Proceedings of CONF-MPCS 2025 Symposium: Mastering Optimization: Strategies for Maximum Efficiency
DOI: 10.54254/2753-8818/2025.CH24595

Modeling the Environmental Footprint and Sustainability of High-Performance Computing

Weihang Ma[1*], Lanxi Gao[1], Qinyue Gong[1], Zhixing Wang[1]

[1]The High School Affiliated to Renmin University of China, Beijing, China
**Corresponding Author. Email: coringree810@gmail.com*

Abstract. The rapid development and expansion of High-Performance Computing (HPC) systems present signif- icant environmental challenges, primarily due to substantial energy consumption and associated carbon emissions, often from non-renewable sources. This paper provides a comprehensive analysis of the en- vironmental footprint of HPC. It models global energy consumption, considering both full capacity and average utilization rates (estimated at 12-18%). A model is developed to quantify total carbon emis- sions, accounting for diverse energy sources, regional energy mixes, and conversion efficiencies. The analysis explores future trends by considering projected HPC growth, increasing energy demand from other sectors, and potential shifts in energy sources, forecasting impacts up to 2030 using time series analysis. The study further investigates the potential for mitigation by modeling the relationship be- tween increased renewable energy adoption and carbon emission reductions, including a scenario for 100% renewables. Additionally, the model is expanded to include water usage, another critical environmental factor, analyzing its relationship with energy consumption. Based on these models, actionable technical and policy recommendations are proposed to enhance energy efficiency and promote sustainability in the HPC sector, emphasizing the need to integrate these concerns into future development and climate change mitigation strategies.

Keywords: High-Performance Computing, energy consumption, carbon emissions, environmental impact, renewable energy, data centers

1. Introduction

In recent years, the expansion of high-powered computing (HPC) capabilities has transformed numerous sectors, including artificial intelligence (AI), data science, and cryptocurrency mining. These areas rely heavily on massive data processing and computational power, driving demand for high-performance hardware and extensive data center infrastructures. However, this rapid advancement is not without environmental cost [1, 2].

The primary issue with HPC is energy consumption, as these systems rely on large-scale data centers that require vast amounts of electricity, often sourced from non-renewable energy. This results in significant carbon emissions. To fully understand HPC's environmental footprint, it is crucial to estimate annual energy use, considering both full and average utilization rates. Developing

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.CH24595

a model that quantifies carbon emissions based on energy consumption, while accounting for energy sources and regional energy mixes, can offer valuable insights into current and future impacts, especially given the expected growth in demand by 2030 [3, 4].

Beyond energy use, the manufacturing and disposal of HPC hardware also contribute to environmental harm, generating significant electronic waste (e-waste) and depleting natural resources due to rare earth extraction. In addition, the large physical footprint of data centers, along with concerns about air quality and noise pollution, can negatively affect nearby communities and ecosystems [1]. As HPC continues to expand, addressing these complex challenges is essential, and strategies must be developed to balance technological progress with environmental sustainability.

This paper aims to：

Describe the scope of current global energy consumption associated with HPC, considering full capacity and average utilization.

Develop a model to assess the environmental impact, focusing on total carbon emissions based on energy sources, efficiency, and regional mixes.

Apply the model to analyze future scenarios (up to 2030) considering HPC growth, energy demand changes, and shifts in energy sources.

Expand the model to assess the impact of increased renewable energy adoption and to incorporate water usage as another key environmental factor.

Provide actionable recommendations for reducing the environmental impact of HPC through technical and policy solutions.

2. Methodology and models

This section outlines the models developed to estimate energy consumption, carbon emissions, future trends, and the impact of mitigation strategies for HPC systems.

2.1. Estimating global HPC energy consumption

To estimate the total global power consumption of HPC, energy consumption data from various data center types (traditional, cloud non-hyperscale, hyperscale) were collected for the years 2015-2021 [4]. The annual total energy consumption is assumed to be the sum of these types, as shown in Table 1.

Table 1: Total energy consumption of data centers worldwide, under real life utilization rate conditions

Year	Total Energy Consumption Worldwide under utilization rate (terawatt hours)
2015	190.7
2016	195.26
2017	195.03
2018	197.69
2019	191.84
2020	190.13
2021	190.81

Data centers typically do not operate at full capacity. Based on industry reports, the average server utilization rate is estimated to be between 12-18% [2]. For simplification, a median utilization

DOI: 10.54254/2753-8818/2025.CH24595

rate (R) of 15% is assumed. The theoretical energy consumption under full capacity ($E_{fullcapacity}$) is calculated from the energy consumption under average utilization ($E_{utilizationrate}$) using equation (1):

Applying this calculation yields the estimated energy consumption under full capacity conditions shown in Table 2, indicating relative stability over this period.

$$E_{fullcapacity} = \frac{E_{utilizationrate}}{R} \quad (1)$$

2.2. Modeling carbon emissions from HPC energy consumption

A model was developed to assess the environmental impact, specifically total carbon emissions, resulting from HPC energy consumption. This requires considering the energy mix and the carbon intensity of different energy sources.

Table 2: The total energy consumption of data centers worldwide, under full capacity conditions

Year	Total Energy Consumption Worldwide under full capacity (terawatt hours)
2015	1271.33
2016	1301.73
2017	1300.2
2018	1317.93
2019	1278.93
2020	1267.53
2021	1272.07

baseline scenarios (e.g., by IPCC). While future changes may alter trends, this provides a baseline impact assessment. Data splitting into training/validation sets is used to mitigate overfitting.

Variables: Key variables include Carbon Intensity (direct, CI_{direct}; LCA, CI_{LCA}), Low Carbon Percent- age (LCP), and Renewable Percentage (RP) [3], detailed in Table 3.

Table 3: Variable definitions for carbon emissions

Symbol	Definition	Unit
CI_{direct}	Carbon Intensity (direct)	gCO2eq/kWh
CI_{LCA}	Carbon Intensity (LCA)	gCO2eq/kWh
LCP	Low Carbon Percentage	Percentage (%)
RP	Renewable Percentage	Percentage (%)

The energy mix and corresponding carbon emissions per kWh for major fossil fuels based on U.S. data [5] inform the model. Using the shares ($percent_i$) and carbon emission factors ($carbon\ emission_i$) for each source i, the average carbon emission per unit of electricity consumption ($carbon\ emission_{average}$) is calculated using a weighted average (Equation (2)):

$$\text{carbon emission}_{average} = \Sigma\left(percent_i \text{ carbon emission}_i\right) \quad (2)$$

DOI: 10.54254/2753-8818/2025.CH24595

Based on sample U.S. data (53% coal @ 2.30 lb/kWh, 45% natural gas @ 0.97 lb/kWh, 2% petroleum @ 2.38 lb/kWh), the average carbon emission is calculated as 1.70 pounds per kWh. Total carbon emissions are then derived by multiplying this average intensity by the total energy consumption estimated in Section 2.1.

2.3. Forecasting future energy demand and environmental impact (to 2030)

To project future impacts, the model considers: (a) the growth of HPC facilities, (b) increasing energy demand from other major sectors, and (c) potential changes in the energy source mix.

Growth of HPC: Data on the number of global data centers from 2010-2023 was analyzed using polynomial regression (Degree 3) to model growth trends. The analysis indicated rapid growth followed by a slowdown after 2019.

Increasing demand from other sectors: Energy consumption data for residential, commercial, industrial, and transportation sectors were analyzed. Generalized linear models (polynomial regression) were developed to study the energy demand trends for the major consuming sectors (residential and commercial), showing increasing demand over time.

Actual energy sources and mixes: The complex mix of energy sources influences the carbon emission per unit of energy. A Long Short-Term Memory (LSTM) model was developed using historical data (1970- 2023) to analyze and forecast the temporal change of carbon emission intensity (kg CO_2eq/kWh). Model performance analysis (RMSE) and forecasts for the US and World were generated. The models generally captured historical downward trends but showed some divergence in recent years and future projections, suggesting limitations in extrapolating current trends without accounting for potential systemic shifts.

These model components provide inputs for estimating the upper limit of total carbon emissions from HPC by 2030 under various scenarios.

2.4. Modeling mitigation strategies: renewable energy and water usage

This section expands the analysis to quantify the impact of mitigation strategies, specifically increasing renewable energy share and considering water consumption.

2.4.1. Impact of renewable energy on carbon emissions

Assumptions:

The relationship between renewable energy percentage (P_{RE}) and carbon intensity (CI_{direct}, CI_{LCA}) observed in 2023 data [3,6] is causally interpretable and can be extrapolated.

A 100% renewable scenario can be simulated via extrapolation, with qualitative consideration of grid/s- torage challenges.

The P_{RE}-Carbon Intensity relationship remains relatively stable in the short term.

Justification: Daily variations in the dataset allow for causal inference analysis. Exponential decay models can account for diminishing returns as P_{RE} approaches 100%. Sensitivity analysis and BAU comparisons address uncertainties.

Variables: In addition to variables in Table 3, Total Energy Production (E_{total}) and Reduced Emissions ($E_{reduced}$) are used [3], defined in Table 4.

DOI: 10.54254/2753-8818/2025.CH24595

Table 4: Variable definition (subset for renewable impact)

Symbol	Definition	Unit
CI_{direct}	Carbon Intensity (direct)	gCO2eq/kWh
$P_{renewable}$	Proportion of renewable energy in total energy production	Percentage (%)
E_{total}	Total energy generated in the U.S. on day	kWh
$E_{reduced}$	Reduction in carbon emissions from increasing renewables	kg CO2eq/day

Model Development: Baseline daily emissions ($E_{baseline}$) are calculated as:

$$E_{baseline}(t) = CI_{direct}(t) \cdot E_{total}(t) \tag{3}$$

Emissions under increased renewable share ($P_{renewable}$) are modeled assuming negligible direct carbon intensity from renewables:

$$E_{reduced}(t) = (1 - P_{renevable}(t)) \cdot CI_{divect}(t) \cdot E_{total}(t) \tag{4}$$

To simulate a gradual transition and diminishing returns for 100% renewables:

$$E_{reduced}(t) = E_{baseline}(t) \cdot \ e^{-\alpha \cdot P_{renewable}(t)} \tag{5}$$

where α reflects the efficiency of renewable integration.

A linear regression model validates the relationship using 2023 U.S. daily data:

$$E_{actual}(t) = \beta_0 + \beta_1 \cdot CI_{direct}(t) + \beta_2 \cdot CI_{LCA}(t) + \beta_4 \cdot P_{low}(t) + \epsilon_t \tag{6}$$

where P_{low} is the low-carbon energy percentage and ϵ_t is the error term.

2.4.2. Incorporating water usage

Assumptions:

Statistical trends between energy consumption and water withdrawals (2016-2023 data [7, 8]) continue (BAU).

Water withdrawals (W) are directly proportional to energy consumption (E) via an average intensity factor (α).

Baseline assumes equal water intensity across sources; adjustments made for renewables.

HPC energy (EHP C) and water use (WHP C) scale proportionally to their share of total energy.

Justification: Simplifies projections, consistent with IPCC practices and empirical evidence of linear energy- water links. Allows focus on HPC impact.

Variables: See Table 5.

DOI: 10.54254/2753-8818/2025.CH24595

Table 5: Variable definition (water usage)

Symbol	Definition	Unit
E	Total energy consumption	Gigawatt-hours (GWh)
W	Total water withdrawals	Million gallons (Mgal)
α	Water withdrawal intensity per unit of energy	Mgal/GWh
αi	Water withdrawal intensity of source i	Mgal/GWh
EHPC	HPC system energy consumption	Gigawatt-hours (GWh)
WHPC	HPC system water withdrawals	Million gallons (Mgal)

Model Development: Baseline relationship between total energy (E) and water (W):

$$W = \alpha \cdot E + \epsilon \tag{7}$$

where α is estimated using regression on historical data (2016-2023), and ϵ is the residual error.
HPC-specific water demand:

$$W_{HPC} = \alpha \cdot E_{HPC} \tag{8}$$

Impact of renewables on water intensity: The overall intensity factor α changes based on the proportion (P_i) and water intensity (α_i) of each energy source i:

$$\alpha_{new} = \Sigma P_i \cdot \alpha_i \tag{9}$$

This α_{new} is then used in equation (8) to calculate $W_{HP\ C}$ under different renewable penetration scenarios (e.g., baseline, 50% renewable, 100% renewable).

3. Results and interpretation

3.1. Renewable energy impact on carbon emissions

Analysis of 2023 U.S. daily data[3] established a strong linear relationship between the renewable energy percentage and carbon intensity, with parameters shown in Table 6.
The linear regression equation (keeping two decimals) is:

$$y = -6.93x + 489.73 \tag{10}$$

where y is carbon intensity and x is the renewable energy percentage. This negative slope indicates that as the percentage of renewable energy increases, carbon emissions decline linearly. Simulations using this model quantify the potential reduction in carbon emissions as renewable energy penetration increases.

DOI: 10.54254/2753-8818/2025.CH24595

Table 6: Regression model parameters (renewable % vs carbon intensity)

Parameter	Value
Slope	-6.927
Intercept	489.727
R-squared	0.894

3.2. Energy consumption and water usage

Analysis of Google data center annual water usage (Mgal) and energy consumption (GWh) from 2016 to 2023 [7, 8] revealed a strong linear correlation, with parameters detailed in Table 7.

Table 7: Regression model parameters (energy consumption vs water usage)

Parameter	Value
Slope	0.30715634
Intercept	772.68644612
R-squared	0.9871740908384165

The relationship (keeping two decimals) is:

$$y = 0.31x + 772.69 \tag{11}$$

where y is water withdrawal (Mgal) and x is energy consumption (GWh). This positive linear relationship indicates that as the energy consumption increases, the amount of water withdrawal also increases linearly.

Interpretation:

Baseline: HPC systems contribute significantly to water withdrawals (WHP C,baseline) under the current energy mix, calculated using equation (8) and the derived baseline α.

Renewable Impact: Increasing renewable energy (Prenew) generally decreases water withdrawal inten- sity (α), reducing WHP C. However, the extent depends on the type of renewables used (e.g., solar/wind are low-water, hydropower can be high-water).

Challenges: Transitioning fully to renewables might reduce overall water demand, but heavy reliance on sources like hydropower could offset gains compared to low-water renewables[9,10].

3.3. Strengths and weaknesses of the models

Strengths:

Data-driven: Utilizes real-world data from authoritative sources (EIA, IEA, Electricity Maps, IBM, Google).

Statistically rigorous: Employs established techniques (regression, time series, causal inference tools) providing a robust framework.

Actionable: Informs specific technical and policy recommendations.

Unified framework: Integrates carbon and water footprints for a holistic assessment.

Weaknesses:

DOI: 10.54254/2753-8818/2025.CH24595

Modeling complexity: Precise modeling of challenges like energy intermittency, storage, and infrastruc- ture limitations is beyond the scope, potentially affecting long-term 100% renewable scenario accuracy.

Data dependency: Accuracy relies on the availability and quality of historical data; gaps or errors could undermine results.

Assumption limitations: Reliance on BAU assumptions and stable relationships might not capture disruptive technological or policy changes.

4. Recommendations and conclusion

4.1. Actionable recommendations

Reducing the environmental impact of HPC requires a combination of technological innovation and policy measures.

Technical Solutions:

Hardware Optimization: Promote energy-efficient processors, advanced liquid cooling systems, and hardware designed for renewable/alternative energy integration.

Software Optimization: Implement dynamic power management, energy-aware scheduling algorithms, and virtualization to improve resource utilization and reduce computational waste.

AI/ML Integration: Utilize AI/ML for predictive maintenance (reducing downtime/waste), optimizing resource allocation, real-time energy monitoring, and streamlining workloads.

Policy Interventions:

Incentivize Renewables: Offer subsidies, tax benefits, or renewable energy credits for HPC operations transitioning to clean energy.

Regulations: Implement carbon emission caps or standards for large data centers.

Green Finance: Develop green credit programs with favorable rates for financing sustainable HPC initiatives.

Table 8 summarizes these recommendations.

Table 8: Actionable recommendations

Category	Optimization/Management	Technology/Strategy	Policy Support
Hardware	Energy-saving processors	Liquid cooling systems	Incentives for efficient hardware adoption
Software	Dynamic power management	Energy-aware scheduling	Standards for software energy efficiency
Energy	Optimize resource usage	Renewable/alternative energy integration	Subsidies, credits for renewable transition
AI/ML	Predictive hardware failures	Optimize resource usage	Funding for AI in green computing research
Cross-cutting	Identify efficiency trends	Streamline workloads	Carbon emission caps, green credit programs

DOI: 10.54254/2753-8818/2025.CH24595

4.2. Impact of recommendation implementation

Implementing the recommendation to increase the share of renewable energy in the HPC energy mix shows significant potential. Based on the model derived in Section 3.1 (specifically, the slope from Table 6), for every 5% increase in the proportion of renewable energy utilized by HPC facilities, carbon emissions can be reduced significantly (quantified by 5 × slope value, demonstrating the principle, though precise units depend on underlying scale). This highlights a scalable and impactful strategy.

4.3. Conclusion

High-Performance Computing is a critical enabler of scientific and technological progress, but its substantial environmental footprint, driven primarily by energy consumption and associated carbon emissions, neces- sitates urgent attention. This analysis has quantified the current energy use of HPC systems under both average and full capacity, modeled the resulting carbon emissions based on prevailing energy mixes, and projected future impacts considering growth trends and potential energy transitions up to 2030.

The models demonstrate a clear link between HPC operations, energy consumption, carbon emissions, and water usage. Importantly, they also quantify the significant potential for mitigating these impacts through the adoption of renewable energy sources. A strong linear relationship exists between increasing renewable energy penetration and decreasing carbon emissions. While challenges remain, particularly concerning grid stability and energy storage for 100% renewable scenarios, the path towards sustainability involves a concerted effort combining technological advancements (hardware/software efficiency, AI-driven optimization) and supportive policy measures (incentives, regulations).

Addressing the environmental impact of HPC is crucial not only for environmental stewardship but also for the long-term viability of the technology itself. Integrating sustainability considerations into HPC development and operation, supported by informed policy, is essential to align technological progress with global climate change mitigation and sustainable development goals.

Authorship contribution

All authors contributed equally to this work and should be considered co-first authors.

References

[1] Lovett, C. (2024). How to Advance Sustainable High Performance Computing? TierPoint. Retrieved from https: //www.tierpoint.com/blog/sustainable-high-performance-computing/
[2] Zaloumis, C. (2022, June 7). Are Your Data Centers Keeping You From Sustainability? IBM. ibm.com
[3] Electricity Maps. (2024, January 17). The United States 2023 Daily Carbon Intensity Data. Electricity Maps Data Portal. https: //www.electricitymaps.com/data-portal
[4] IEA. (2020, March). Energy demand in data centers worldwide from 2015 to 2021, by type (in terawatt hours). IEA. iea.org.
[5] U.S. Energy Information Administration. (2023, December 7). How much carbon dioxide is produced per kilowatthour of U.S. electricity generation? EIA. eia.gov
[6] Roundy, J. (2024). How to manage data center water usage sustainably. TechTarget. Retrieved fromhttps: //www.techtarget.com/searchdatacenter/tip/ How-to-manage-data-center-water-usage-sustainably
[7] U.S. Energy Information Administration. (2023, December 7). How much carbon dioxide is produced per kilowatthour of U.S. electricity generation? EIA. eia.gov

DOI: 10.54254/2753-8818/2025.CH24595

[8] Alves, B. (2024, October 24). Google water withdrawals 2016-2023. Statista. https: //www.statista.com/statistics/1498212/google-water-withdrawals-worldwide/
[9] Fernández, L. (2024, October 11). Google energy consumption 2011-2023. Statista. https: //www.statista.com/statistics/788540/energy-consumption-of-google/
[10] Hölzle, U. (2022). Our commitment to climate conscious data center cooling. Google Blog. Retrieved from https: //blog.google/outreach-initiatives/sustainability/ our-commitment-to-climate-conscious-data-center-cooling/