AI Energy Carbon Footprint Estimator (Lite)

Powered by PlanetAI Carbon Estimation Baseline Framework (PCE-BF v1.0)

Shiladitya Munshi, Kamalesh Karmakar
PlanetAI Research Lab

1 Executive Summary

The AI Carbon Footprint Estimator (Lite) is a public-facing sustainability assessment tool developed by PlanetAI Research Lab (PARL) under the PlanetAI Carbon Estimation Baseline Framework (PCE-BF v1.0). The tool enables researchers, developers, students, and organizations to estimate the energy consumption and carbon emissions associated with training and deploying artificial intelligence models.

As AI systems scale in complexity and deployment, their energy footprint has grown significantly (Strubell et al., 2019; Patterson et al., 2021), yet this impact often remains invisible to practitioners. The absence of accessible, lightweight carbon estimation tools creates a gap between AI innovation and sustainability awareness. The AI Carbon Footprint Estimator (Lite) addresses this gap by translating technical parameters—such as hardware type, number of GPUs, training duration, inference workload, deployment environment, and geographic location—into interpretable sustainability metrics.

Using structured baseline tables grounded in manufacturer specifications, benchmarking studies, cloud infrastructure efficiency data, and national carbon intensity statistics, the tool estimates:

- Training energy consumption (kWh)
- Inference energy consumption (kWh)
- Total carbon emissions (kg CO_2)
- Real-world equivalency indicators (e.g., driving distance, tree absorption)

The estimator is designed to be educational, transparent, and methodologically defensible. While it does not replace detailed lifecycle carbon accounting or hardware-level monitoring, it provides a standardized baseline framework for sustainability benchmarking and responsible AI deployment planning.

Strategically, this tool serves multiple purposes for PlanetAI Research Lab (PARL) :

- Demonstrates PARL's commitment to Green & Sustainable AI
- Establishes a measurable framework for planet-aware AI practice

- Provides a foundation for future research publications and empirical calibration
- Creates a scalable pathway toward advanced versions incorporating real-time carbon intensity data and telemetry integration

Planned for release on 30th March 2026, Version 1.0 (Lite) represents the first step in building an open, research-aligned sustainability intelligence ecosystem for artificial intelligence systems.

2 Problem Statement

Artificial Intelligence systems are rapidly increasing in scale, computational complexity, and deployment frequency across industries. Training modern deep learning models—particularly large transformer-based architectures—requires substantial computational resources, often involving multi-GPU clusters operating for extended durations. As AI adoption expands, the cumulative energy demand of model training and inference workloads is rising significantly.

Recent research has highlighted the environmental implications of large-scale AI training. Studies have shown that training state-of-the-art language models can consume substantial electricity and generate measurable carbon emissions, depending on hardware configuration, energy source, and geographic deployment context. While hyperscale cloud providers are improving data center efficiency and renewable integration, the environmental footprint of AI workloads remains non-trivial—especially when scaled across thousands of training experiments, fine-tuning cycles, and inference deployments worldwide.

Despite growing awareness, a critical gap persists:

- Most researchers and developers do not quantify the energy consumption of their AI experiments.
- Carbon impact is rarely reported in research publications.
- Existing enterprise-grade carbon accounting tools are complex, infrastructure-dependent, or inaccessible to students and small research groups.
- Sustainability considerations are often addressed post-deployment rather than integrated during model design and experimentation.

This disconnect creates a structural problem: AI innovation continues to accelerate without a lightweight, standardized mechanism for environmental visibility at the development stage.

Furthermore, AI sustainability discussions frequently rely on abstract metrics or high-level global statistics, making it difficult for individual practitioners to understand their own contribution to energy use and emissions. Without accessible estimation tools, sustainability remains conceptual rather than operational.

There is therefore a clear need for:

- A lightweight, publicly accessible carbon estimation framework
- A methodologically transparent approach grounded in technical specifications and empirical benchmarking
- A standardized baseline model for comparative sustainability analysis
- An educational instrument that integrates sustainability awareness into everyday AI practice

The AI Carbon Footprint Estimator (Lite) directly addresses this need by operationalizing sustainability measurement at the point of AI experimentation and deployment planning. By translating technical parameters into interpretable carbon impact metrics, the tool enables early-stage awareness, comparative benchmarking, and responsible AI decision-making.

It does not replace lifecycle carbon accounting (ISO 14064) nor hardware telemetry measurement. Instead, it provides structured baseline estimation aligned with Green AI principles (Schwartz et al., 2020).

In doing so, it contributes to aligning AI development practices with broader climate responsibility and sustainable computing principles.

3 Objectives

3.1 Primary Objectives

3.1.1 Enable Accessible Carbon Estimation

To provide a lightweight, publicly accessible tool that estimates the energy consumption and carbon emissions associated with AI training and inference workloads using standardized baseline assumptions.

3.1.2 Promote Sustainability Awareness in AI Development

To integrate environmental impact visibility into routine AI experimentation and deployment planning, encouraging sustainability-conscious design decisions at early stages.

3.1.3 Establish a Standardized Baseline Framework

To operationalize the PlanetAI Carbon Estimation Baseline Framework (PCE-BF v1.0) as a structured, transparent, and version-controlled estimation methodology.

3.1.4 Support Comparative Benchmarking

To enable researchers and organizations to compare carbon implications across different hardware configurations, model types, infrastructure choices, and geographic deployment contexts.

3.2 Secondary Objectives

3.2.1 Generate Foundational Research Infrastructure

To create a baseline dataset and methodological foundation that can be refined through empirical calibration and future research publications.

3.2.2 Strengthen PlanetAI's Sustainable AI Positioning

To demonstrate practical implementation of Green AI principles through a measurable, publicly accessible tool.

3.2.3 Encourage Responsible Deployment Planning

To assist developers and decision-makers in evaluating environmental trade-offs between local vs. cloud deployment, model scaling choices, and infrastructure selection.

3.2.4 Provide Educational Utility

To serve as a teaching instrument in academic and professional settings for illustrating the relationship between computational scale and environmental impact.

3.2.5 Enable Future Tool Evolution

To establish a scalable foundation for advanced versions incorporating real-time carbon intensity data, telemetry integration, institutional dashboards, and carbon-aware scheduling mechanisms.

3.3 Measurable Outcomes (Initial Targets)

- Within the first 12 months of release, the tool aims to:
- Achieve measurable public usage engagement
- Support sustainability discussions in academic or workshop settings
- Serve as a reference framework for at least one research or whitepaper output
- Provide structured feedback for refinement into Version 2.0

4 Scope of the Tool

The AI Carbon Footprint Estimator (Lite) is designed as a lightweight, structured sustainability estimation instrument under the PlanetAI Carbon Estimation Baseline Framework (PCE-BF v1.0). This section defines the functional boundaries of Version 1.0 and clarifies included and excluded components.

4.1 Functional Scope (Included in Version 1.0)

4.1.1 Training Energy Estimation

The tool shall estimate electricity consumption (kWh) associated with AI model training based on:

- Hardware type (GPU/CPU/TPU)
- Number of processing units
- Average utilization factor
- Training duration (hours)
- Model compute intensity multiplier
- Multi-GPU efficiency adjustment

4.1.2 Inference Energy Estimation

The tool shall estimate electricity consumption associated with inference workloads based on:

- Number of inference queries
- Average inference duration
- Hardware selection
- Inference utilization multiplier

4.1.3 Cloud Deployment Overhead Modeling

The tool shall incorporate infrastructure-level overhead factors using predefined PUE-based cloud multipliers to account for cooling, networking, and facility energy consumption.

4.1.4 Regional Carbon Emission Conversion

The tool shall convert total estimated energy consumption into carbon emissions (kg CO_2) using predefined national or regional electricity carbon intensity values.

4.1.5 Real-World Impact Translation

The tool shall translate estimated emissions into interpretable equivalency indicators, including:

- Driving distance equivalent
- Tree absorption equivalent
- Electricity consumption comparison

4.1.6 Version-Controlled Baseline Tables

All calculations shall rely exclusively on the locked baseline tables defined under PCE-BF v1.0. Any modification to assumptions requires formal version increment.

4.2 Operational Scope

- Public web-based interface
- Static baseline data tables
- Client-side calculation engine (no personal data storage required)
- Awareness and benchmarking use cases

4.3 Out of Scope (Version 1.0 Limitations)

To maintain clarity and prevent overextension, the following are explicitly excluded from Version 1.0:

- Real-time hardware-level power monitoring
- Direct integration with GPU telemetry APIs
- Hourly grid carbon intensity tracking
- Marginal emission intensity modeling
- Lifecycle carbon accounting (manufacturing, hardware supply chain, end-of-life disposal)
- Renewable energy procurement verification
- Institutional compliance certification
- Automated carbon offset purchasing mechanisms

4.4 Intended Use Context

The tool is intended for:

- Educational purposes
- Sustainability awareness
- Comparative benchmarking
- Early-stage deployment planning
- Research methodology illustration

It is not intended for regulatory reporting, audited carbon accounting, or formal environmental compliance documentation.

4.5 Versioning Commitment

Any expansion beyond the above scope—including dynamic carbon APIs, hardware telemetry integration, or lifecycle accounting—will be introduced only under formally declared future versions (e.g., v2.0, v3.0).

5 Methodology Overview

The AI Carbon Footprint Estimator (Lite) is built upon the **PlanetAI Carbon Estimation Baseline Framework (PCE-BF v1.0)**, which integrates hardware specifications, workload characteristics, infrastructure overhead assumptions, and regional carbon intensity factors into a structured estimation model.

This section outlines the computational methodology and underlying assumptions used in Version 1.0.

5.1 Estimation Philosophy

The methodology follows a deterministic engineering-based estimation approach. It does not rely on runtime telemetry or empirical measurement, but instead uses:

- Manufacturer-reported hardware specifications (NVIDIA, AMD, Intel documentation)
- Empirical utilization assumptions (Uptime Institute; Google Data Center Reports)
- Standardized cloud infrastructure overhead factors (Peer reviewed research)
- National electricity carbon intensity averages (IEA, Our World in Data)

The framework is designed for awareness, benchmarking, and comparative sustainability analysis rather than precise lifecycle accounting.

5.2 Core Variables

The following variables are used in the estimation model:

- **P** = Hardware power rating (Watts)
- **U** = Average utilization factor (0–1)
- **T** = Training duration (hours)
- **N** = Number of GPUs / processing units
- **E_s** = Multi-GPU scaling efficiency factor
- **M** = Model compute intensity multiplier
- **C** = Cloud overhead multiplier (PUE-based)
- **I** = Regional carbon intensity (kg CO_2 per kWh)

5.3 Training Energy Estimation

Training energy consumption (in kWh) is estimated as:

$$E_{train} = \frac{P \times U \times T \times N \times M \times C \times E_s}{1000}$$

Where:

- Power (P) is expressed in Watts
- Division by 1000 converts Watt-hours to kilowatt-hours

This formulation accounts for:

- Hardware energy demand
- Workload intensity
- Distributed training efficiency
- Infrastructure overhead

5.4 Inference Energy Estimation

Inference workload energy is computed separately.

Let:

- **Q** = Number of inference queries
- **t_i** = Average inference time per query (seconds)
- **U_i** = Inference utilization multiplier

Total inference time in hours:

$$T_i = \frac{Q \times t_i}{3600}$$

Inference energy:

$$E_{inf} = \frac{P \times U_i \times T_i \times N \times C}{1000}$$

5.5 Total Energy Consumption

$$E_{total} = E_{train} + E_{inf}$$

5.6 Carbon Emission Estimation

Total carbon emissions are calculated using regional carbon intensity:

$$CO_2 = E_{total} \times I$$

Where:

- **I** represents average grid emission intensity or regional carbon intensity in kg CO_2 per kWh.

  Carbon intensity references derived from International Energy Agency (IEA, 2023) and Our World in Data electricity carbon intensity dataset

5.7 Real-World Equivalency Conversion

Carbon output is translated into public-facing equivalency metrics using predefined reference constants:

$$Driving\ Distance = \frac{CO_2}{\text{Emission per KM}}$$

$$Tree\ Equivalent = \frac{CO_2}{\text{Annual Tree Absorption}}$$

These are illustrative indicators intended for interpretability.

5.8 Multi-GPU Efficiency Adjustment

To prevent unrealistic linear scaling assumptions, a distributed efficiency factor ($\mathbf{E_s}$) is applied based on GPU count ranges. This accounts for communication overhead and synchronization cost in parallel training environments.

5.9 Cloud Overhead Modeling

Cloud overhead multiplier (**C**) approximates Power Usage Effectiveness (PUE):

$$C \approx PUE$$

This accounts for cooling, networking, and facility-level infrastructure energy demand.

5.10 Baseline Data Tables

All calculations in the AI Carbon Footprint Estimator (Lite) rely on version-controlled baseline tables defined under the PlanetAI Carbon Estimation Baseline Framework (PCE-BF v1.0). These include:

- Hardware Power & Utilization Table
- Model Compute Multiplier Table
- Cloud Overhead Multiplier Table
- Regional Carbon Intensity Table
- Carbon Equivalency Reference Table

- Inference Utilization Table
- Multi-GPU Efficiency Table

The complete tables are documented in Appendix A.

5.11 Uncertainty Considerations

The estimator uses deterministic baseline values subject to uncertainty in:

- Runtime power variation
- Workload utilization variability
- Regional carbon intensity fluctuation
- Distributed training efficiency

Estimated uncertainty ranges across components are:

- Hardware power: ±10–20%
- Model multiplier: ±20–30%
- Carbon intensity: ±5–15%
- Overall combined estimate: approximately ±20–30%

The tool does not propagate statistical uncertainty but communicates approximation through disclaimer and transparency.

5.11 Methodological Limitations

The current framework does not include:

- Embodied carbon in hardware manufacturing
- Data center construction emissions
- Renewable energy credit adjustments
- Marginal grid intensity modeling
- Time-of-day energy variation

These may be considered in future versions.

5.12 Version Control

All parameters and formulas correspond to **PCE-BF v1.0 (March 2026)**. Any modification to baseline tables or estimation logic requires formal version increment.

6 Impact Assessment & Success Metrics

The AI Carbon Footprint Estimator (Lite) is intended not only as a computational tool but as a strategic intervention in promoting sustainability-aware AI development. This section defines the anticipated impact dimensions and measurable success indicators for Version 1.0.

6.1 Impact Dimensions

The tool's impact is expected across four primary domains:

6.1.1 Educational Impact

- Enhances awareness of AI energy consumption among students and researchers
- Integrates sustainability into AI curricula and workshops
- Supports teaching modules on Green and Sustainable AI

6.1.2 Research Impact

- Establishes the PlanetAI Carbon Estimation Baseline Framework (PCE-BF v1.0) as a reference methodology
- Provides a foundation for empirical calibration studies
- Supports publication of whitepapers or peer-reviewed research

6.1.3 Behavioral Impact

- Encourages developers to consider energy implications before scaling models
- Promotes comparative decision-making (e.g., hardware choice, cloud vs. local deployment)
- Encourages reduction-oriented thinking rather than offset-only thinking

6.1.4 Institutional & Policy Impact

- Positions PlanetAI Research Lab as a contributor to Green AI discourse
- Provides an accessible benchmarking framework for academic and small research institutions
- Supports sustainability-oriented research proposals and collaborations

6.2 Key Performance Indicators (KPIs)

The following measurable indicators will be tracked during the first 12 months post-release:

6.2.1 Usage Metrics

- Total number of tool visits
- Unique user sessions

- Average interaction duration
- Number of completed estimation runs

6.2.2 Engagement Metrics

- Downloads of result summaries (if enabled)
- Workshop or academic usage instances
- Institutional references to the tool

6.2.3 Research Metrics

- Citations or references to PCE-BF framework
- Inclusion in academic teaching material
- Publication of at least one methodological note or whitepaper

6.2.4 Sustainability Awareness Indicators

- Feedback indicating increased awareness
- Requests for advanced versions or institutional deployment
- Adoption in sustainability discussions or AI ethics forums

6.3 Qualitative Impact Assessment

Beyond quantitative metrics, qualitative impact will be assessed through:

- User feedback forms
- Expert review comments
- Workshop evaluations
- Academic collaboration discussions

These insights will inform Version 2.0 development priorities.

6.4 Success Criteria for Version 1.0

Version 1.0 will be considered successful if it achieves:

- Demonstrable public engagement
- Methodological credibility without major technical issues
- Positive academic or expert feedback
- Foundation establishment for framework evolution

6.5 Long-Term Impact Vision

Over time, the tool aims to evolve into:

- A calibrated sustainability estimation framework
- An API-enabled carbon-aware planning system
- An institutional dashboard for research labs
- A publishable sustainability benchmarking standard

The Lite version represents the foundational step toward that broader vision.

6.6 Review Cycle

Impact assessment will be conducted:

- 6 months post-launch (initial review)
- 12 months post-launch (comprehensive evaluation)

Findings will inform version updates and methodological refinements.

Disclaimer

PlanetAI Carbon Estimation results are approximate engineering estimates derived from publicly available technical specifications, benchmarking studies, and national carbon intensity datasets. Values are intended for awareness, sustainability benchmarking, and comparative analysis. They do not replace detailed hardware-level power measurement or full lifecycle carbon accounting.

References

Andrae, A. S. G., & Edler, T. (2015). On global electricity usage of communication technology. *Challenges*, 6(1), 117–157.

Google. (2023). *Data Center Efficiency & Sustainability Report*. Google LLC.

International Energy Agency (IEA). (2023). *Emissions Factors 2023*. Paris: IEA.

ISO. (2018). *ISO 14064-1:2018 Greenhouse gases — Specification with guidance*. International Organization for Standardization.

MLCommons. (2023). *MLPerf Training Benchmark Results*. https://mlcommons.org

NVIDIA Corporation. (2023). *GPU Architecture Whitepapers and Product Specifications*. NVIDIA.

Our World in Data. (2023). *Carbon intensity of electricity production*. University of Oxford.

Patterson, D., Gonzalez, J., Le, Q., et al. (2021). Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*.

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63.

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *ACL 2019*, 3645–3650.

Uptime Institute. (2023). *Global Data Center Survey Report*.

APPENDIX – 1

TABLE 1 - Hardware Power & Utilization

🔷 Consumer GPUs

Hardware	Power (W)	Avg Utilization
GTX 1660	120	0.70
RTX 2060	160	0.75
RTX 3060	170	0.75
RTX 3070	220	0.80
RTX 3080	320	0.85
RTX 3090	350	0.85
RTX 4090	450	0.85

🔷 Data Center GPUs

Hardware	Power (W)	Avg Utilization
Tesla T4	70	0.70
V100	250	0.85
A100 40GB	350	0.85
A100 80GB	400	0.90
H100	700	0.90
AMD MI250X	560	0.85

🔷 TPUs

Hardware	Power (W)	Utilization
TPU v2	280	0.85
TPU v3	450	0.90
TPU v4	600	0.90

🔷 CPUs

Hardware	Power (W)	Utilization
Laptop CPU	30	0.60
Desktop i7	95	0.70
Xeon Server CPU	150	0.75
Dual Xeon	300	0.80

TABLE 2 — Model Compute Multipliers

Model Type	Multiplier
Logistic Regression	0.3
Decision Tree	0.4
Random Forest	0.6
Small CNN	0.7
ResNet-50	0.9
EfficientNet	0.9
Transformer Base	1.0
BERT Base	1.0
BERT Large	1.3
GPT-2 Small	1.2

Model Type	Multiplier
GPT-3 Scale (normalized)	3.0
Large LLM Fine-tuning	1.5
Diffusion Model	1.4

TABLE 3 — Cloud Overhead (PUE-Based Multipliers)

Deployment Type	Multiplier
Local Desktop	1.0
On-Prem Server	1.1
AWS (Std Region)	1.2
Azure	1.2
GCP	1.2
Energy-Optimized Data Center	1.1
Hyperscale AI Cluster	1.3

TABLE 4 — Regional Carbon Intensity (kg CO_2 per kWh)

🔷 Asia

Country	Intensity
India	0.70
China	0.58
Japan	0.43
Singapore	0.40
South Korea	0.45

◆ Europe

Country	Intensity
Germany	0.35
France	0.06
UK	0.20
Norway	0.02
EU Average	0.25

◆ Americas

Country	Intensity
USA	0.40
Canada	0.15
Brazil	0.10
Global Average	0.475

TABLE 5 — Carbon Equivalency Reference Values

Activity	CO_2 Equivalent
1 km driving (petrol)	0.12 kg
1 hour streaming	0.055 kg
1 kWh electricity (global avg)	0.475 kg
1 tree absorption per year	21 kg
Short-haul flight (per passenger)	150 kg

TABLE 6 — Inference Utilization Multipliers

Inference Type	Multiplier
Batch inference	0.7
Real-time API	0.8
High-throughput inference	0.9
Edge inference	0.6

TABLE 7 — Multi-GPU Efficiency Scaling

GPU Count Range	Efficiency Factor
1–4	1.0
5–16	0.95
17–64	0.90
65+	0.85