR COOLING ENERGY REDUCTION IN URBAN DATA CENTERS THROUGH PHOTOVOLTAIC INTEGRATION

Aldair R. Auccapuma¹, Antonio F.C. Maia², Cristhian A. Lopez³, Joel D. Melo⁴

¹Federal University of ABC – aldair.auccapuma@ufabc.edu.br; ²Federal University of ABC – antonio.maia@ufabc.edu.br

³Federal University of ABC – anderson.lopez@ufabc.edu.br; ⁴ Federal University of ABC – joel.melo@ufabc.edu.br

ABSTRACT

The rising energy demand of data centers, primarily driven by intensive cooling requirements, presents significant economic and environmental challenges in densely populated urban areas. This paper presents an optimization model designed to reduce cooling energy cost in small urban data centers by integrating photovoltaic generation with advanced thermal management strategies. The proposed model incorporates a simplified heat transfer characterization that accounts for environmental heat exchanges and air-conditioning dynamics. The proposal was evaluated using 15-minute climate data from the São Paulo metropolitan region to capture daily and seasonal variations in cooling loads over a full operational year. Results indicate the optimization model achieved monthly costs of 129.58 USD in November and 95.16 USD in April, while maintaining safe CPU operating temperatures (≤ 72 °C). Photovoltaic generation supplied a significant share of the cooling demand, reducing dependence on the power grid and demonstrating the feasibility of solar integration for sustainable data center operations.

Keywords: Data Centers, Thermal Management, Photovoltaic Systems.

1 INTRODUCTION

The rapid growth of digital infrastructure, driven by cloud computing and artificial intelligence, has established data centers as strategic pillars of the modern digital economy (SHAH et al., 2025). However, this advancement comes with a substantial increase in global energy consumption, positioning data centers among the largest electricity users worldwide (MYTTON et al., 2022). Their electrical energy demand is expected to grow further due to the expansion of computational applications and the increasing digitalization of business and social processes.

Among the subsystems within a data center, the air conditioning system (HVAC) accounts for a significant portion of total energy use, reaching up to 40% in conventional facilities (OFFUTT et al., 2025). This high demand stems from the need to maintain strict environmental conditions, particularly temperature control, to ensure reliable and safe IT operation (LIN et al., 2018). Elevated temperatures not only increase the risk of server failures but also trigger backup units, further reducing energy efficiency (AHMED et al., 2023).

Accurate thermal modeling is essential for energy efficiency strategies, enabling simulation of workload variations, cooling configurations, and external conditions (CUPPELLI et al., 2018). Models that couple CPU and server thermal dynamics with the environment allow more effective optimization of energy use (ZHANG et al., 2025). Operational reliability is closely tied to thermal management, as degradation can compromise service performance and availability, especially under high service-level requirements (AHMED et al., 2023).



In this context, integrating renewable energy sources, particularly photovoltaic (PV) generation, emerges as a promising solution to reduce dependence on the power grid and mitigate environmental impacts (ZHANG et al., 2025). Solar energy, being clean, abundant, and increasingly cost-effective, can partially supply the cooling demand, which is highly sensitive to temporal variations in thermal load.

This work proposes an energy cost optimization model for data centers, aimed at reducing air conditioning energy consumption through the strategic integration of local photovoltaic generation. The model incorporates a dynamic representation of thermal demand and its effect on energy consumption, based on energy balances and thermal coupling between the CPU, servers, and environment, building on established methodologies for thermal demand modeling. The objective is to quantify the economic and energy benefits of the solution and assess its technical feasibility, contributing to the sustainable transition of the IT sector.

2 METHODOLOGY

A. Thermal System Modeling

A.1) Cold Aisle Temperature: The variation of the cold aisle air temperature within the container is described by a thermal energy balance. The model accounts for the energy supplied by the HVAC system, the energy extracted by the server racks, convective exchanges with the cold wall, and heat generated by auxiliary equipment (Eq. 1).

$$T_t^{cr} = T_{t-1}^{cr} + \frac{1}{C_{cr}} \left(\dot{m}_{HVAC} \cdot c_{air} \cdot (T_{t-1}^{in} - T_{t-1}^{cr}) + \frac{T_{t-1}^{cw} - T_{t-1}^{cr}}{R_{cw,cr}} + \delta_1 \right) \cdot \tau \quad \text{Eq. (1)}$$

$$C_{cw} = \rho_{air} \cdot V_{cr} \cdot c_{air} \quad \text{Eq. (2)}$$

In these expressions, T_t^{cr} denotes the cold aisle temperature at time step t , τ is the simulation time interval, and C_{cr} represents the thermal capacitance of the aisle air, computed from the product of air density (ρ_{air}), aisle volume (V_{cr}), and air specific heat (c_{air}). The term $\dot{m}_{HVAC}$ is the mass flow rate of the cooling air, T_{t-1}^{in} is the inlet air temperature to the cold aisle, T_{t-1}^{cw} is the cold wall temperature, $R_{cw,cr}$ is the thermal resistance between the cold wall and the aisle, and δ_1 accounts for additional internal thermal loads.

A.2) Cold Wall Temperature: The thermal dynamics of the cold wall are described by a heat balance equation that accounts for convective interaction with the cold aisle air, heat exchange with the external environment, and internal wall conduction (Eq. 3).

$$T_t^{cw} = T_{t-1}^{cw} + \frac{1}{C_{cw}} \left(\frac{T_{t-1}^{amb} - T_{t-1}^{cw}}{R_{amb,cr}} - \frac{T_{t-1}^{cw} - T_{t-1}^{cr}}{R_{cw,cr}} \right) \cdot \tau \quad \text{Eq. (3)}$$

$$C_{cw} = \rho_{cw} \cdot A_{cw} \cdot d_{cw} \cdot c_{cw} \quad \text{Eq. (4)}$$

Where, T_t^{cw} represents the cold wall temperature at time step t , and τ is the time interval. C_{cw} is the thermal capacitance of the cold wall, calculated as the product of the wall material density (ρ_{cw}), wall area (A_{cw}), thickness (d_{cw}), and specific heat (c_{cw}). T_{t-1}^{amb} is the ambient temperature exchanging heat with the cold wall through the thermal resistance $R_{amb,cr}$ while T_{t-1}^{cr} is the cold aisle temperature exchanging heat with the wall through $R_{cw,cr}$.

A.3) Server-Rack Assembly Temperature: The variation of the internal air temperature within the rack is modeled using a thermal energy balance equation. This balance accounts for the heat extracted from the cold aisle by the server-rack assembly, the heat



transferred from the server-rack assembly to the hot aisle, the thermal exchanges from the CPU to the server-rack, and the heat generated by the server fans, as expressed in Eq. (5).

$$T_t^{server} = T_{t-1}^{server} + \frac{1}{C_{server}} \left(\dot{m}_{HVAC} \cdot c_{air} \cdot (T_{t-1}^{cr} - T_{t-1}^{server}) + \frac{T_{t-1}^{cpu} - T_{t-1}^{server}}{R_{eq}} + Q_{sfan} \right) \cdot \tau \quad \text{Eq. (5)}$$

$$C_{server} = n_{rack} \cdot n_{server} \cdot \rho_{air} \cdot V_{sf} \cdot c_{air} \quad \text{Eq. (6)}$$

$$R_{eq} = c_3 + \frac{c_4}{RPM^{c_5}} \quad \text{Eq. (7)}$$

$$Q_{sfan} = n_{rack} \cdot n_{server} \cdot (c_6 + c_7 \cdot RPM + c_8 \cdot RPM^2 + c_9 \cdot RPM^3) \quad \text{Eq. (8)}$$

This equation describes the evolution of the server temperature T_t^{server} over a time step τ . Where, C_{server} represents the thermal capacitance of the server assembly, calculated as the product of the number of racks n_{rack} , the number of servers per rack n_{server} , the air density ρ_{air} , the airflow volume per server V_{sf} , and the specific heat of air c_{air} . T_{t-1}^{cr} is the cold aisle temperature exchanging heat with the servers through the HVAC airflow, while T_{t-1}^{cpu} represents the CPU temperature transferring heat to the server assembly via the equivalent thermal resistance R_{eq} . The heat generated by the fans, Q_{sfan} , depends on the fan rotational speed (RPM), as detailed in Ahmed et al., 2021.

A.4) CPU-Server Assembly Temperature: The thermal model of the CPU-server assembly is based on energy balances that capture temperature dynamics through heat generation and conductive and convective transfers. It accounts for CPU dissipation, the effect of server fans, and thermal coupling with the surroundings, as expressed in Eq. (9).

$$T_t^{cpu} = T_{t-1}^{cpu} + \frac{1}{C_{cpu}} \left(Q_{cpu} + \frac{T_{t-1}^{cpu} - T_{t-1}^{server}}{R_{eq}} \right) \cdot \tau \quad \text{Eq. (9)}$$

$$C_{cpu} = m_{cpu} \cdot c_{cpu} + m_{rack} \cdot c_{rack} \quad \text{Eq. (10)}$$

$$Q_{cpu} = P_{s,idle} (c_0 + c_1 \cdot U_{cpu} + c_2 \cdot T_{in}) \quad \text{Eq. (11)}$$

The thermal dynamics of the CPU-server assembly are modeled through an energy balance that considers heat generation in the CPU, conductive transfer between the processor and the server via the equivalent thermal resistance R_{eq} and additional heat generated by the server fans, as detailed in Cuppelli et al., 2018. This approach captures the temporal evolution of CPU temperature in response to workload and environmental thermal conditions, providing a physically consistent representation of thermal coupling within the system.

The temperatures of the hot aisle and the hot wall are modeled analogously, following the same principles as equations (1) and (3).

B. HVAC System Power Consumption

The HVAC power is calculated based on the thermal balance of the air supplied to the data center and the system's energy efficiency, expressed through the Coefficient of Performance (COP). This model allows for an accurate estimation of the HVAC contribution to the total energy consumption of the data center.

$$P_t^{HVAC} = \frac{\dot{m}_{HVAC} \cdot c_{air}}{COP_{HVAC}} (T_{t-1}^{hr} - T_{t-1}^{in}) \quad \text{Eq. (12)}$$



This equation considers the air mass flow rate $\dot{m}_{HVAC}$, the specific heat of air c_{air} , and the temperature difference between the hot aisle air temperature T_{t-1}^{hr} and the inlet air temperature T_{t-1}^{in} , allowing a dynamic estimation of HVAC energy consumption according to the actual thermal load. In this way, the model provides a precise representation of the HVAC contribution to the total data center energy consumption.

C. Objective Function

The proposed model is formulated as an optimization problem to simultaneously represent the thermal and energy dynamics of the data center, as well as the associated operational decisions. The objective function aims to minimize the electricity consumption of the HVAC system, thereby reducing the total energy costs.

$$\text{Min } z = \sum_{t \in \mathcal{T}} \left[\left(u_t^{HVAC} \cdot \frac{\dot{m}_{HVAC} \cdot c_{air}}{COP_{HVAC}} (T_{t-1}^{hr} - T_{t-1}^{in}) + P_{IT} - P_{PV} \right) \cdot \tau \cdot C_e \right] \quad \text{Eq. (13)}$$

The simulation horizon covers a full day, divided into time intervals of $\tau = 15$ minutes, resulting in a set of $N = 96$ discrete steps, where $t \in \{1, 2, 3, \dots, N\}$. Equations (1), (3), (5) and (9) represent the dynamic system constraints, modeling the energy balances and thermal dynamics of the components. These equations ensure a physically consistent evolution of the system while respecting operational and safety limits. Here, C_e represents the electricity tariff.

D. Additional Constraints

The proposed model includes additional constraints to ensure the stable operation of the data center, including operational temperature limits and power constraints for the IT room and the HVAC system:

$$T_{min}^{cr} \leq T_t^{cr} \leq T_{max}^{cr} \quad \forall t \in \mathcal{T} \quad \text{Eq. (14)}$$

$$T_{min}^{cw} \leq T_t^{cw} \leq T_{max}^{cw} \quad \forall t \in \mathcal{T} \quad \text{Eq. (15)}$$

$$T_{min}^{server} \leq T_t^{server} \leq T_{max}^{server} \quad \forall t \in \mathcal{T} \quad \text{Eq. (16)}$$

$$T_{min}^{cpu} \leq T_t^{cpu} \leq T_{max}^{cpu} \quad \forall t \in \mathcal{T} \quad \text{Eq. (17)}$$

$$T_{min}^{hr} \leq T_t^{hr} \leq T_{max}^{hr} \quad \forall t \in \mathcal{T} \quad \text{Eq. (18)}$$

$$T_{min}^{hw} \leq T_t^{hw} \leq T_{max}^{hw} \quad \forall t \in \mathcal{T} \quad \text{Eq. (19)}$$

E. Electrical Power Constraints

The proposed model ensures power balance and operational limits for both the IT and HVAC systems

$$P_t^{grid} + P_t^{PV} = P_t^{IT} + P_t^{HVAC} \quad \text{Eq. (20)}$$

This equation guarantees that the total power supplied by the grid (P_t^{grid}) and the photovoltaic system (P_t^{PV}) matches the combined consumption of the IT equipment and the HVAC system at each time step.

$$0 \leq P_t^{HVAC} \leq u_t^{HVAC} \cdot P_{max}^{HVAC} \quad \forall t \in \mathcal{T} \quad \text{Eq. (21)}$$

$$u_t^{HVAC} \in \{0, 1\} \quad \forall t \in \mathcal{T} \quad \text{Eq. (22)}$$

These constraints define the operational limits of the HVAC system, where u_t^{HVAC} is a binary control variable representing the ON/OFF state, and P_{max}^{HVAC} is the maximum HVAC power capacity.



The IT power consumption was calculated using the equation (23).

$$P_t^{IT} = n_{rack} \cdot n_{server} \cdot P_{s,idle} (c_0 + c_1 \cdot U_{cpu} + c_2 \cdot T_{in}) \quad \text{Eq. (23)}$$

Where, n_{rack} , is the number of racks, n_{server} is the number of servers per rack, and $P_{s,idle}$ is the power consumption of a single server in idle state. The terms U_{cpu} and T_{in} represent the CPU utilization ratio and the inlet air temperature to the server, respectively, while c_0, c_1, c_2 are empirical coefficients capturing the influence of baseline power, CPU load, and inlet temperature on server power consumption.

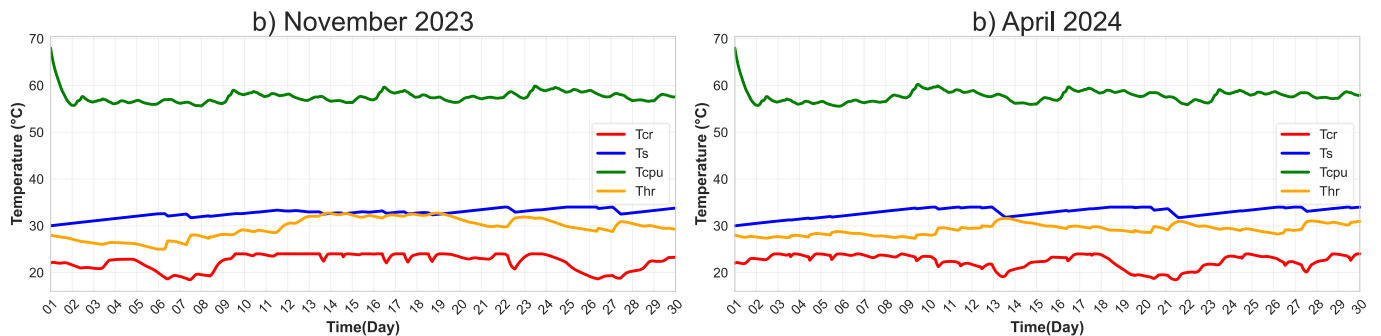
3 RESULTS AND DISCUSSION

The dynamic thermal model was formulated as an optimization problem in Pyomo using the BONMIN solver to minimize HVAC electricity consumption and reduce total energy costs. The model simulated the temporal evolution of temperatures and power in critical subsystems (cold aisle, cold wall, servers/racks, CPUs, hot aisle, hot wall) with a 15-minute time step. Key parameters included thermal capacities (12.3 MJ/°C and 19.7 MJ/°C for cold and hot aisles), wall areas, convection coefficients, CPU and rack properties, HVAC efficiency (COP = 3.7), and airflows (0.05 to 2 kg/s). Time-dependent variables encompassed ambient temperature, supply air temperature, CPU utilization, photovoltaic generation, and electricity costs. Simulations used historical weather data from November 2023 and April 2024 in Santo André to evaluate solar-data center interactions and seasonal effects on energy consumption. The model achieved optimal solutions of 129.58 USD in November and 95.16 USD in April, demonstrating how thermal dynamics directly impact HVAC performance and energy costs.

A. Thermal Analysis

The thermal behavior of the data center was evaluated by monitoring temperatures in key subsystems. Cold aisle temperatures ($T_{cr} = 22.3 - 22.5$ °C) and supply air temperatures ($T_{server} = 32.8 - 33$ °C) remained stable and within operational ranges, with low dispersion (standard deviations < 2 °C). CPU temperatures ($T_{cpu} = 57.6 - 57.8$ °C) were effectively controlled with a maximum point of 68 °C, while hot aisle temperatures ($T_{server} = 29 - 29.6$ °C) exhibited greater seasonal variability. The comparison between November and April revealed relevant seasonal differences: November presented slightly higher temperature averages and peaks due to warmer ambient conditions, while April showed lower and more homogeneous temperatures, indicating more stable thermal behavior characteristic of the colder period.

Figure 1 – Temporal Evolution of Temperatures

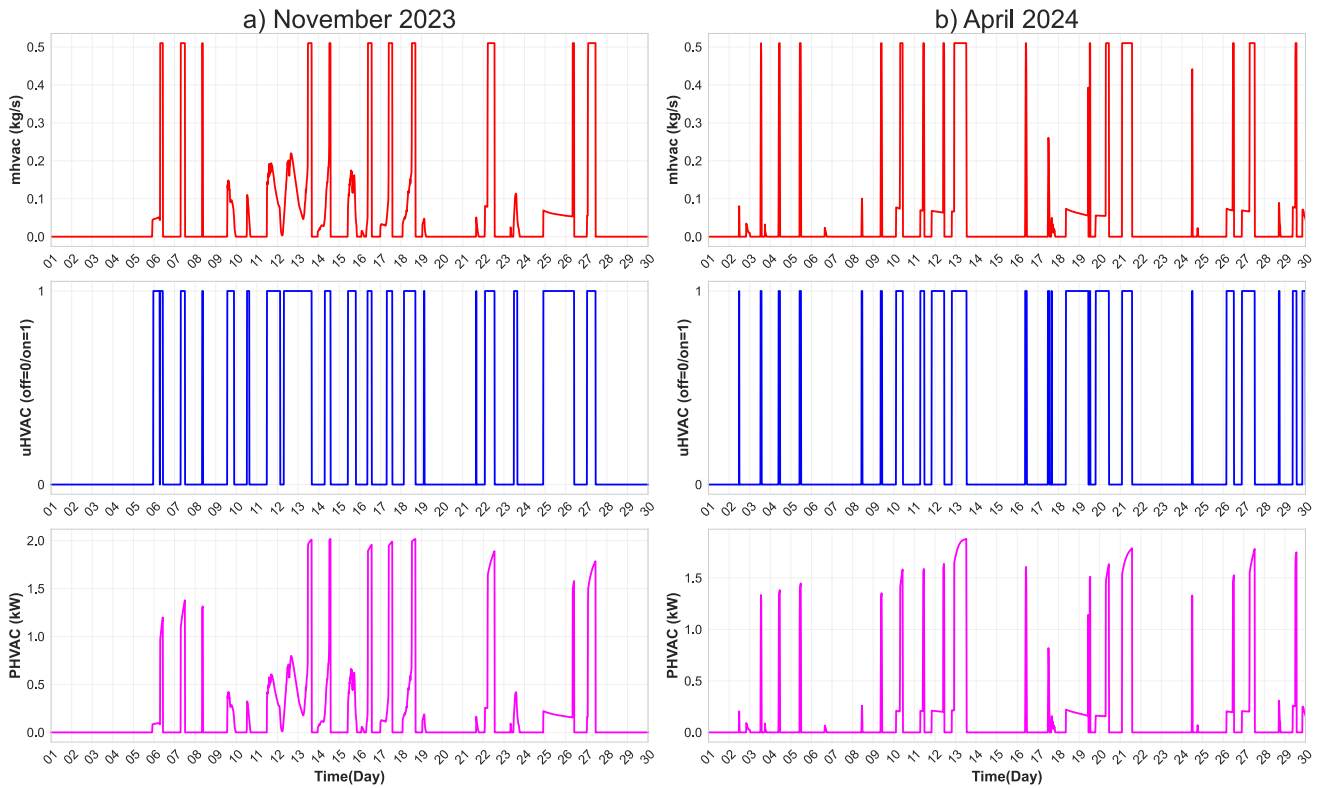




B. HVAC Operation

The HVAC operation analysis shows how cooling demand responds to seasonal thermal variations. Figure 2 presents the operation results for November and April. In November, the fraction of time the HVAC was active (u_{HVAC}) reached approximately 25.1 %, with an average mass flow $\dot{m}_{HVAC} = 0.054 \text{ kg/s}$ and mean power consumption $P_{HVAC} = 0.184 \text{ kW}$. Peaks reached up to 2 kW, reflecting intermittent and more intensive operation during the warmer period. In contrast, April showed lower values and a more uniform activation ($u_{HVAC} = 23.2 \%$, $\dot{m}_{HVAC} = 0.049 \text{ kg/s}$, $P_{HVAC} = 0.156 \text{ kW}$), evidencing lower cooling demand. These results demonstrate how the HVAC system dynamically adapts to thermal conditions and internal load, optimizing electricity consumption.

Figure 2 – HVAC System Operation

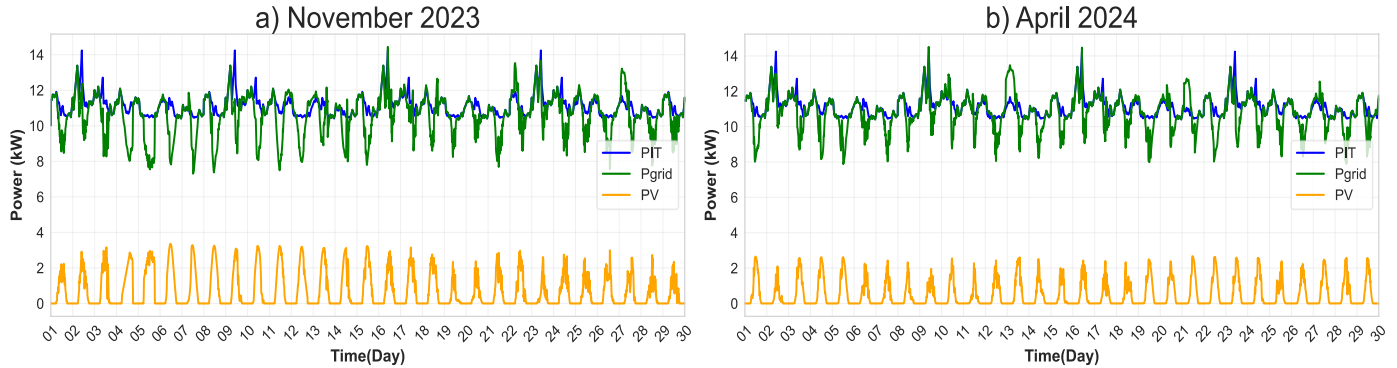


C. Consumption Analysis

Figure 3 provides a detailed view of each component's contribution to total data center consumption. Server demand remained practically constant at around 11.2 kW in both months. In November, the HVAC average consumption was slightly higher (0.18 kW) than in April (0.16 kW), consistent with the higher cooling needs during warmer months. Photovoltaic generation was also higher in November (0.72 kW) than in April (0.50 kW), partially reducing dependence on the grid. Although solar generation was not sufficient to cover the entire data center load, it was enough to offset the HVAC consumption, underscoring the relevance of photovoltaic integration as a complementary energy source, particularly during periods of high demand.



Figure 3 – HVAC Operation and Electrical Consumption

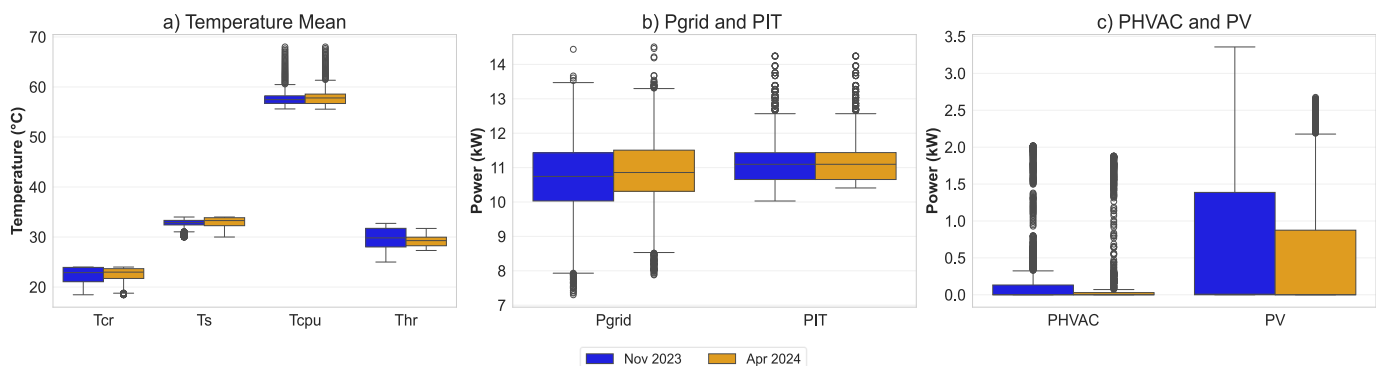


D. Discussion

The results indicate that warmer months incur higher operating costs, with optimal solutions of 129.58 USD in November and 95.16 USD in April. Thermal analysis shows that the system remains stable, with practically constant average temperatures (35.6 °C) and lower variability in April, indicating more uniform thermal control. Total internal server demand (P_{IT}) remained stable, while grid dependence increased slightly by 1.8 %, reflecting marginally higher electricity use. HVAC consumption decreased by 14.8 % in April compared to November, although photovoltaic generation also fell by 30.7 %, reducing solar coverage of HVAC from 100 % to 100 %, still maintaining a surplus of 0.34 kW.

Figure 4 presents a comparative boxplot of temperature variables, HVAC consumption, and power (P_{HVAC} , P_{grid} , P_{PV}), visually illustrating seasonality and data dispersion. These findings emphasize the importance of storage strategies and surplus solar management to maximize photovoltaic integration and minimize grid reliance, ensuring efficient and sustainable operation.

Figure 4 – Seasonal Comparison of Data Center Performance



4 CONCLUSIONS

In conclusion, the study demonstrates that integrating photovoltaic systems and implementing an advanced thermal management model can effectively optimize the operation of urban data centers. The model achieved optimal solutions of 129.58 USD in November and 95.16 USD in April, highlighting the impact of seasonal variation on operating costs. Cold aisle and supply



INTERNATIONAL SYMPOSIUM ON ENERGY

Optimizing renewable energy: advanced strategies for enhanced efficiency and sustainability

air temperatures remained stable within operational ranges, while CPU temperatures were effectively controlled, with a maximum of 72 °C. HVAC operation showed an average consumption of 0.184 kW in November and 0.156 kW in April. In November and April, photovoltaic generation covered up to 100% of HVAC demand, reducing grid dependency. These results show that the proposed strategy optimizes energy efficiency and solar utilization, providing clear guidelines for sustainable urban data center operations.

ACKNOWLEDGMENTS

The authors would like to thank the Federal University of ABC (UFABC) for institutional support. This work was financially supported by the São Paulo Research Foundation (FAPESP), under grants \#2025/02554-0 and \#2024/14894-8; by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Funding Code 001, under grants 88887.709682/2022-00, 88887.692379/2022-00, and 88887.939062/2024-00; and by the National Council for Scientific and Technological Development (CNPq), through the Universal Project no. 407244/2023-9 and 309013/2025-9.

REFERENCES

- AHMED, K. M. U.; ALVAREZ, M.; BOLLEN, M. H. J. A novel reliability index to assess computational resource adequacy in data centers. *IEEE Access*, v. 9, p. 54530-54541, 2021. DOI: <https://doi.org/10.1109/ACCESS.2021.3070915>
- CUPELLI, L.; SCHÜTZ, T.; JAHANGIRI, P.; FUCHS, M.; MONTI, A.; MÜLLER, D. Data center control strategy for participation in demand response programs. *IEEE Transactions on Industrial Informatics*, v. 14, n. 11, p. 5087-5099, 2018. DOI: <https://doi.org/10.1109/TII.2018.2806889>
- MYTTON, D.; ASHTINE, M. Sources of data center energy estimates: A comprehensive review. *Joule*, v. 6, n. 9, p. 2032-2056, 2022. DOI: <https://doi.org/10.1016/j.joule.2022.07.011>
- LIN, Weiwei; WU, Wentai; WANG, Haoyu; WANG, James Z.; HSU, Ching-Hsien. Experimental and quantitative analysis of server power model for cloud data centers. *Future Generation Computer Systems*, v. 86, p. 940–950, 2018. Disponível em: <https://doi.org/10.1016/j.future.2016.11.034>. Acesso em: 29 ago. 2025.
- OFFUTT, Martin C.; ZHU, Ling. Data centers and their energy consumption: frequently asked questions. Washington, D.C.: Congressional Research Service, 2025. (CRS Report R48646). Disponível em: <https://crsreports.congress.gov>. Acesso em: 29 ago. 2025.
- SHAH, Dhruv; VORA, Hitesh. Green Data Centers: Merging IT Innovation with Energy-Efficient Cooling and Renewable Power. *International Journal of Mechanical Engineering and Technology (IJMET)*, v. 16, n. 3, p. 132–160, maio/jun. 2025. DOI: https://doi.org/10.34218/IJMET_16_03_007
- ZHANG, Wei; LI, Ming; GARCIA, José; et al. Stochastic optimization of photovoltaic-integrated data centers with hybrid free-mechanical cooling and waste heat recovery. *Renewable & Sustainable Energy Reviews*, Londres, v. 200, n. 3, p. 123–141, jul. 2025. DOI: 10.1016/j.rser.2025.137912.