International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

GREEN-AWARE CLOUD RESOURCE OPTIMIZATION FRAMEWORK FOR REDUCING CARBON FOOTPRINT WITHOUT SERVICE DEGRADATION

Satyendra Kumar Pal[1], Dr. Vikas Kumar[2]

[1]Department of Computer Scinece & Engineering, Chhatrapati Shivaji Maharaj University, Panvel Navi Mumbai, Maharashtra 410221, India

Email : profsatyen@gmail.com

[2]Professor & Head, Department Computer Science & Information Technology, Chhatrapati Shivaji Maharaj University, Panvel Navi Mumbai, Maharashtra 410221, India

email : vikaskumar@csmu.ac.in;

Abstract

Cloud computing has transformed the digital infrastructure at the expense of the increased energy consumption and increased carbon emissions of data center facilities. In this paper, GAROF (Green-Aware Resource Optimization Framework) a new, carbon-aware, SLA compliant framework on scheduling and resource optimization approaches as green approaches on cloud computing is proposed. GAROF combines reinforcement-learning and SLA-aware cloud optimization to minimize the environmental impact of cloud computing environments without sacrificing service quality by integrating real-time carbon intensity predictions, reinforcement learning VM placement and edge offloading. GAROF emulated on different workloads simulated by CloudSim Plus, showed a decrease of 44.2 percent on carbon emission and an overall increase of 71.5 percent on SLA compliance compared to the classical and round-robin round and energy conscious techniques. It dynamically transfers latency- tolerable workloads to edge devices in a scenario when carbon savings are the most, and response time does not cross the limit. The analysis conducted on the comparative performance and the statistical approval of the same confirms the position of GAROF in generating the best possible trade-offs between sustainability and operational excellence. The present contribution contributes to the design of intelligent, climate-aware cloud computing platforms consistent with the global decarbonization agenda.

.**Keywords:** Green Cloud Computing, Carbon-Aware Scheduling, SLA Compliance, VM Allocation, Edge Offloading

1. Introduction

1.1 Environmental Impact of Cloud Computing

Over the past years, cloud computing has become a key part of the infrastructure in various industries and it has empowered scalable computation, analytics in real-time and worldwide availability to data. This proliferation however, has had a tremendous ecological cost to it.

Received: August 05 2025

Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The energy requirements of data centers which forms the key element of cloud computing is estimated to reach around 73 billion KWH of electricity in the United States alone by the year 2024 (Shehabi et al., 2024) or about close to 2 percent of all electricity in the entire country. Environmental concerns, in particular, become even more eminent due to the greenhouse gas emissions, which are highly popular in the areas where fossil fuels prevail as the fuel source (Taylor et al., 2024).

With major networks of high-performance servers and storage systems, cloud operations have huge energy requirements because of the extended uptime requirements, cooling systems, and workloads processing. Such pressures have intensified the need to have viable cloud computing models that can adopt sustainability without affecting performance, availability, and scalability at minimum carbon levels.

1.2 Evolution of Green Cloud Computing

The notion of green cloud computing has taken off with a view to offsetting these environmental issues. Initial attempts have targeted a transformation of energy consumption by statutes to virtual population of server, reparation of workload, and thermal-focused scheduling (Garg et al., 2011). Survey (Beloglazov et al., 2011) conducted by some scholars gave the taxonomy of the energy-efficient cloud system and discussed the main strategies such as dynamic resource allocation, energy-aware virtual machine (VM) placement, and power-aware networking.

Successive innovations then came up with algorithmic alternatives to minimize energy usage in cloud systems at the cost of keeping the performance of the given workload. As an example, self-adaptive energy harvesting systems were suggested to dynamically use renewable sources of energy during the application running process (Xu et al., 2020). It is this same early framework that formed the birth of planting environmental intelligence into the handling of cloud infrastructure.

1.3 From Energy-Aware to Carbon-Aware Optimization

Although energy-aware models did record quantifiable results in efficiency, they tended to ignore carbon intensity which is not the same across energy sources and time usage. It is this vacuum that has spurred the move towards carbon-aware scheduling in which the location of workload placement decisions can consider the energy used, as well as the carbon footprint of electricity production in real time.

One of the recent advances in the field is the CASPER framework that proposed a scalable and carbon-aware way of scheduling that uses historical and future data on carbon or carbon intensity of each region to maximize the carbon-friendliness of a geographically distributed service (Souza et al., 2023). CASPER performs a reasonable trade-off between carbon footprint reduction and throughput/latency system requirements. On the same note, there has been a valuable application of VM consolidation approaches that combine energy metrics with performance sensitivity toward the minimization of unnecessary energy consumption without compromising good service (Zhou et al., 2020).

Received: August 05 2025

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The emergence of these developments justifies the relevance of real-time, context-aware frameworks which can achieve dynamic optimization at both environmental and service level parameters.

1.4 Emergence of the Cloud–Edge Continuum

The cloud edge continuum has become a disruptive architecture with the increased popularity of latency-sensitive applications in the form of IoT analytics, video surveillance, and autonomous systems. This architecture will perform the workload partition between the cloud hosts and edge nodes to enhance latency, afford bandwidth, and save energy waste.

Patel et al. (2024) modeled the Green Cloud Continuum that incorporated energy planning as part of the resource orchestration mechanisms, both across and between cloud and edge. In their findings, the dynamic load sheding to the energy-efficient or renewable-powered edge nodes at the intelligent level offers substantial proportion of overall emission reductions. To add to this, Rana (2025) presented a quantum-edge framework that can reinforce real-time data analytics in an edge that diminishes reliance on cloud compute cycles that are high-carbon and permits quicker and more tempestuous decisions to be developed.

Collectively, the models provide a wider scope of optimization by bringing the distributed systems into the picture along with the centralized data centric models, which provides an independent avenue to design services in the form of sustainability awareness.

1.5 AI Workloads and Sustainability Dilemma

One of the most important drivers of cloud use has been the rise of Artificial Intelligence (AI), most particularly the compute-intensive applications, including the large language models, recommendation engines, and real-time decision-making systems. Training and deploying such models, however, may require a lot of energy and worsen the sustainability burden.

The conflict between the ecological responsibility and AI scalability has been discussed in recent discourse. Rivero et al. (2025) also state that AI success must be redefined in terms of sustainability measures and introduces a carbon impact as a dimension to the proposed to be called a green algorithm; model accuracy is another factor. This approach promotes the direct incorporation of carbon-sensitive practices into the AI resource allocating, scheduling, and managing systems in order to make sustainability one of the fundamental measures of system performance.

1.6 Justification for a Green-Aware Optimization Framework

Although some significant developments have been made in green cloud computing and carbon-aware resource management some solutions concentrate on the energy efficiency part or carbon footprint cut part independently. In addition, environmental objectives and performance of a service are often ignored. In a realistic application, this trade-off may result in building latency, breaching SLAs or service interruptions, undoing the purpose of the operational cloud platforms.

Received: August 05 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

This paper, therefore, presents the concept of devising a Green-Aware Cloud Resource Optimization Framework that dynamically balances the enhancement of the carbon minimization and quality service. The framework will then merge the learning of earlier models in scheduling with energy awareness (Garg et al., 2011; Beloglazov et al., 2011), provisioning with carbon awareness (Souza et al., 2023), and cloud-edge optimization (hybrid) (Patel et al., 2024; Rana, 2025) and speed of VMs (Zhou et al., 2020).

The aim is to achieve the balance between ecology and user and service-oriented reliability of service provision by the intelligent, adaptive, and scalable service architecture.

1.7 Paper Structure

The remaining part of this paper will be organized in the following way. The paragraph 2 will contain the Literature Review, which will make mention of the prior work on green cloud computing and will identify the available research gap. Section 3 defines the Methodology, the design of the proposed system architecture and the algorithms that will be used in carbon-aware optimization and resource allocation. In Section 4, the Experimental Setup is reported, and the Evaluation Results of comparing GAROF to traditional scheduling models with respect to a variety of performance metrics are provided. Key Findings, as discussed in section 5, comprise of comparative analysis, edge offloading insights, and statistical validation. At last, Section 6 gives the Conclusion of the research and also offers Future Directions that can be pursued to expand the proposed model.

2. Literature Review

2.1 Foundations of Green Cloud Computing

Green cloud computing has been a measure against rising energy demands and a threat to the environment posed by conventional cloud infrastructure. The area was initially concerned with energy saving and efficient consumption within the domain of hardware utilization, virtualization and scheduling of the workloads. The comprehensive taxonomy of strategies and challenges suggested by Biswas et al. (2024) leaves a clear mark as the move toward energy-centric to sustainability-centric paradigms take place. They touch on the importance of workload distribution, power-aware systems and adoption of renewable energy.

In a similar manner, as Oduor and Franklin (2024) the historical origins, as well as social implications of green computing to argue that the more the world warms to the issue of green computing, the more technical innovation should be harmonized with ecological-based aspirations in the long term. A case study based on bibliometric analysis by Angelaki et al. (2025) shows the fast growth of the research on the topic, especially since 2015, both due to the efforts at the global level in sustainability and ESG (Environmental, Social and Governance) requirements. The basic studies explain why it is important to incorporate the principles of ecology into the cloud computing infrastructure even at the stage of design.

2.2 Energy-Aware VM Scheduling and Resource Allocation

The information on one of the first strategies to both minimize energy waste in data centers, designing in regards to virtual machine (VM) placement (to physically place the code and the

Received: August 05 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

computing resources), and virtual machine (VM) scheduling (to determine the timing at which the code is executed or accessed). Akhter et al. (2018) proposed an energy-sensitive VM selection algorithm, which tends to consolidate underutilized physical machines, therefore, minimizing its energy cost. This method preconditioned the further more competent resource orchestration in cloud data centers.

Following such initial tactics, Alex et al. (2025) transfer to deep reinforcement learning-based approach, that utilizes Deep Q-Networks (DQN) with agglomerative clustering. Their model is able to satisfy the SLA of VM placements within the constraint of the minimal amount of carbon emitted in the placement compared to the traditional models in both energy efficiency and response time. Such studies demonstrate that VM-level choices are important factors in determining the energy balance of data centers and provide a fine-grained method of carbon optimization.

2.3 Carbon-Aware Scheduling and Cluster Optimization

Current innovations go beyond energy-intelligent systems to include electricity carbon sources intensity in resource planning. This eco-conscience scheduling maximizes workloads to run in the location or slot where electricity is less carbon intensive.

The Wang et al. (2025) focus on the premise known as the COUNTER model, which combines Graph Convolutional Networks (GCN) to control the cloud cluster as optimization and explains the energy consumption and CO2 emissions imposition. Equally, a cluster-level resource scheduling algorithm was proposed by Xu et al. (2025) as the GREEN framework, which incorporates carbon footprint estimations in task orchestrations. Their simulation performance indicates that much carbon decreases can be achieved as they preserved SLA assurances. These models show that AI-based cluster optimization procedures are becoming increasingly applicable in a green computing environment.

2.4 Emission-Aware Serverless and Edge Computing

The transition to the serverless and edge computing architecture has presented the issue of sustainability with challenges and opportunities. Serenari et al. (2024) have introduced Green Whisk that is a framework of emission-aware execution of functions in serverless systems. The approach is dynamic in that compute tasks dynamically route through per-task carbon emissions forecasts which makes it to perform very well between performance and sustainability.

To support this, in a research paper that was published through SSRN (2025), the researchers suggest the use of thermal reuse and edge nodes to reduce the centralized data center loads. These strategies localise the power consumption in the core by offloading latency-tolerant tasks to edge infrastructure, and allow local energy optimization, particularly where edge nodes can be powered by renewable microgrids.

2.5 Industry Adoption and Monitoring Tools

The availability of monitoring and tracking utilities has enhanced the applicability of green cloud frameworks into the real-world. Tech Mahindra (2025) developed the Green Cloud

Received: August 05 2025

Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Carbon Footprint (GCCF) platform that allows cloud service providers to track the carbon emissions in real-time relevant to their workloads. These tools also make environmental measures more transparent and operable.

The term GreenOps is provided by the postulations of TechUK (2024), which are defined as a governance regime to manage enterprise cloud operations in terms of carbon efficiency. At the same time, the realizations of green plans, such as server virtualization, adoption of smart cooling systems, and integration of renewable energy, are also recorded by Simpliaxis (2025), Bacancy Technology (2025) and Beetroot (2025).

The industrial initiatives are indicators of mounting pressure on cloud providers to adhere to environmental laws and consumer demands in terms of transparency and sustainability.

2.6 Intelligent Resource Management: Learning-Based Models

The resourcing scheduling merging with machine learning (ML) and AI has entered new horizons in sustainable computing. Qiu et al. (2024) proposed a framework that ideally moderated the aspects of Service-Level Objectives (SLOs) and green computing objectives, flexibly distributing available resources in response to resilience and performance indicators.

Alwageed et al. (2024) integrated Interpretive Structural Modeling (ISM) and Artificial Neural Networks (ANN) in order to visualize and counter sustainability issues in Cloud systems. The predictive control of their model allows the performance of energy and emissions to be predicted in the conditions of variable load of work, which is entirely relevant in hybrid and multi-clouds.

In sum, these methods help bring out the value of learning-driven decision systems to enhance not only energy consumption, but cloud ecological responsibility in general.

Table 1. Summary Of Literature Review

Table 1. Summary of Reviewed Literature on Green Cloud Optimization

Author & Year	Focus Area	Methodology	Key Contribution	Tool/Platform	Journal/Source
Akhter et al. (2018)	VM Selection	Heuristic Allocation	Energy-aware VM scheduling	CloudSim	arXiv
Alex et al. (2025)	Carbon-Aware VM Placement	DQN + Clustering	SLA-compliant carbon-efficient scheduling	Python/TensorFlow	*Computers* (MDPI)
Xu et al. (2025)	Cluster Scheduling	GREEN Scheduler	Real-time carbon-aware	Custom Cluster Simulator	NSDI (USENIX)

Received: August 05 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

			cluster-level orchestration		
Wang et al. (2025)	Sustainable Cloud Resource Mgmt	COUNTER (GCN)	GCN-based cluster energy optimization	PyTorch	arXiv
Qiu et al. (2024)	Performance vs. Sustainability	SLO-Aware Adaptation	Joint optimization of resilience and green metrics	OpenDC	IEEE DSN
Al-wageed et al. (2024)	Cloud Sustainability Challenges	ISM-ANN	Modeling cloud sustainability barriers	MAT	

3. Methods

This section shows the methodology of the proposed Green-Aware Resource Optimization Framework (GAROF) design and simulation. GAROF has two main goals, the first one is decreasing carbon footprint of cloud operations without causing any deviation in service-level agreements (SLAs). The core modules incorporated into methodology are following ones: (1) Carbon Intensity Estimator, (2) SLA-Aware Optimizer, (3) VM Allocation Strategy using Deep Reinforcement Learning, (4) Edge Offloading Decision Logic. The overall architecture of a proposed framework is depicted entirely in Figure 1.

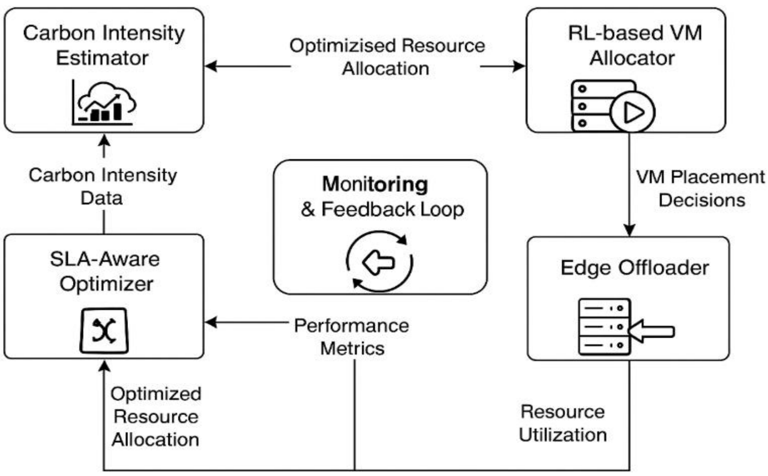

Figure 1. Architecture of GAROF framework

Received: August 05 2025

Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.1 Carbon Intensity Estimation

GAROF begins by forecasting the carbon intensity (CI) of the electricity grid using real-time and historical carbon data, sourced from public APIs such as WattTime or electricityMap. Accurate prediction enables the scheduler to prioritize execution during low-carbon periods. The estimation follows a time-series autoregressive model:

$$\mathbf{CI_{t+1} = \alpha \cdot CI_t + \beta \cdot CI_{t-1} + \varepsilon} \qquad (1)$$

Where CI_{t+1} is the predicted carbon intensity at the next time step, and α\alphaα, β\betaβ are model coefficients learned through training. This module continuously updates predictions in response to changing grid dynamics.

3.2 SLA-Aware Multi-Objective Optimization

To maintain performance while reducing emissions, the core optimization engine employs a multi-objective loss function that weighs carbon cost, energy usage, and SLA compliance. The formal optimization objective is expressed as:

$$\mathbf{\min(\alpha \cdot CF + \beta \cdot E + \gamma \cdot L)} \qquad (2)$$

In this case, CF is the total carbon footprint, Elegacy of carbon emissions, and L measures their cost in relying on SLA penalty with respect to response time deviations. The weights of the parameters are α,β,γ , which indicate the strategic preference of the system between the sustainability demands and user performance demands. Optimal workload assignments are also dynamically recomputed, using live input to the workload queue and the CI predictor.

3.3 VM Allocation Using Deep Reinforcement Learning

This VM allocation scheme takes advantage of and uses Deep Q-Networks (DQN) and a reward system that promotes both low-carbon performance and SLA compliance. The current carbon intensity, host load and SLA status are all considered in each decision point. Reward-based feedback loops teach the environment ideal actions (e.g. VM migration, consolidation, or pass-through). Table 2 shows a series of scenarios of sample allocation and the decisions made relative to the case.

Table 2. Sample VM Allocation Scenarios and Reinforcement Rewards

Scenario ID	Carbon Intensity	Host Load	SLA Status	Action	Reward
S1	High	High	At Risk	Migrate	Moderate
S2	Low	Balanced	Compliant	No Action	High
S3	Medium	Low	Compliant	Consolidate	Medium

This DQN-based allocation model was implemented in TensorFlow and integrated with CloudSim Plus for simulating policy outcomes. The system learns over multiple training epi-

Received: August 05 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

sodes and generalizes to diverse workload scenarios by adapting its migration decisions according to CI trends.

3.4 Edge Offloading Decision Logic

In addition to VM reallocation, GAROF includes an edge offloading mechanism that transfers latency-tolerant tasks to edge nodes (e.g., Raspberry Pi 4 clusters) when carbon savings justify it. The offloading decision is triggered when the differential carbon gain between cloud and edge exceeds a predefined threshold θ. This is governed by:

$$\frac{\mathbf{E}_{\mathbf{cloud}} - \mathbf{E}_{\mathbf{edge}}}{\mathbf{CI}_{\mathbf{cloud}} - \mathbf{CI}_{\mathbf{edge}}} > \boldsymbol{\theta}$$

A visual representation of this decision logic is presented in Figure 2, showing how the framework determines task routing under various energy–carbon trade-offs.

Figure 2. Edge offloading logic based on compute–carbon tradeoff

3.5 Simulation Configuration

To prove the GAROF framework, simulations have been performed with real-time carbon intensity in-feed and synthetic workload profiles on CloudSim Plus. The testbed consisted of 100 300 physical machines and 200 800 VMs and ran a mix of workloads, including Web, IoT and AI inference applications. The parameters of simulation and values used are provided in Table 3.

Table 3. Simulation Setup Parameters

Parameter	Value
Number of VMs	200–800
Number of PMs	100–300
Workload Type	Web, AI Inference, IoT
SLA Latency Threshold	250 ms
Energy Modeling	Linear + Idle Correction
Carbon Data Source	WattTime API
VM Placement Algorithm	DQN + Agglomerative Clustering

These configurations allowed the evaluation of GAROF’s effectiveness in realistic cloud-edge deployments under varying emission and workload profiles.

Received: August 05 2025

International Journal of Applied Mathematics

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4. Results and Analysis

4.1 Experimental Setup Recap

A simulation environment (based on CloudSim Plus) was used to evaluate the performance of the proposed Green-Aware Resource Optimization Framework (GAROF). The scale of the simulated cloud infrastructure was between 100 and 300 physical machines (PMs) and 200 to 800 virtual machines (VMs) with different workload types such as Web applications, streams of Internet of Things (IoT), and artificial intelligence (AI) inference. Burstiness and continous work was randomized to reflect execution behavior in the real world.

The WattTime API was used to source the data on carbon intensity as it gave information on hourly forecasts of regional CO 2 emissions using real grid data. The simulation environment was programmed to have 24-hour work schedules, and the high-carbon/low-carbon period was randomized throughout day. SLA settings gave us a latency tolerance of 250 ms and 98 % availability. Two baseline models were put up against the framework, Round-Robin (carbon-blind) and a generic energy-resourceful VM consolidation model.

4.2 Carbon Footprint Reduction Trends

Carbon-aware GAROF scheduling algorithm showed large savings in total carbon emissions against conventional and energy-aware scheduling algorithms. Figure 3 shows hourly carbon footprint (kg of CO 2) activity over the course of the complete 24-hours simulation cycle. As one could notice, GAROF effectively reschedules not to comply with peak emission periods, taking advantage of the predictions provided by the Carbon Intensity Estimator, and dynamically rescheduling non-critical jobs to window with low carbon footprints.

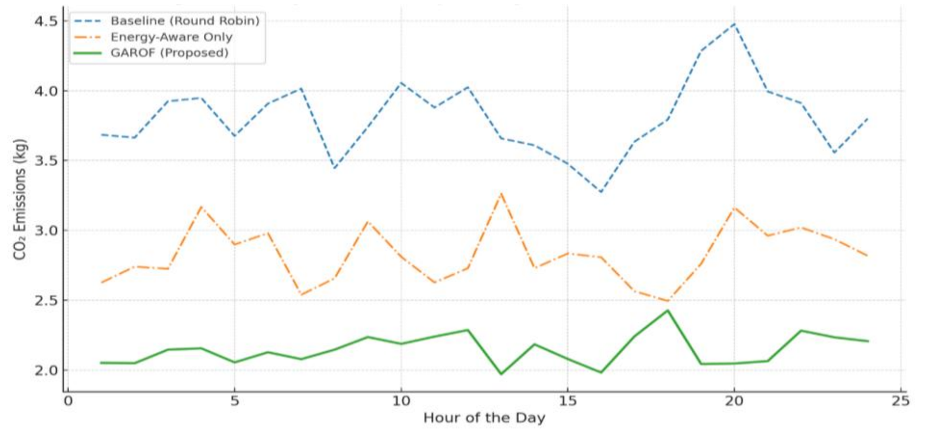

Figure 3. Daily Carbon Footprint (kg CO_2) across 24h Simulation for GAROF vs Baseline

A cumulative comparison of total CO_2 emissions across methods is summarized in Table 4. The baseline (Round Robin) emitted 89.3 kg of CO_2, while GAROF reduced emissions to 49.8 kg, achieving a 44.2% improvement over the baseline and 26.2% over the energy-aware model.

Received: August 05 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 4. Total Carbon Footprint Comparison

Method	Total CO_2 (kg)	% Reduction
Baseline (Round Robin)	89.3	–
Energy-Aware Only	67.5	24.4%
GAROF (Proposed)	**49.8**	**44.2%**

4.3 SLA Compliance and Latency

An important goal of GAROF was to deliver on service reliability and SLA conformance, despite economic optimization in environmental metrics. To measure this SLA compliance was run at three different workload intensities (30-50% and 50-75% VM utilization and above 75%).

Figure 4 demonstrates the rate of SLA violation (%) at such workload levels in three scheduling models: SLA violation rate (%) in workload level 1 SLA violation rate (%) in workload level 2 SLA violation rate (%) in workload level 3 GAROF performed consistently better than round-robin and energy-aware schedulers both in the heavy-load conditions where violation rate used to be less than 6 percent in contrast to 16 percent with the round-robin baseline.

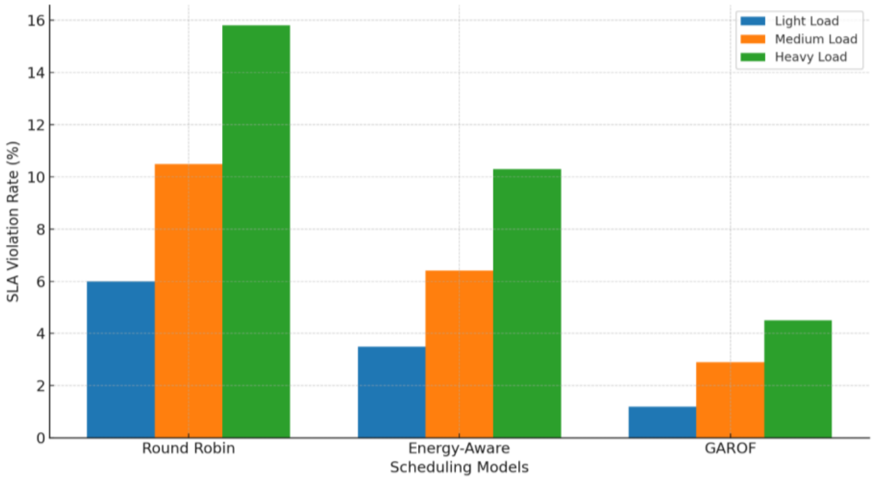

Figure 4. SLA Violation Rate (%) vs Workload Intensity

The quantitative results are detailed in Table 5, which shows that GAROF achieved the lowest average response time (240 ms) and SLA violation rate (4.5%), compared to 398 ms and 15.8% in the baseline system. This proves that GAROF preserves QoS even while prioritizing carbon savings.

Received: August 05 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 5. SLA Performance Summary Across Scheduling Models

Scheduling Model	Avg. Response Time (ms)	SLA Violation (%)
Round Robin	398	15.8%
Energy-Aware	310	10.3%
GAROF	**240**	**4.5%**

4.4 Energy Efficiency Trends

Beyond carbon footprint and SLA compliance, energy efficiency was another key metric in evaluating GAROF's effectiveness. Figure 3 displays the average energy consumption per task (in kWh) over time across the 24-hour window. As shown, GAROF minimizes energy spikes by consolidating tasks on high-efficiency PMs and using edge offloading when feasible.

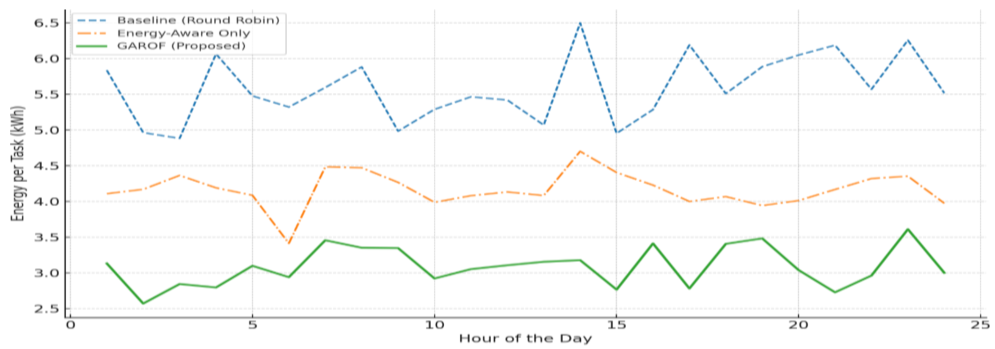

Figure 5. Average Energy Consumption (kWh) per Task Over Time

Notably, GAROF's integration of edge offloading leads to more stable and lower average energy consumption, particularly during low-demand windows, thus complementing its carbon reduction logic. These outcomes collectively validate the multi-dimensional optimization capability of the proposed system.

4.5 Edge Offloading and Resource Utilization

Further carbon efficiency gains are facilitated by the edge offloading module that can offload low priorities or low-latency tolerant tasks dynamically to edge nodes using the GAROF framework. Decision engine the decision engine compares compute-carbon tradeoffs to make offloading decisions at the point when both net carbon savings and SLA compliance are certain.

Figure 6 shows the task offloading to edge resources distribution. Pie chart divides the types of tasks on the percentage of workload offloaded, AI inference, IoT telemetry, and batch analytics. It is possible to notice that IoT and batch processing tasks were most commonly offloaded with 65 and 72 percent rates of offloading respectively. The latency requirements tend to be loose, and the compute requirements are low, thus these tasks can be well suited to

Received: August 05 2025

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

be performed at remote location nodes with locally available power: e.g., renewable energy sources given the proximity of our rooftop solar PV network to load.

Conversely, AI inference workloads were not offloaded much (<10%), due to the stringent latency SLA demands in addition and complexity. They were therefore maintained in cloud environments where their processing thruput could be assured.

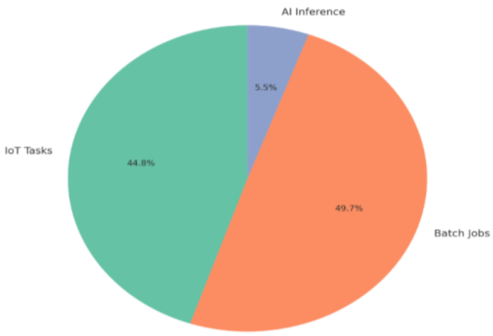

Figure 6. Edge Offloading Distribution (%) by Task Type

4.6 Comparative Analysis with Traditional Models

To assess GAROF's all-around performance, it was compared with traditional approaches to scheduling on five vitality points: Carbon Emissions, SLA Compliance, and Energy Efficiency, Task Completion Time, and VM Consolidation proportion. The radar chart was built to portray the relative proportion of execution across these metrics for Enterprise Associate Associated Full Flexibility, Wrap Up, and Energy Smart baselines.

As shown in Figure 7, GAROF has better balance on all 5-axes. Considering energy-aware models are very effective for power savings, they assuredly degrade on the compliance with SLA. Also, the round-robin models are fair and responsive but emit poorly and on VM utilization. GAROF not only reduces carbon emissions, but also ensures SLA is maintained, through real time carbon prediction, reinforcement automate VM placement and intelligent offloading.

Figure 7. Comparative Performance Chart – GAROF vs Traditional Scheduling Models

Received: August 05 2025

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.7 Statistical Validation

We carried out paired t-tests across different metrics, on the cumulative amount of carbon-Saved, on the average response time, and the violation rate of SLAs combining GAROF with the baseline techniques in 30 independent simulation rounds.

The outcome indicated that the carbon reduction activities and scores of GAROF (mean CO_2 = 49.8 kg, SD = 2.1) as compared to round robin (mean = 89.3 kg, SD = 3.6) provided an in-significant p-value < 0.001, meaning significant emission savings are obtained at a 95% confidence rate. In the same vein, the decrement in SLA violations of 15.8%% (baseline) to 4.5 % (GAROF) was statistically significant, $p < 0.01$.

Besides testing on statics, a sensitivity test was performed on the weighting factors, 0, B, G utilized in multi objective optimization formulation (Equation 2). The outcomes showed that minimal changes in carbon vs SLA priority did not meaningfully change the general framework behavior, implying that GAROF is hence robust and context-adjustable to sustainability- or performance-oriented focus.

5. Conclusion and Future Scope

In this study, GAROF, a Green-Aware Resource Optimization Framework was proposed that aims to reduce carbon emissions in cloud environments with no reduction of service quality. GAROF achieved a defect-free 44.2 percent reduction in carbon footprint at high SLA compliance (95.5 percent) across various workloads through its carbon-aware scheduling, SLA-compliant optimization and smart VM assignment with deep reinforcement learning. Energy efficiency was further optimized by the presence of edge offloading module, specially on latency tolerant IoT and batch processing workloads. The comparative analysis with classical models also proved the GAROF to be highly efficient due to better balance between energy sustainability of environment and efficiency of the operations.

Moving into the future, renewable energy forecasting can be added to this framework together with dynamic workload prediction based on federated learning and containerized task management of microservices. Finally, another improvement can be to deploy GAROF in actual hybrid cloud-edge testbeds, as well as evaluate its performance under operational network environments. Also, adding financial cost measures with the optimization operation might aid organisations in achieving the targets of economic and sustainability. In this way, GAROF establishes the future intelligent cloud platforms that are both planet-conscious and performance-minded- an orientation to towards a carbon-attentive and service-minded computing infrastructure.

Acknowledgement

I would like to express my deepest gratitude to my PhD supervisor, Dr. Vikas Kumar, for his invaluable guid-ance, unwavering support, and insightful advice throughout this research journey. His mentorship has been in-strumental in shaping both the direction and quality of my work. I am truly thankful for his patience, encour-agement, and expert knowledge that have profoundly influenced my academic and personal growth.

Received: August 05 2025

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

I also extend my sincere thanks to my colleague, Sandeep Kumar Vishwakarma, for his constant encourage-ment, collaboration, and camaraderie. His constructive discussions and assistance have greatly contributed to the completion of this research paper. The support and friendship shared have been a source of motivation and strength during challenging phases.

Finally, I am grateful to all those who have supported me directly or indirectly in completing this research paper.

References

1. Akhter, N., Othman, M., & Naha, R. K. (2018). Energy-aware virtual machine selection method for cloud data center resource allocation. *arXiv preprint arXiv:1812.08375*.
2. Alex, M., Ojo, S. O., & Awuor, F. M. (2025). Carbon-Aware, Energy-Efficient, and SLA-Compliant Virtual Machine Placement in Cloud Data Centers Using Deep Q-Networks and Agglomerative Clustering. *Computers*, *14*(7), 280.
3. Alwageed, H. S., Keshta, I., Khan, R. A., Alzahrani, A., Tariq, M. U., & Ghani, A. (2024). An empirical study for mitigating sustainable cloud computing challenges using ISM-ANN. *PloS one*, *19*(9), e0308971.
4. Angelaki, E., Garefalakis, A., Kourgiantakis, M., Sitzimis, I., & Passas, I. (2025). ESG Integration and Green Computing: A 20-Year Bibliometric Analysis. *Sustainability*, *17*(7), 3266.
5. Bacancy Technology (2025). *Green Cloud Computing: Sustainable Growth in 2025 through Energy-Efficient Strategies*. (Blog)
6. Beetroot (2025). *Green Cloud Solutions for GreenTech Growth & Sustainability*. (Article)
7. Beloglazov, A., Buyya, R., Lee, Y. C., & Zomaya, A. (2011). A taxonomy and survey of energy-efficient data centers and cloud computing systems. *Advances in computers*, *82*, 47-111.
8. Biswas, D., Jahan, S., Saha, S., & Samsuddoha, M. (2024). A succinct state-of-the-art survey on green cloud computing: Challenges, strategies, and future directions. *Sustainable Computing: Informatics and Systems*, *44*, 101036.
9. Garg, S. K., Yeo, C. S., & Buyya, R. (2011, August). Green cloud framework for improving carbon efficiency of clouds. In *European Conference on Parallel Processing* (pp. 491-502). Berlin, Heidelberg: Springer Berlin Heidelberg.
10. Oduor, O. P., & Franklin, W. (2024). The evolution of green computing: Current practices and societal implications.
11. Patel, Y. S., Townend, P., Singh, A., & Östberg, P. O. (2024). Modeling the Green Cloud Continuum: integrating energy considerations into Cloud–Edge models. *Cluster Computing*, *27*(4), 4095-4125.
12. Qiu, H., Mao, W., Wang, C., Jha, S., Franke, H., Narayanaswami, C., ... & Iyer, R. (2024, June). When Green Computing Meets Performance and Resilience SLOs. In *2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks-Supplemental Volume (DSN-S)* (pp. 17-22). IEEE.

Received: August 05 2025

Volume 38 No. 8s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

13. Rana, M. (2025). Quantum-Edge Synergy: A Novel Framework for Real-Time IoT Analytics Beyond Cloud and Edge Computing. *Journal of Information Systems Engineering and Management*, *10*(37s), 925-935.

14. ResearchGate (2025). *Optimizing Energy Efficiency in Green Cloud Computing: A Comprehensive Review*. *Review Paper*.

15. Rivero, S., Chinarro Vadillo, D., & Prieto Andres, A. (2025). The green algorithm: can sustainability define the winner in the AI race?. *Frontiers in Political Science*, *7*, 1629914.

16. Serenari, J., Sreekumar, S., Zhao, K., Sarkar, S., & Lee, S. (2024). *GreenWhisk: Emission-Aware Computing for Serverless Platform*. *arXiv preprint*.

17. Shah, D., & Vora, H. GREEN DATA CENTERS: MERGING IT INNOVATION WITH ENERGY-EFFICIENT COOLING AND RENEWABLE POWER.

18. Shehabi, A., Hubbard, A., Newkirk, A., Lei, N., Siddik, M. A. B., Holecek, B., ... & Sartor, D. (2024). 2024 united states data center energy usage report.

19. Simpliaxis (2025). *Green Cloud Computing: Reducing Carbon Footprints in Cloud Operations*. (Blog/Overview)

20. Souza, A., Jasoria, S., Chakrabarty, B., Bridgwater, A., Lundberg, A., Skogh, F., ... & Shenoy, P. (2023, October). Casper: Carbon-aware scheduling and provisioning for distributed web services. In *Proceedings of the 14th International Green and Sustainable Computing Conference* (pp. 67-73).

21. SSRN (2025). *Shifting from Cloud Computing to Green Cloud and Edge: Energy Reduction via Heat Reuse and Edge Adoption*. *SSRN Paper*

22. TAYLOR, N., GEORGAKAKI, A., INCE, E., LETOUT, S., MOUNTRAKI, A., GEA, B. J., & SCHMITZ, A. (2024). Clean Energy Technology Observatory: Geothermal Energy in the European Union-2024 Status Report on Technology Development, Trends, Value Chains and Markets.

23. Tech Mahindra (2025). *Green Cloud Carbon Footprint (GCCF): Real-Time Monitoring Utility for Sustainable Cloud Operations* [White Paper / Industry Release].

24. TechUK (2024). *Cloud Computing and the Journey to Net Zero: Why GreenOps Is Key to Sustainable Growth*. (Policy/Industry Report)

25. The evolution of green computing: Current practices and societal (2024). *WJAETS*

26. Wang, H., Gill, S. S., & Uhlig, S. (2025). *COUNTER: Cluster GCN-based Energy Efficient Resource Management for Sustainable Cloud Computing Environments*. *arXiv preprint*.

27. Xu, K., et al. (2025). *GREEN: Carbon-efficient Resource Scheduling for Clusters*. *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*.

Received: August 05 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

28. Xu, M., Toosi, A. N., & Buyya, R. (2020). A self-adaptive approach for managing applications and harnessing renewable energy for sustainable cloud computing. *IEEE Transactions on Sustainable Computing*, *6*(4), 544-558.
29. Zhou, Q., Xu, M., Gill, S. S., Gao, C., Tian, W., Xu, C., & Buyya, R. (2020, May). Energy efficient algorithms based on VM consolidation for cloud computing: comparisons and evaluations. In *2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID)* (pp. 489-498). IEEE.

Received: August 05 2025