

**Carbon-Aware AI Computing: Optimizing Data Center Power Consumption Using
Workload, Environmental, and Grid Emissions Data**

EMS 2026
Doctoral Consortium Submission

By

Dana McDonald

Temple University, Fox School of Business

12 April 2026

Data centers are becoming one of the largest and least visible consumers of energy worldwide, consuming over 240 terawatt-hours (TWh) annually, or approximately 1% of global electricity demand. The growth of the AI-enabled digital economy is accelerating this energy demand and its associated carbon impact. With rapidly increasing GPU-accelerated workloads hitting data centers, data center operators will soon be tasked not only with managing energy consumption but also with doing so intelligently, while considering application performance, cost, and environmental impact. Current methods of managing these optimizations largely silo workloads, external environmental conditions, and energy/resource sourcing. This research proposes a carbon-aware AI computing framework that leverages workload telemetry, along with real-time environmental conditions and grid carbon intensity to inform intelligent, sustainable decision-making in high-performance (HPC) data centers.

This research combines information theory and sustainable computing to explain how data centers can improve energy and carbon efficiency. It reframes energy optimization as an information problem, where workload and environmental data act as meaningful signals that reduce uncertainty and enable more intelligent system-level decisions. When these signals exhibit structure rather than randomness, they reduce entropy and improve the ability to predict energy consumption. Machine learning models then act as decoders, transforming these signals into actionable insights for forecasting and optimization.

The research question that guides this study is:

RQ1: How can workload characteristics, environmental conditions, and grid carbon intensity be integrated to optimize power consumption and reduce the carbon footprint of AI workloads in HPC data centers?

Prior literature shows that data center energy demand is rapidly increasing, driven by rack power density rising from 2.4 kW to over 8.4 kW, intensifying pressure on infrastructure (Isazadeh et al., 2023). Data center operations are estimated to contribute 2-4% of global carbon emissions (Zhu et al., 2023). Cooling systems account for 30-40% of total energy use (Ni & Bai, 2017), and environmental conditions strongly influence their efficiency. Traditional metrics like Power Usage Effectiveness (PUE) fail to capture workload variability and carbon impact, suggesting the need for more integrated approaches (Safari et al., 2025). Machine learning has demonstrated strong potential in energy optimization. AI-driven systems have reduced cooling energy by up to 40% (Ma et al., 2025) and improved overall efficiency by approximately 15% in large-scale data centers (Kumar et al., 2021). Workload characterization research shows that GPU-intensive workloads produce predictable power consumption patterns, enabling predictive modeling opportunities (Veigas et al., 2025).

This research study employs a quantitative, empirical research design using three real-world datasets. The MIT Supercloud dataset provides visibility into workload behavior and energy consumption at the job level. It includes over 500,000 anonymized jobs, with approximately 100,000 GPU-intensive workloads, with telemetry such as CPU/GPU utilization and power draw (S. Samsi et al., 2021). Historical weather data from Weather Underground provides environmental variables, including temperature, humidity, wind speed, and pressure (The Weather Company, n.d.). Grid emissions data from WattTime (WattTime API, n.d.) provide marginal and average carbon intensity metrics, enabling measurement of the real-time carbon impact of electricity consumption.

The research study proceeds in three phases. The first phase combines data integration with statistical and econometric analysis. Workload, environmental, and emissions time-series

data are synchronized at the job level, followed by statistical analysis using STATA. These analyses establish baseline relationships and test the moderating effects of environmental variables. The second phase is predictive modeling. Standard machine learning models are trained and tested using workload characteristics only, workload characteristics plus environmental variables, and fully integrated models with emissions data. The third phase is optimization and simulation of carbon-aware scheduling policies. These simulations assess trade-offs among energy consumption, carbon emissions, and system performance metrics such as latency and throughput.

This research contributes to theory by extending information theory into the domain of energy optimization in HPC data centers. It shows how workload characteristics, environmental conditions, and grid emissions data serve as informational signals that reduce uncertainty and improve the accuracy of energy-use predictions. This research advances sustainable computing through an integrated approach that connects energy efficiency with carbon-aware decision-making.

This research offers practical insights for data center operators. It demonstrates how predictive models can better anticipate energy consumption and manage costs more effectively. It shows how carbon-aware scheduling can reduce emissions by shifting workloads to times when cleaner energy is available. The framework also helps improve resource allocation, reduce peak demand, and support broader sustainability goals. It tackles the challenge of managing energy and carbon impact in increasingly complex, AI-driven environments. By integrating workload, environmental, and grid emissions data, it offers a more adaptive and intelligent way to run data centers not only more efficiently, but also more responsibly.

REFERENCES

- Isazadeh, A., Ziviani, D., & Claridge, D. E. (2023). Global trends, performance metrics, and energy reduction measures in datacom facilities. *Renewable and Sustainable Energy Reviews*, 174, 113149. <https://doi.org/10.1016/j.rser.2023.113149>
- Kumar, R., Khatri, S. K., & Diván, M. J. (2021). Power usage efficiency (PUE) optimization with counterpointing machine learning techniques for data center temperatures. *International Journal of Mathematical, Engineering and Management Sciences*, 6(6), 1594.
- Ma, S., Liu, Y., Liu, Y., Wang, J., Fang, Q., & Huang, Y. (2025). Artificial intelligence-enabled predictive energy saving planning of liquid cooling system for data centers. *Advanced Engineering Informatics*, 65, 103283. <https://doi.org/10.1016/j.aei.2025.103283>
- Ni, J., & Bai, X. (2017). A review of air conditioning energy performance in data centers. *Renewable and Sustainable Energy Reviews*, 67, 625–640. <https://doi.org/10.1016/j.rser.2016.09.050>
- S. Samsi, M. L. Weiss, D. Bestor, B. Li, M. Jones, A. Reuther, D. Edelman, W. Arcand, C. Byun, J. Holodnack, M. Hubbell, J. Kepner, A. Klein, J. McDonald, A. Michaleas, P. Michaleas, L. Milechin, J. Mullen, C. Yee, ... V. Gadepally. (2021). The MIT Supercloud Dataset. *2021 IEEE High Performance Extreme Computing Conference (HPEC)*, 1–8. <https://doi.org/10.1109/HPEC49654.2021.9622850>
- Safari, A., Sorouri, H., Rahimi, A., & Oshnoei, A. (2025). A Systematic Review of Energy Efficiency Metrics for Optimizing Cloud Data Center Operations and Management. *Electronics*, 14(11). <https://doi.org/10.3390/electronics14112214>

The Weather Company. (n.d.). *Weather Underground*. Historical Weather Data. Retrieved <https://www.wunderground.com/history>

Veigas, K. L., Chinnici, A., De Chiara, D., & Chinnici, M. (2025). Towards Energy Efficiency of HPC Data Centers: A Data-Driven Analytical Visualization Dashboard Prototype Approach. *Electronics*, 14(16), 3170. <https://doi.org/10.3390/electronics14163170>

WattTime API. (n.d.). *CO2 Marginal Operating Emissions Rate* [Dataset]. Retrieved March 18, 2026, from api.watttime.org

Zhu, H., Zhang, D., Goh, H. H., Wang, S., Ahmad, T., Mao, D., Liu, T., Zhao, H., & Wu, T. (2023). Future data center energy-conservation and emission-reduction technologies in the context of smart and low-carbon city construction. *Sustainable Cities and Society*, 89, 104322. <https://doi.org/10.1016/j.scs.2022.104322>