Carbon Emissions from Cloud Storage:

An Empirical Analysis toward Actionable User Practices for Sustainable Storage

GOPIKA PREMSANKAR, Aalto University, Finland
LIN WANG, Paderborn University, Germany
SUZAN BAYHAN, University of Twente, The Netherlands

The growing volume of data stored in cloud storage is a contributor to operational emissions. Reducing emissions from cloud usage is constrained by a limited set of factors. Among them, cloud providers recommend choosing regions with lower carbon intensity (CI) or higher carbon-free energy (CFE) percentages as the primary means of reducing emissions. In this paper, we examine whether following this guidance reliably reduces reported emissions for object storage. We store data for six months across four regions of two major cloud object storage providers (namely, Amazon S3 and Google Cloud Storage), and analyze the scope 2 location-based emissions reported in each provider's carbon-reporting tool. We find that within a single provider, lower CI (or higher CFE) regions do not consistently yield lower reported emissions; in some cases, they result in more emissions than higher-CI regions. This points to an implicit assumption behind the prevailing guidance that electricity consumption is invariant between regions. In practice, regional differences in IT equipment and data center power usage effectiveness can produce electricity variations that offset CI differences. We discuss the implications of this finding and outline future research directions in choosing carbon-efficient regions.

CCS Concepts: • **Information systems** → **Cloud based storage**; • **Social and professional topics** → **Sustainability**.

Additional Key Words and Phrases: Object storage, cloud storage, carbon emissions, region selection, sustainable cloud computing, carbon reporting

1 Introduction

The scale and pace at which digital data is being generated by smart sensors to humans via our daily interaction with cloud services has reached an unprecedented level, with profound implications on the sustainability of digital infrastructures. The data must be moved, stored, computed, communicated, and secured [41]; each step of this data life-cycle results in an environmental cost, e.g., typically represented as carbon-equivalent emissions. While sustainability concerns of the cloud have been in the spotlight recently, they mostly focus on computing workloads in the context of artificial intelligence applications and large language models in particular [13, 30, 39]. However, the environmental cost of retaining data in cloud storage infrastructure remains under-examined [28, 32, 40]. As data volumes continue to grow [29, 44], storage is becoming an increasingly important target for energy and emissions reduction [28, 32, 45, 49].

Object storage services such as Amazon S3 [5], Azure Blob Storage [34], and Google Cloud Storage [23] allow software developers and enterprise customers (henceforth, *users*) to store unstructured data at scale. Such services underpin a broad range of applications, from backups and media delivery to machine learning pipelines [11, 29]. The emissions associated with these services have two components: embodied emissions from manufacturing the underlying storage, networking, and compute devices, and operational emissions from running them. Although embodied emissions from storage have recently received attention [32, 45], operational emissions are also substantial: storage infrastructure accounts for 33% of operational emissions in Azure's cloud [32]. More importantly, operational emissions are the component that users can most directly influence, and are therefore the focus of this article. Users have growing reasons to track and reduce these emissions. Beyond voluntary sustainability commitments, regulatory frameworks such as the GHG Protocol and the Corporate Sustainability Reporting Directive increasingly require organizations to report scope 3 emissions from purchased services, including cloud usage [37, 48]. To address these reporting and tracking requirements, cloud providers offer per-service carbon-reporting tools to their customers [2, 19], reflecting real demand for emissions visibility.

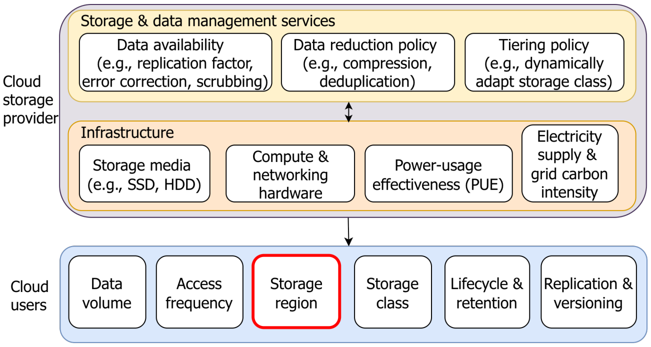


Fig. 1. Factors affecting operational emissions from object storage, separated into cloud provider-controlled (top) and those that users can influence (bottom). This paper focuses on storage region selection, highlighted in red.

Recent work [32, 40] attributes operational emissions from cloud storage primarily to storage media (hard disk drives and solid state drives), replication strategies, and background management tasks, with read/write activity itself contributing minimally (§ 2). As depicted in Figure 1, these dominant drivers represent hardware and system design choices outside of users' control. Among the factors that users can control, region selection emerges as one of the few levers available to reduce emissions. In fact, cloud providers explicitly recommend selecting regions powered by carbon-free energy as a primary means of reducing emissions [20, 38].

To examine how region choice affects storage emissions in practice (§ 3), we stored data for six months in four regions of two major cloud providers, namely Amazon Web Services (AWS) and Google Cloud Platform (GCP), and obtained the scope 2 location-based emissions from each provider's carbon-reporting tool. Location-based

---

Authors' Contact Information: Gopika Premsankar, gopika.premsankar@aalto.fi, Aalto University, Finland; Lin Wang, lin.wang@uni-paderborn.de, Paderborn University, Germany; Suzan Bayhan, suzan.bayhan@utwente.nl, University of Twente, The Netherlands.

emissions are calculated from the carbon intensity (CI) of the local electricity grid where a service operates, independent of any renewable energy purchased by the provider. We observe that within each provider[1], regions with lower CI (or a higher percentage of carbon-free energy) did not reliably yield lower reported emissions. For example, AWS reported that emissions for data stored in Virginia were roughly 56% lower than for the same data volume in Singapore, even though the CI of the two regions was similar during our experiments. A complementary pattern appears within GCP, where regions with markedly different CI nonetheless reported similar emissions – implying that the lower-CI region consumed proportionally more electricity. Together, these observations point to an implicit assumption in the prevailing guidance that storage electricity consumption is invariant across regions. In practice, regional differences in hardware and facility characteristics such as power usage effectiveness can cause electricity variation large enough to dominate CI differences. We provide some directions for future work (§ 4) on using alternate metrics for reliably reducing emissions through region selection, and outline some limitations of our study (§ 5).

2 Background

This section provides a brief overview on object storage, carbon-reporting tools, and carbon emissions from storage services.

2.1 Object storage

All major cloud providers—Amazon Web Services, Google Cloud and Microsoft Azure—provide *object storage* services, namely S3 [5], Cloud Storage [23] and Azure Blob Storage [34]. Modern applications rely on object storage for several data management needs, including data backups, application storage, AI data storage, and content delivery [5, 22, 34]. Object storage services store data as unstructured objects in the provider's data centers, where each object is assigned a unique identifier and made accessible through a network API [10, 22]. Each object is also accompanied by metadata that describes properties of the object including its *storage class*. In Amazon S3 and Google Cloud Storage, the storage class determines, among other properties, performance (i.e., data access latency and throughput), durability (i.e., the probability that data remains intact and uncorrupted over time), availability (i.e., the probability that data is accessible upon request), and replication level used to meet durability and availability guarantees [4, 25]. In contrast, Azure Blob Storage decouples these properties: users select a storage tier [35] and a redundancy level separately, where the redundancy level controls how data is replicated across zones or regions [33]. In all three providers, users can either explicitly select the storage class or rely on provider tools to manage them automatically.

2.2 Carbon reporting from cloud providers

AWS, Google Cloud, and Microsoft Azure each provide dedicated tools for customers to estimate the emissions associated with their cloud usage; the tools are Customer Carbon Footprint Tool [9], Carbon Footprint [21], and Carbon Optimization [36], respectively. All three tools cover object storage services, enabling estimation of emissions from data stored in such services. The tools report emissions according to the Greenhouse Gas Protocol's carbon reporting and accounting standard [48]. As defined in the protocol, scope 1 emissions cover direct on-site emissions from sources such as fuel combustion in backup generators and refrigerant leakage from cooling systems at data centers; scope 2 emissions cover indirect emissions from purchased electricity used to operate data centers; and scope 3 emissions include other indirect emissions such as embodied carbon of data center equipment and buildings, upstream fuel extraction, employee commuting, and business travel [2, 19]. The emissions reports are generated every month and provide a breakdown of emissions per cloud service and per region of a customer's use. The AWS and Google Cloud reporting tools include both location-based and market-based emissions, whereas Azure reports only market-based emissions. Location-based emissions comprise emissions associated with electricity used at specific locations where the cloud service operates, independent of any renewable energy procurement by the provider [3, 19, 31]. Market-based emissions account for contractual renewable energy instruments acquired by the provider (such as Renewable Energy Certificates and Power Purchase Agreements), and can therefore be lower than location-based emissions even when the underlying infrastructure is not directly supplied by low-carbon energy sources [31].

Our analysis focuses primarily on scope 2 location-based emissions that correspond to the electricity used to provide the cloud service. Such emissions are defined as the product of electricity consumed in a region and the carbon intensity (CI) of the local electricity grid. The CI expresses the amount of GHG emissions produced per unit of electricity consumed (gCO2eq/kWh). AWS and GCP reporting tools differ in three key aspects of CI data used to calculate emissions: the source of such data, its temporal granularity, and its life-cycle boundary [2, 19]. GCP sources hourly grid CI from Electricity Maps [16], falling back to annual country-level averages from the IEA where Electricity Maps data is unavailable [19]. AWS derives its grid CI data from region-, state-, or country-level average grid mix data, sourced from the IEA, the EPA's eGRID database, or similar government agencies; the temporal granularity of these factors is not disclosed [2, 3]. Finally, AWS accounts for direct power plant emissions, upstream emissions, and transmission and distribution losses [3], whereas GCP's Electricity Maps CI values (called direct emission factors) cover only electricity generation emissions, excluding other life-cycle stages [19].

2.3 Carbon emissions from storage services

The factors contributing to operational emissions from cloud storage have recently drawn attention [32, 40]. A primary source of these emissions is the energy consumed by storage media. Different storage media can be used for both data and metadata – hard disk drives (HDDs), solid state drives (SSDs) and specialized archival media such as tape, glass or DNA-based storage [32]. To ensure data durability and availability, cloud systems typically store data redundantly, either by replicating it across multiple devices or by using erasure coding, which allows data to be reconstructed if a limited number of servers fail. Both approaches introduce storage overhead beyond

---

[1] Differences in carbon reporting methodologies between GCP and AWS prevent a cross-provider comparison.

what the raw data alone would require. In addition to the storage media, storage servers contain hardware components such as CPUs, RAM, fans, and network interface cards; these components draw power for background tasks and data management operations. Such operations include data scrubbing, metadata management, catalog summarization, deduplication and storage space reclamation [40]. Finally, the data centers housing the storage servers have a power usage effectiveness (PUE) rating, which captures the additional energy overhead for lighting, cooling, and other facility operations. A higher PUE directly increases total electricity consumption and, by extension, overall operational emissions.

An analysis of emissions from Azure's infrastructure for cloud storage [32] found that storage media account for the largest share of operational emissions within storage server racks: SSDs contribute roughly 39% of emissions in SSD-dominant racks, while HDDs contribute approximately 48% in HDD-dominant racks. CPUs are the next largest contributor, responsible for 26–32% of emissions. Notably, the emissions impact of actual data read/write operations is minimal. Since the dominant drivers of operational emissions are tied to hardware and system design choices (namely, storage media type, redundancy schemes and operational management tools used) rather than to workload activity, one remaining lever to reduce emissions is the choice of region where data is stored.

3 Measurement and Analysis

Methodology. We experiment with object storage services of AWS and GCP[2], namely Amazon S3 [5] and Google Cloud Storage [23], and use the following storage classes for each provider: Amazon S3 Standard [4] and standard storage for Google Cloud [25]. Both storage classes are designed for low-latency access and general-purpose storage for cloud applications and websites. Additionally, they provide similar SLOs for availability (99.99% and 99.9% for AWS and GCP respectively) and durability (11 nines for both). Amazon S3 standard storage class stores data in at least three availability zones in the respective region. For Google Cloud standard storage, data is stored in at least two availability zones in the region. Thus in both cases, data remains within the chosen region.

The four chosen regions are Singapore (Asia), Virginia (East USA), Frankfurt (Europe) and São Paulo (South America) as these locations are common between both cloud providers, and represent different geographic continents (Table 1). The carbon intensity (CI) of the electricity grid in these locations exhibit different characteristics. Figure 2 shows the CI obtained from Electricity Maps in 2025[3]: Singapore has the highest average CI (364 gCO2eq), followed by Virginia (340 gCO2eq), Frankfurt (274 gCO2eq) and São Paulo (70 gCO2eq). Furthermore, Singapore demonstrates a low variation in CI over the months, whereas there are more fluctuations in CI values over the year for Virginia and Frankfurt. The coefficient of variation in the monthly CI are low for Singapore (0.33%) and Virginia (8.39%), whereas Frankfurt and São Paulo exhibit a higher variation with 22.35% and 42.77% respectively.

Table 1. Storage locations, their corresponding cloud region identifiers, and the Electricity Maps (EM) zones from which carbon intensity data is sourced.

Location	AWS region [8]	GCP region [24]	EM zone [15]
Singapore	ap-southeast-1	asia-southeast1	SG
Virginia	us-east-1	us-east4	US-MIDA-PJM
Frankfurt	eu-central-1	europe-west3	DE
São Paulo	sa-east-1	southamerica-east1	BR-CS

We store 1 TB of randomly-generated data (comprising 500 GB of binary and 500 GB image data) in the chosen locations between June 2025 to November 2025 for each cloud provider. Throughout the experiment period, the data is retained in storage and remains untouched, i.e., no reads or retrievals are performed. Any storage class optimizations, e.g., automatic storage class transitions triggered by data access patterns, were explicitly disabled to ensure storage conditions remained consistent throughout.

We retrieve scope 2 location-based method (LBM) emissions data from each cloud provider at gram-level precision using their respective data export tools. The emissions data reported by AWS uses carbon footprint tool version 3.0.1 [7], and GCP uses carbon model version 15 [18] for all months except June 2025 (version 13)[4]. A direct comparison of GCP and AWS emissions is not made, as each platform follows a distinct carbon reporting methodology (Section 2.2). The goal is to understand how the choice of storage locations affects carbon emissions reported by AWS and GCP over the experimental period. To better understand this impact, we first approximate the scope-2 LBM emissions attributed to cloud storage at a particular location as follows:

$$C_{\text{storage,loc}} = E_{\text{storage,loc}} \times \text{PUE}_{\text{loc}} \times \text{CI}_{\text{loc}}, \tag{1}$$

where $E_{\text{storage,loc}}$ is electricity consumption attributed to storage at the particular location, PUE is the power usage effectiveness of the data center(s) in the region, and CI_{loc} is the average CI of the grid.

Each location stores the same volume of data and under similar conditions (no access of data and with the same storage class) allowing a fair comparison of scope 2 LBM emissions between different locations for a single cloud provider. We aim to understand the impact of the choice of locations on reported emissions, by first focusing on the carbon intensity (RQ1), followed by an analysis of electricity consumption attributed to storage (RQ2), as discussed next. Note that we use the terms region and location interchangeably throughout this paper.

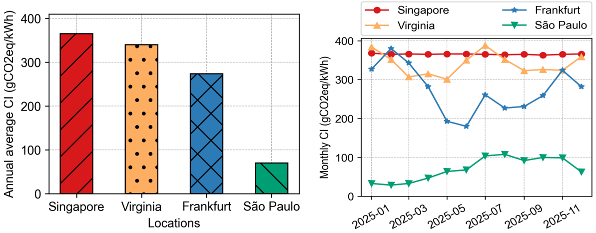
(RQ1) Do locations with lower CI result in lower carbon emissions reported for data stored in those regions? Cloud providers recommend choosing regions with a low CI of electricity (AWS [38]) or regions with high CFE percentages (Google [20]) to lower scope 2 emissions. Given the average CI for the chosen locations in 2025 (Figure 2a) and the CFE percentages reported by Google [20]—São Paulo:

---

[2] We exclude Azure Blob Storage from our experiments, as Azure's Carbon Optimization tool reports only market-based emissions [36] preventing a meaningful analysis of emissions associated with storing data in different regions.

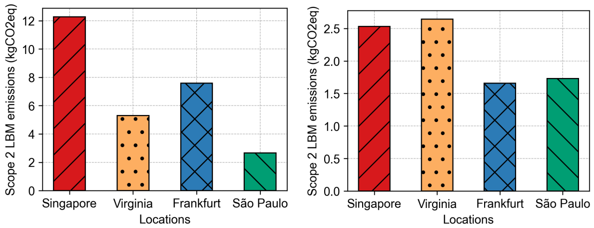
[3] We retrieve the direct emission factors from Electricity Maps that account for the emissions associated with operating the power plant(s) and not those associated with the life-cycle of the power plant(s) (e.g., building the plant, extracting fuel, etc.)

[4] The reported data for June 2025 was not updated version 15 despite running a backfill.



(a) (b)

Fig. 2. (a) Average annual carbon intensity (CI) and (b) monthly CI in 2025 for the chosen locations.



(a) AWS (b) GCP

Fig. 3. Total location-based scope 2 carbon emissions reported for data stored between June to Nov 2025 in (a) AWS and (b) GCP for the chosen locations.

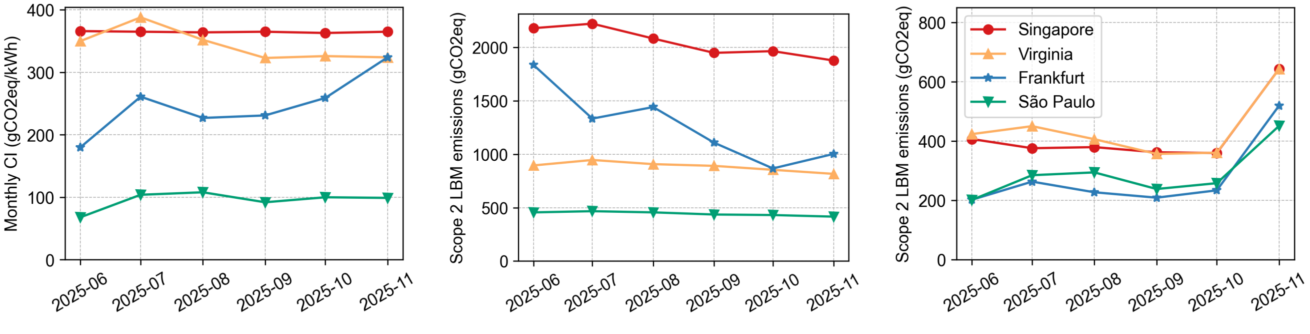
88%, Frankfurt: 68%, Virginia: 62% and Singapore: 4%—we expect lowest scope 2 LBM carbon emissions in São Paulo, followed by Frankfurt, Virginia and Singapore. Note that Google Cloud sources grid-level CI data directly from Electricity Maps to generate its carbon emissions reports and thus we expect this ranking to hold. On the other hand, AWS does not disclose the source of CI data in its official carbon footprint methodology (§ 2.2); however, AWS recommends using publicly available data for CI when choosing regions in a carbon-aware manner [6, 38]. We revisit the sensitivity of our findings to this CI source in § 5.

Figure 3 shows the total scope 2 LBM emissions reported in each location for AWS and GCP. For AWS (Figure 3a), Singapore reports the highest carbon emissions and São Paulo lowest. The annual average carbon intensity is ~80% lower in São Paulo than in Singapore, and the reported scope 2 emissions are similarly ~78% lower. However, interestingly, the scope 2 emissions in Virginia (5.3 kg) are much lower than expected, 56% lower than reported in Singapore, despite the fact that the average CI of the electricity supply in both locations is quite high and similar in magnitude: 365 gCO2eq/kWh in Singapore and 340 gCO2eq/kWh in Virginia. Since we do not have the CI values used by AWS in generating their reports, we cannot explain whether electricity distribution losses or other factors affect this discrepancy (see Section 5 for a sensitivity analysis). In contrast, for GCP (Figure 3b), Virginia and Singapore report the highest emissions with 2.64 kgCO2eq and 2.52 kgCO2eq respectively. The scope 2 emissions in Frankfurt are lowest at 1.65 kgCO2eq followed closely by São Paulo at 1.73 kgCO2eq. The emissions reported in São Paulo are higher than expected, considering the low annual average CI of its electricity supply. The emissions reported are also surprisingly similar in Frankfurt despite the electricity grid in São Paulo (CI: 95 g, CFE: 88%) being more carbon-free than Frankfurt (CI: 247 g, CFE: 68%). Finally, although the CFE of Virginia and Frankfurt are similar (62% and 68% respectively), the scope 2 emissions in Frankfurt are ~38% lower than in Virginia.

Next, we examine whether the reported scope 2 LBM emissions vary over each month of the experiment period. For AWS, it is unclear whether reported emissions are calculated using monthly or annual CI values [3], whereas for GCP, hourly CI values are used to calculate location-based emissions [19]. Thus, we expect the reported scope 2 emissions for GCP to follow the monthly CI values obtained from Electricity Maps, whereas for AWS, we expect minimal variation between the months assuming that electricity used remains constant between regions and months. For AWS (Figure 4b), the emissions reported in São Paulo (456 g in June to 416 g in November) and Virginia (895 g to 816 g) show little variation. There are fluctuations in the emissions reported in Frankfurt (a 27% drop from June to July 2025, followed by an increase again in August 2025). The LBM emissions in Singapore show a declining trend from 2.179 kg in June to 1.876 kg in November, representing a drop of about 14%. For GCP (Figure 4c), the reported emissions follow a trend similar to CI obtained from Electricity Maps between June to October 2025 (Figure 4a). Although the trends are similar, the magnitude of emissions differ from the monthly CI values. Specifically, different from the CI values, emissions reported in São Paulo are higher than in Frankfurt, and emissions reported in Virginia are higher than in Singapore. All regions report a significant increase in scope 2 emissions in November 2025: between 74 to 78% in Singapore, Virginia and São Paulo, and ~120% in Frankfurt compared to October. We cannot explain this increase as the carbon model used by GCP remained the same (version 15) since August 2025, and no issues are reported for storage emissions in the tool's release notes [18].

Takeaway 1. Choosing a region purely based on lower CI or higher CFE (in the case of Google) does not necessarily result in lower reported scope 2 location-based emissions. Moreover, the reported LBM emissions do not always follow the monthly trends observed in the CI values of the respective region, indicating that the electricity used varies between the months and also differs between regions for the same amount of data stored.

(RQ2) How does the electricity used in each location vary? Both AWS and GCP do not directly report per-customer energy consumption in their carbon footprint tools; however, it can be estimated indirectly as follows: $E_{\text{storage,loc}} = C_{\text{reported,loc}}/\text{CI}$. For AWS, energy consumption can be derived by dividing the scope 2 location-based emissions by the corresponding location's annual CI as recommended by their documentation [6]. For GCP, it can similarly be estimated by dividing the reported monthly location-based emissions by the effective CI value for that period. Although GCP uses

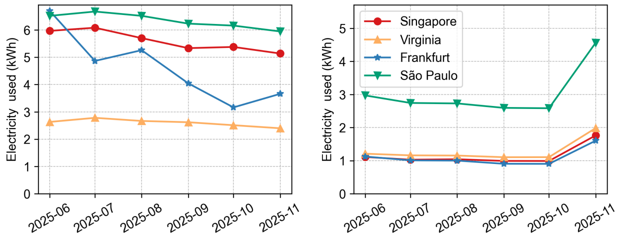


(a) Carbon Intensity

(b) AWS

(c) GCP

Fig. 4. (a) Carbon intensity, and monthly scope 2 emissions for (b) AWS and (c) GCP between June to November 2025.



(a) AWS

(b) GCP

Fig. 5. Estimated electricity used (kWh) per month in (a) AWS and (b) GCP for the chosen locations.

hourly CI values from Electricity Maps [19], hourly carbon emissions reports from Google are not available; thus, we rely on the highest available temporal resolution of monthly data.

We observe that the electricity used is highest in São Paulo for both AWS (Figure 5a) and GCP (Figure 5b). This provides a partial explanation for the high scope 2 LBM emissions reported in this location for both AWS and GCP despite the low CI of the electricity grid. For GCP, the electricity used in the remaining locations is similar in magnitude. For AWS, Virginia uses the lowest amount of electricity in all months, possibly indicating more energy-efficient hardware in use in this location.

As noted in Eq. (1), the PUE also impacts the electricity consumed and thus reported carbon emissions. GCP explicitly states that its scope 2 boundary includes ancillary data center electricity uses such as cooling and lighting [19], whereas AWS does not disclose whether such overhead is included in its scope 2 customer allocation [2]. For AWS, the PUE values[5] are Singapore: 1.32, Virginia: 1.15, Frankfurt: 1.32, and for GCP, Singapore: 1.13, Virginia: 1.09. The relative order of the electricity used reported in Figure 5a does not change after dividing by the PUE values, so we do not report these results. However, the PUE values do indicate that the AWS data center facilities in Virginia are more energy-efficient, which could contribute to the lower reported emissions in this location (if they are indeed included in the emissions reported by AWS). For GCP, there is no public information about PUE values in Frankfurt or São Paulo, so we cannot draw further conclusions about the overhead of the facility itself.

Takeaway 2. Electricity consumption is highest in the region with the lowest CI, indicating that its IT equipment and/or facilities are less energy-efficient than those in other regions. This suggests that additional metrics capturing electricity consumption are needed to meaningfully choose regions that reduce carbon emissions.

4 Discussion

Our empirical analysis shows that lower carbon intensity (CI) or higher carbon-free energy (CFE) percentages alone are insufficient signals to choose regions that minimize storage emissions; additional metrics that capture electricity consumption are needed. As a first step toward closing this gap, we propose a carbon-aware score (S_{loc}) that combines CI with publicly available data: CFE and PUE. The score is defined as:

$$S_{\text{loc}} = \left(1 - \frac{\text{CFE}_{\text{loc}}\%}{100}\right) \times \text{CI}_{\text{loc}} \times \text{PUE}_{\text{loc}}. \tag{2}$$

In Eq. (2), we include CFE as it complements CI by capturing the carbon-free composition of the grid, which CI alone does not indicate. Although the two terms are correlated, they together help to identify regions where the grid is cleaner and has a larger share of carbon-free energy. Additionally, PUE is added as a multiplier to incorporate the power usage overhead (e.g., cooling) in the particular location. Validating this score and its impact on reported emissions requires further empirical study, which we leave to future work. Cloud providers may eventually disclose per-region and user-specific electricity use; private previews of electricity use are available on GCP [19]. However, a proxy-based approach remains useful when disclosure is delayed and future disclosures can serve as ground truth for estimates such as ours.

S_{loc} could underpin an optimization-based method [11, 50] to select carbon-efficient regions. We aim to devise an optimization

---

[5] Most recent data available at the time of writing: 2024 for AWS [1], 2025 for GCP [26]. AWS's PUE was stable over 2022–2024, so 2024 values are used. GCP values are averaged across facilities per location where applicable.

framework to choose regions that reduce emissions while also considering storage pricing. Storage pricing includes actual storage costs, data retrieval and network costs [11, 14]. The inclusion of pricing supports a cost-emissions tradeoff analysis for applications with different SLOs and access patterns. Beyond region selection, other levers in Figure 1 such as storage class can further affect emissions by reducing storage overhead or determining which media are used. Open questions include how frequently data placement plans can be re-evaluated in practice, and whether meaningful carbon savings remain achievable when location choices are constrained by regulatory or compliance policies.

Finally, although we have focused on storage emissions, our findings and proposed proxy-based approach are not storage-specific. The optimization framework can be applied to region selection for compute workloads (e.g., where VMs or compute functions are deployed). However, unlike storage, compute introduces utilization variations that affect electricity use; separate experiments are needed to validate such effects.

5 Limitations

Sensitivity to carbon intensity data. We rely on Electricity Maps [16] for carbon intensity (CI) data for each chosen region, as it provides a consistent methodology and granularity of data for comparison between regions in different countries. However, AWS's cloud carbon footprint tool draws on sources such as the IEA, EPA's eGRID database and other national data providers [2], which may yield different CI values than estimated by Electricity Maps.

In particular, in the US, there are two main ways in which the methodology used by Electricity Maps [17] differs from that of eGRID [46]. First, Electricity Maps uses a flow-tracing algorithm to trace electricity flows between interconnected grids and estimate a consumption-based CI in each zone. This accurately represents the mix of electricity available to consumers on the grid. In contrast, eGRID tracks emissions for power plants at the point of generation, without accounting for power purchases, imports, or exports of electricity. Second, the geographic granularity of reported emissions are different: per zone (which maps to a balancing authority) in Electricity Maps, and per balancing authority, subregion or North American Electricity Reliability Corporation (NERC) region in eGRID. According to eGRID documentation [46], the subregions are eGRID-specific and developed as a compromise between large NERC regions and small balancing authority-based regions.

To assess the robustness of our findings, we additionally compute the estimated electricity consumption for data stored on AWS in Virginia (`us-east-1`) using the emission factor for the SRVC subregion from the eGRID 2023 database [47], assuming the data centers are located in Loudoun County, Virginia. SRVC represents a smaller area (CI: 270 gCO2eq/kWh in 2023[6]) than the US-MIDA-PJM zone (or balancing authority) used by Electricity Maps (CI: 332 gCO2eq/kWh in 2023); this represents an approximate 20% reduction in the CI for Virginia. Among the other three regions, Frankfurt (Germany) and Singapore use country-wide CI (no geographic granularity consideration applies), and thus, Electricity Maps already provides the appropriate value. For Brazil, however, sub-regional CI values are not available from official data sources[7], so the sensitivity check is not extended beyond Virginia. With Virginia's CI adjusted to 270 gCO2eq/kWh (the SRVC value), it remains the region with the lowest estimated electricity consumption in all six months. The gap relative to Frankfurt (the second-lowest region for AWS in terms of estimated electricity consumption) narrows on average by 33% from an average of 2.01 kWh (EM) to 1.35 kWh (eGRID SRVC). The narrowing is non-uniform across months (similar to Figure 5a), with the largest gap in June 2025 (4.06 kWh with EM vs 3.39 kWh with eGRID), and a tie in values in October 2025 (gap of 0.65 kWh with EM vs 0.01 kWh with eGRID). The ordering of regions holds under both CI sources, supporting our key takeaway that understanding regional electricity consumption is necessary to meaningfully choose regions that reduce emissions.

The choice of which geographic granularity and CI data source to use for US regions in our carbon scoring framework remains an open question. An open-source tool for estimating emissions from cloud use, Cloud Carbon Footprint, uses NERC-region granularity for US data centers rather than the smaller eGRID subregion [12]. On the other hand, zone-based CI data from Electricity Maps has been widely used in workload-shifting frameworks in cloud computing [27, 42, 43]. We plan to further investigate and quantify the differences between sources of CI data.

Carbon reporting methodology. The carbon reporting methodologies of cloud providers are continuously updated and refined [7, 18]. For instance, AWS has released four major methodology versions since 2022 (v1.0, v2.0, v3.0.0, v3.0.1), most recently moving from estimated to invoiced utility data for scope 2 emissions [7]. Google Cloud issues semi-annual updates to its carbon model, for example, revising scope 1–3 emission inputs and allocation factors [18]. This implies that the reported emissions may be revised later when a new carbon model is introduced. Our key takeaways were found to hold with different carbon emissions models (v3.0.0 to v3.0.1 for AWS [7] and models 13 to 15[8] for GCP [18]), but future methodology changes to the carbon model may affect reported values in ways our current analysis cannot anticipate. A complementary limitation is that current reporting tools include only emissions and not the underlying electricity consumption. Disclosing regional electricity consumption alongside emissions reports would allow users to distinguish the effect of grid CI from that of IT equipment and facility efficiency, which is necessary for meaningful region selection.

6 Acknowledgments

This research is supported by the Research Council of Finland grant 368053, and the Modular Integrated Sustainable Data Center (MISD) project funded by the Netherlands Enterprise Agency (RVO) under the 8ra initiative, IPCEI-CIS. We thank Electricity Maps for providing data access through their academic program, Sujay Srivastava for setting up the experiments on the cloud, and Mehrdad Aslani for collecting Electricity Maps data.

---

[6] The latest public dataset available at the time of writing.

[7] https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/sirene/dados-e-ferramentas/fatores-de-emissao

[8] Model 15 appears in the data export used in our study but is not documented in the release notes [18] at the time of writing.

References

[1] Amazon Sustainability. AWS Cloud - Power usage effectiveness. https://sustainability.aboutamazon.com/products-services/aws-cloud. Accessed on 20.04.2026.

[2] Amazon Sustainability. AWS Customer Carbon Footprint Methdology, October 2025. https://sustainability.aboutamazon.com/aws-customer-carbon-footprint-methodology.pdf. Accessed 24.03.2026.

[3] Amazon Sustainability. Carbon Methdology, July 2025. https://sustainability.aboutamazon.com/carbon-methodology.pdf. Accessed 23.02.2026.

[4] Amazon Web Services. Amazon S3 Storage Classes. https://aws.amazon.com/s3/storage-classes/. Accessed 17.02.2026.

[5] Amazon Web Services. Amazon Simple Storage Service (Amazon S3). https://docs.aws.amazon.com/s3/. Accessed 17.02.2026.

[6] Amazon Web Services. Calculating your energy usage. https://docs.aws.amazon.com/awsaccountbilling/latest/aboutv2/ccft-energy.html. Accessed on 24.03.2026.

[7] Amazon Web Services. Customer carbon footprint tool release notes. https://docs.aws.amazon.com/ccft/latest/releasenotes/what-is-ccftrn.html. Accessed on 24.06.2026.

[8] Amazon Web Services. S3 Pricing. https://aws.amazon.com/s3/pricing/. Accessed 17.02.2026.

[9] Amazon Web Services. Understanding your customer carbon footprint tool. https://docs.aws.amazon.com/awsaccountbilling/latest/aboutv2/what-is-ccft.html. Accessed on 24.03.2026.

[10] Amazon Web Services. What's the Difference Between Block, Object, and File Storage? https://aws.amazon.com/compare/the-difference-between-block-file-object-storage/. Accessed 16.03.2026.

[11] Bang, T., Douglas, C., Crooks, N., and Hellerstein, J. M. SkyPIE: A fast & accurate oracle for object placement. *Proceedings of the ACM on Management of Data 2*, 1 (2024), 1–27.

[12] Cloud Carbon Footprint. Methodology. https://www.cloudcarbonfootprint.org/docs/methodology/#carbon-estimates-co2e, 2025. Accessed on 27.06.2026.

[13] Dodge, J., Prewitt, T., Tachet des Combes, R., Odmark, E., Schwartz, R., Strubell, E., Luccioni, A. S., Smith, N. A., DeCario, N., and Buchanan, W. Measuring the carbon intensity of AI in cloud instances. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (2022), pp. 1877–1894.

[14] Durner, D., Leis, V., and Neumann, T. Exploiting cloud object storage for high-performance analytics. *Proceedings of the VLDB Endowment 16*, 11 (2023), 2769–2782.

[15] Electricity Maps. Data center zones. https://github.com/electricitymaps/electricitymaps-contrib/blob/master/config/data_centers/data_centers.json. Accessed on 18.05.2026.

[16] Electricity Maps. Signals: Carbon Intensity. https://app.electricitymaps.com/developer-hub/api/signals#carbon-intensity. Accessed on 18.05.2026.

[17] Electricity Maps. Methodology for the Electricity Mix and Derived Signals. Whitepaper, https://www.electricitymaps.com/resources/publications/whitepaper-methodology-for-the-electricity-mix-and-derived-signals, 2025. Accessed on 29.06.2026.

[18] Google Cloud. Carbon Footprint release notes. https://docs.cloud.google.com/carbon-footprint/docs/release-notes. Accessed on 24.06.2026.

[19] Google Cloud. Carbon Footprint reporting methodology (Last updated 2026-02-19 UTC). https://docs.cloud.google.com/carbon-footprint/docs/methodology. Accessed 23.02.2026.

[20] Google Cloud. Carbon free energy for google cloud regions. https://cloud.google.com/sustainability/region-carbon. Accessed on 20.04.2026.

[21] Google Cloud. View carbon footprint data. https://cloud.google.com/carbon-footprint/docs/view-carbon-data. Accessed on 24.03.2026.

[22] Google Cloud. What is Object storage? https://cloud.google.com/learn/what-is-object-storage?hl=en. Accessed 07.05.2026.

[23] Google Cloud Platform. Cloud Storage. https://cloud.google.com/storage?hl=en. Accessed 17.02.2026.

[24] Google Cloud Platform. Cloud storage pricing. https://cloud.google.com/storage/pricing?hl=en. Accessed 17.02.2026.

[25] Google Cloud Platform. Storage Class. https://docs.cloud.google.com/storage/docs/storage-classes. Accessed 17.02.2026.

[26] Google Data Centers. Power usage effectiveness. https://datacenters.google/efficiency/. Accessed on 20.04.2026.

[27] Hanafy, W. A., Liang, Q., Bashir, N., Souza, A., Irwin, D., and Shenoy, P. Going green for less green: Optimizing the cost of reducing cloud carbon emissions. In *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 3* (2024), pp. 479–496.

[28] Kopczyk, D., and Chandra, A. Re-evaluating storage carbon emissions in machine learning workloads. *ACM SIGENERGY Energy Informatics Review 5*, 2 (2025), 111–117.

[29] Li, J., Cao, B., Jian, J., Li, C., Han, S., Wang, Y., Wu, Y., Chen, K., Yin, Z., Chen, Q., et al. Mantle: Efficient hierarchical metadata management for cloud object storage services. In *Proceedings of the ACM SIGOPS 31st Symposium on Operating Systems Principles* (2025), pp. 928–943.

[30] Lyu, J., Wang, J., Frost, K., Zhang, C., Irvene, C., Choukse, E., Fonseca, R., Bianchini, R., Kazhamiaka, F., and Berger, D. S. Myths and misconceptions around reducing carbon embedded in cloud platforms. In *Proceedings of the 2nd Workshop on Sustainable Computer Systems* (2023), HotCarbon '23.

[31] Maji, D., Bashir, N., Irwin, D., Shenoy, P., and Sitaraman, R. K. The green mirage: Impact of location-and market-based carbon intensity estimation on carbon optimization efficacy. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems* (2024), pp. 256–267.

[32] McAllister, S., Kazhamiaka, F., Berger, D. S., Fonseca, R., Frost, K., Ogus, A., Sah, M., Bianchini, R., Amvrosiadis, G., Beckmann, N., et al. A call for research on storage emissions. *ACM SIGENERGY Energy Informatics Review 4*, 5 (2024), 67–75.

[33] Microsoft Azure. Azure Storage Redundancy. https://learn.microsoft.com/en-us/azure/storage/common/storage-redundancy. Accessed 17.02.2026.

[34] Microsoft Azure. Introduction to Azure Blob Storage. https://learn.microsoft.com/en-us/azure/storage/blobs/storage-blobs-introduction. Accessed 17.02.2026.

[35] Microsoft Azure. Introduction to Azure Blob Storage. https://learn.microsoft.com/en-us/azure/storage/blobs/storage-blobs-introduction. Accessed 16.03.2026.

[36] Microsoft Azure. View and analyze emission data with Carbon optimization in Azure. https://learn.microsoft.com/en-us/azure/carbon-optimization/quickstart-view-analyze-emission-data. Accessed 24.03.2026.

[37] Microsoft Sustainability. The Corporate Sustainability Reporting Directive (CSRD). https://www.microsoft.com/en-us/sustainability/learning-center/corporate-sustainability-reporting-directive. Accessed on 07.05.2026.

[38] Mokhtari, S., Dua, I., and Khanal, A. How to select a Region for your workload based on sustainability goals. https://aws.amazon.com/blogs/architecture/how-to-select-a-region-for-your-workload-based-on-sustainability-goals/. Accessed 24.03.2026.

[39] Patterson, D., Gonzalez, J., Hölzle, U., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D. R., Texier, M., and Dean, J. The carbon footprint of machine learning training will plateau, then shrink. *Computer 55*, 7 (2022), 18–28.

[40] Rao, V., and Chien, A. A. Understanding the operational carbon footprint of storage reliability and management. *ACM SIGENERGY Energy Informatics Review 4*, 5 (2024), 180–187.

[41] Semiconductor Research Corporation, and Semiconductor Industry Association. Decadal Plan for Semiconductors: Full Report. https://www.src.org/about/decadal-plan/decadal-plan-full-report.pdf, Jan. 2021. Accessed 29.06.2026.

[42] Souza, A., Jasoria, S., Chakrabarty, B., Bridgwater, A., Lundberg, A., Skogh, F., Ali-Eldin, A., Irwin, D., and Shenoy, P. CASPER: Carbon-aware scheduling and provisioning for distributed web services. In *Proceedings of the 14th International Green and Sustainable Computing Conference* (2023), pp. 67–73.

[43] Sukprasert, T., Souza, A., Bashir, N., Irwin, D., and Shenoy, P. On the limitations of carbon-aware temporal and spatial workload shifting in the cloud. In *Proceedings of the Nineteenth European Conference on Computer Systems* (2024), pp. 924–941.

[44] Sébastien Stormacq (AWS Blog). Twenty years of Amazon S3 and building what's next. https://aws.amazon.com/blogs/aws/twenty-years-of-amazon-s3-and-building-whats-next/. Accessed 16.03.2026.

[45] Tannu, S., and Nair, P. J. The dirty secret of SSDs: Embodied carbon. *ACM SIGENERGY Energy Informatics Review 3*, 3 (2023), 4–9.

[46] U.S. Environmental Protection Agency. The emissions generation resource integrated database: eGRID technical guide with year 2023 data. https://www.epa.gov/system/files/documents/2025-01/egrid2023_technical_guide.pdf, Jan. 2025. Accessed on 29.06.2026.

[47] U.S. Environmental Protection Agency (EPA). Emissions Generation Resource Integrated Database (eGRID). https://www.epa.gov/egrid. Accessed on 27.06.2026.

[48] World Resources Institute and World Business Council for Sustainable Development. The greenhouse gas protocol: A corporate accounting and reporting standard. https://ghgprotocol.org/sites/default/files/standards/ghg-protocol-revised.pdf. Accessed on February 16, 2024.

[49] Xie, D., Stavrinos, T., Zhu, K., Peter, S., Kasikci, B., and Anderson, T. Can storage devices be power adaptive? In *Proceedings of the 16th ACM Workshop on Hot Topics in Storage and File Systems* (2024), pp. 47–54.

[50] Zhang, Q., Li, S., Li, Z., Xing, Y., Yang, Z., and Dai, Y. CHARM: A cost-efficient multi-cloud data hosting scheme with high availability. *IEEE Transactions on Cloud computing 3*, 3 (2015), 372–386.