FootPrinter: Quantifying Data Center Carbon Footprint

Dante Niewenhuis
d.niewenhuis@vu.nl
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

Sacheendra Talluri
s.talluri@vu.nl
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

Alexandru Iosup
a.iosup@vu.nl
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

Tiziano De Matteis
t.de.matteis@vu.nl
Vrije Universiteit Amsterdam
Amsterdam, Netherlands

ABSTRACT

Data centers have become an increasingly significant contributor to the global carbon footprint. In 2021, the global data center industry was responsible for around 1% of the worldwide greenhouse gas emissions. With more resource-intensive workloads, such as Large Language Models, gaining popularity, this percentage is expected to increase further. Therefore, it is crucial for data center service providers to become aware of and accountable for the sustainability impact of their design and operational choices. However, reducing the carbon footprint of data centers has been a challenging process due to the lack of comprehensive metrics, carbon-aware design tools, and guidelines for carbon-aware optimization. In this work, we propose FootPrinter, a first-of-its-kind tool that supports data center designers and operators in assessing the environmental impact of their data center. FootPrinter uses coarse-grained operational data, grid energy mix information, and discrete event simulation to determine the data center's operational carbon footprint and evaluate the impact of infrastructural or operational changes. Foot-Printer can simulate days of operations of a regional data center on a commodity laptop in a few seconds, returning the estimated footprint with marginal error. By making this project open source, we hope to engage the community in the development of methodologies and tools for systematically assessing and exploring the sustainability of data centers.

CCS CONCEPTS

• **Hardware → Impact on the environment**; **Renewable energy**; • **Computing methodologies → Agent / discrete models**.

KEYWORDS

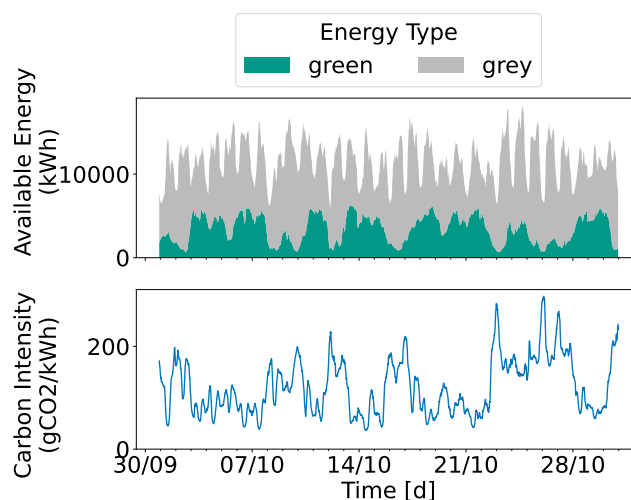Carbon Footprint, Carbon Emission, Data Center, Simulation

Figure 1: The energy mix and carbon intensity of the energy grid in the Netherlands during the month of October 2023 from the ENTSO-E Transparency Platform[1]. The top graph shows the energy mix during the month into green and non-green energy. The bottom graph shows the resulting carbon intensity of the grid.

1 INTRODUCTION

Climate change is a significant social challenge today, affecting various aspects of our daily lives [34]. In 2015, world leaders reached a breakthrough with the Paris Agreement, which aims "to limit the temperature increase to 1.5°C above pre-industrial levels." [31]. To achieve this goal, the European Union (EU) has established a 55% reduction in greenhouse gas emissions by 2030 for all its member states [12].

Data centers significantly contribute to the global carbon footprint [13], accounting for 1% of global greenhouse gas emissions in 2021 [24]. As a result of demands from governments and users [30], and financial considerations, data center operators have been working to reduce their carbon footprint. The recent energy price crisis and sustainability efforts (e.g., through green bond emissions) have made operational expenses a primary cost factor for data centers [36].

---

[1]https://transparency.entsoe.eu/dashboard/show

So far, data center designers and operators have been focusing mainly on improving their power efficiency. Data centers already use more than 1% of the global energy consumption [28], and some estimate this will rise to as much as 8% in 2030 [3]. Despite the improvements in energy efficiency, aggregate energy usage has increased in the last 15 years [7]. Moreover, efficiency improvements have slowed down significantly in recent years [38]. A bigger problem, however, is that optimizing energy does not directly reduce the carbon footprint. The carbon emitted by a data center depends not just on the amount of energy used but also on the type. For instance, Figure 1 shows how the grid energy mix and its carbon intensity can change continuously over time.

Reducing the carbon footprint of a data center is a challenging process. There is no consensus on measuring carbon emissions [19], and *there is a lack of carbon-aware design tools and guidelines for carbon-aware optimization* [18]. These challenges have resulted in many companies still relying on rule-of-thumb reasoning [4], which has led to carbon-inefficient practices, such as significant overprovisioning of resources [20]. Improving the carbon footprint has been even more difficult for smaller data centers [23], which often lack insight into tenant workloads and their provided energy mix. Besides the technical challenges, significant costs are involved. Data centers operate on a large scale, making experimentation costly and time-consuming. Making uninformed decisions can also have a significant economic impact. Data center projects have been stopped in countries like the Netherlands based on vague, qualitative statements about their potential climate impact[2].

In this work, we make three contributions:

(1) *We discuss what information data center operators need to quantify and optimize their operational carbon footprint.* Measuring a data center's energy consumption requires that operators invest in hardware and software tools. Attributing this to individual applications is complex and requires even more tooling. Therefore, we suggest using coarse-grained execution metrics, as a convenient yet effective way of assessing the data center's energy consumption.

(2) *We introduce FootPrinter*[3], a data center discrete simulator based on the OpenDC[4] framework. FootPrinter takes as input the hardware configuration of a data center and workload traces and uses simulation to determine the corresponding energy footprint. The energy profile is combined with the energy mix of the location region to calculate the operational carbon footprint of the data center when it runs the given workload.

(3) *We validate FootPrinter* using a wall-socket energy trace from SURF, the Dutch national supercomputing center, showing that the simulated data center has the same energy usage as the data center running the same workload in the real world.

With FootPrinter, we aim to contribute with a tool for data center designers and operators to reason about the environmental impact and associated costs of their infrastructures and plan for appropriate measures to improve their sustainability.

---

[2]https://www.datacenterknowledge.com/meta-facebook/scorned-meta-data-center-holland-met-all-environmental-standards
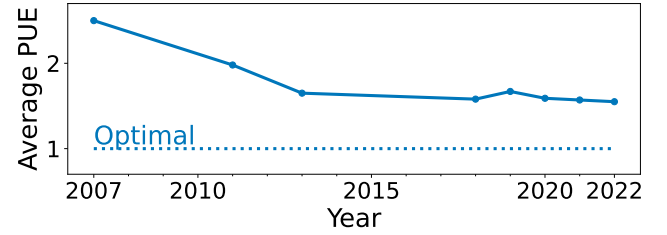[3]https://github.com/atlarge-research/FootPrinter
[4]https://opendc.org/

Figure 2: The average Power Usage Effectiveness (PUE) of 669 data centers from 2007 to 2022 [14]. The dotted line shows the optimal value of 1.0.

2 BACKGROUND

The carbon footprint of a data center is characterized by two types of carbon emission: the *embodied* carbon footprint and the *operational* carbon footprint. Embodied carbon is the carbon emitted from manufacturing and production. Operational carbon footprint is the CO2 emissions caused by energy usage during operations. In this work, we focus on reducing the operational carbon footprint of data centers.

2.1 Power Usage Effectiveness

In recent years, much focus has been placed on improving the efficiency of data centers. The most commonly used metric for energy efficiency is Power Usage Effectiveness (PUE). PUE is calculated using Equation 1:

$$PUE = \frac{E_T}{E_{IT}} \tag{1}$$

In which E_t and E_{IT} denote the total energy used by the data center and the energy used by the IT components of the data center. In an optimal data center, no energy is required for redundant tasks, using all energy for the IT equipment doing the computation. This results in a PUE of 1.0. However, while many data centers have been able to optimize their PUE, with for instance Google getting close to 1.1 [5], the aggregate energy consumption of data centers has still increased over the last 15 years [7]. One reason for this is the rebound effect, which states that if the energy required to perform a task (and thus its price) decreases, the number of tasks performed will increase [40]. Another reason is that the rate of improvement of PUE has slowed down significantly in recent years. Figure 2 shows the average PUE of 669 data centers during the period of 2007 to 2022 [14]. While great improvements were made between 2007 and 2013 (from 2.5 to 1.6), recent years did not bring any more significant improvements, with the lowest average PUE of 1.55 being achieved in 2022.

We suggest two possible reasons for this slowdown of improvement. First, as the PUE is already highly optimized, it is becoming increasingly difficult to optimize it further. Second, the shift to hyperscale data centers had a significant impact on the average PUE. Because this shift is nearly finished, it is unclear where significant improvements will come from [7].

---

[5]https://www.google.com/about/datacenters/efficiency/

2.2 Carbon Intensity

While PUE is a good metric to determine infrastructure energy efficiency, it is not taking everything into account. PUE does not consider the energy efficiency of applications and workloads [43]. PUE also completely ignores the type of energy used. The source of energy can have an enormous impact on the carbon emitted. In some cases, energy sourced from renewable sources, such as wind or solar, can emit up to 20x less CO2 compared to traditional energy sources, such as coal [22]. The Carbon Intensity of an energy source defines the amount of carbon emitted per unit of energy used. Many data centers do, however, not use energy from a single energy source, but get their energy from the grid. Energy provided by the grid is often gathered from many different energy sources with different carbon intensities. The carbon intensity of the grid is calculated by aggregating the different energy sources in Equation 2:

$$CI_g = \sum_{s \in S} CI_s \frac{E_s}{E_g} \qquad (2)$$

In which CI_s is the carbon intensity of energy source s, E_s/E_g is the share of energy that s contributes to the grid, and S is the set of all available energy sources. Green energy is primarily gained from natural phenomena, such as wind or sunlight. This results in a continuously changing mix of available energy (see Figure 1). During this time, the ratio of green and non-green energy varied significantly. As a result, the carbon intensity of the grid also changes significantly over time (100 to 400 gCO2/kWh). This means that to minimize the carbon footprint of a data center, not only the amount of energy used is important, but also when this energy is used.

2.3 Operational Footprint

The operational carbon footprint is characterized by the carbon emitted when the system is running. The operational carbon footprint can be calculated by combining the carbon intensity of the data center CI_d(gCO2/kWh) and the operational energy of the data center E_{op} (kWh) as defined in Equation 3:

$$C_{op} = CI_d E_{op} \qquad (3)$$

We assume that the carbon intensity of the energy used by a data center is proportional to the carbon intensity of the grid ($CI_d = CI_g$). Some data centers have special energy contracts providing them direct access to specific types of energy[6]. However, these data centers still have to resort to using energy from the grid, when not enough energy is available [1]. In this work, we focus on the carbon footprint of a data center. However, several other metrics for data center sustainability exist [35].

2.4 Simulation

FootPrinter uses discrete-time simulation to estimate the carbon footprint of a data center in a time and energy-aware manner. Using simulation for data center research is not new. Simulators such as Grid/CloudSim [9], SimGrid [10], and iCanCloud [32] have demonstrated the ability to simulate complex operations at cluster and data

---

[6]https://www.datacenterdynamics.com/en/news/meta-signs-renewable-energy-deal-in-arizona-with-orsted/
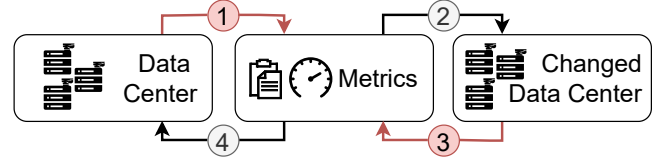

Figure 3: A method of determining the impact of making changes to a data center. 1) Determine initial performance, 2) Change data center infrastructure and/or operations based on metrics and goals, 3) Determine the performance of the changed data center, and when requirements are met 4) Consolidate the changes in the data center. The red lines highlight the challenging steps.

center levels. In this work, we use OpenDC, a trace-based discrete data center simulation framework [29]. OpenDC uses real-world workload traces to drive simulation. A workload trace describes when jobs get submitted and their computational requirements. More advanced workload traces also define their computational demand over time. OpenDC replays the workload on a specified data center and allows users to explore "what-if" scenarios. FootPrinter uses these features and extends them to compute the energy required to run the workload on a user-specified data center and derive its corresponding carbon footprint.

3 PROBLEM STATEMENT

Reducing the carbon footprint of a data center is a challenging task. Due to a lack of carbon-aware tooling, data center designers and operators need to decide between different options with limited insight into their effects [18]. Therefore, determining how to change the data center infrastructure and operations is often a process of trial and error, in which new experiments are executed based on the results of previous experimentation until the imposed requirements are met (see Figure 3). Using a similar approach when working with data centers is ineffective due to the time, energy, and monetary costs involved. Collecting energy metrics on the level of individual servers or server components requires significant investments in hardware and software, such as power meters for measurement and software to process data and storage. The more detailed the information required, the more power meters, storage, and computing are needed. Furthermore, the energy usage of a data center can assist the operator in identifying problems and areas of improvement, such as idle VMs, or inefficient resource management. It does, however, not provide enough information to determine the effect of changes made to address the identified problems. This insight is vital to determine where to invest the available budget and engineering time. Small real-world experiments followed by analysis are often used to quantify efficacy (see Figure 3). However, this feedback loop might be slow because of the long execution time of experiments, or even unfeasible due to economic reasons.

FootPrinter enables a convenient approach to analyzing and optimizing the carbon footprint of a data center. Through the use of discrete simulations, it allows the user to consider several scenarios, keeping costs and operational impact low. FootPrinters' stakeholders are *data center designers*, who architect the data center infrastructure, and *operators* who run the data center operations.

We present three use cases that showcase the difficulties faced by these stakeholders. In the remainder of the paper, we elaborate on how FootPrinter can currently be utilized to tackle the first two, while *UC-Hardware* is utilized to discuss how FootPrinter's capabilities can be expanded.

UC-Footprint *Operational carbon footprint:* Knowing the operational carbon footprint of a data center is an essential part of evaluating its effectiveness. Determining the operational carbon footprint requires knowledge about both the energy usage and the carbon intensity of the used energy sources. As discussed previously, properly monitoring energy usage requires specialized hardware and software.

UC-Location *Selecting a location:* The location of a data center can have a big impact on its operational carbon footprint due to the available energy mix. Choosing the right location is challenging for both data center designers and operators. Designers need to decide where to build new data centers. Operators must decide where to execute submitted jobs when accessing multiple data centers. In both cases, insight into the effect of location on the operational carbon footprint is required.

UC-Hardware *Selecting hardware upgrades:* A designer responsible for upgrading a data center hardware has to make choices within a limited budget. With a wide range of hardware options, deciding what to install can be difficult. To make informed decisions, designers must understand the impact of hardware changes.

4 FOOTPRINTER

We propose FootPrinter, an *energy-aware* discrete data center simulator based on the OpenDC framework. FootPrinter takes as input the hardware configuration of a data center and workload traces, and uses simulation to determine the energy footprint. The energy footprint is combined with the energy mix of the data center's region to determine the operational carbon footprint of the data center during the execution of the given workloads.

Figure 4 shows the architecture of FootPrinter and illustrates how it could be used by data center operators. Using the FootPrinter starts at the real data center ❶. Over time, different workloads ① are submitted to the servers ②, and the operations software ③ is used to decide when, where, and how these workloads are executed. The activity of the data center is monitored during operations and recorded. To use FootPrinter, three pieces of information are required as input data ❷:

④ *Workload traces* that describe when jobs are submitted and hardware requirements of each job. The trace also describes the computational demand over time. FootPrinter is designed to work with traces of any sample frequency. However, providing traces with higher frequency will result in more precise results.

⑤ *Hardware and environment specifications* that describe the hardware used by the datacenter. To determine the carbon footprint, it is also important to define where a data center is located.

⑥ *Operational techniques* that define how and when jobs are run. Important factors are the scheduling and resource allocation policies.
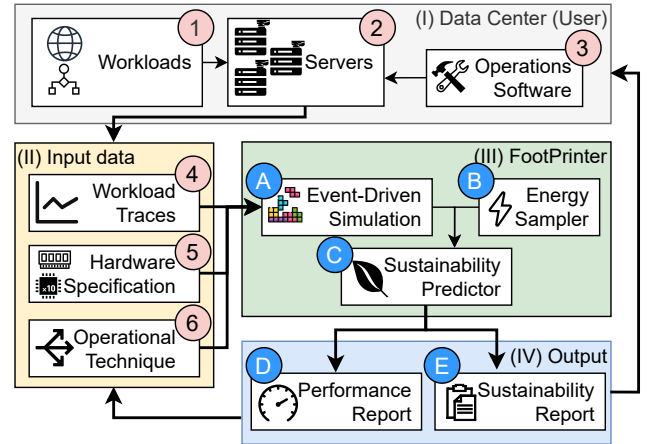

Figure 4: A diagram of the FootPrinter functionality. Four areas are defined: The Data Center which is controlled by the user ❶, the input data gathered from the data center ❷, The FootPrinter which simulated the input data ❸, and the output ❹.

The input data is sent to the FootPrinter to replay. The FootPrinter architecture ❸ consist of the following components:

Ⓐ The *Event-Driven Simulator* replays the given workload traces on the given data center configuration. During the run, the simulator is sampled for performance metrics and energy usage. The frequency of sampling can be chosen to best fit the current experiment. Higher frequency will result in more precision at a cost of increasing the simulation time.

Ⓑ The *Energy Sampler* determines the carbon intensity of the grid while the simulation is run. Whenever the event-driven simulator is sampled, the carbon intensity of the grid is needed. The energy mix of the grid is sampled using the Python API[7] of the ENTSO-E Transparency Platform[8].

Ⓒ The *Sustainability Predictor* aggregates the results of the simulation into sustainability metrics, such as the total carbon emitted and the carbon emission over time. These metrics can be used to determine the operational carbon footprint of the data center during the workload.

FootPrinter generates two types of output ❹. First, the Performance Report Ⓓ shows the performance of the data center during the provided workload. Examples of performance metrics are the time of completion, or average CPU utilization. Next to the performance of the data center, a sustainability report Ⓔ is made. Examples of sustainability metrics are the energy usage, or the carbon emitted. Designing data centers is a difficult process, in which often improvements in sustainability are connected to decreases in performance. FootPrinter reports both sides to provide the data center operators with a complete insight.

---

[7]https://github.com/EnergieID/entsoe-py
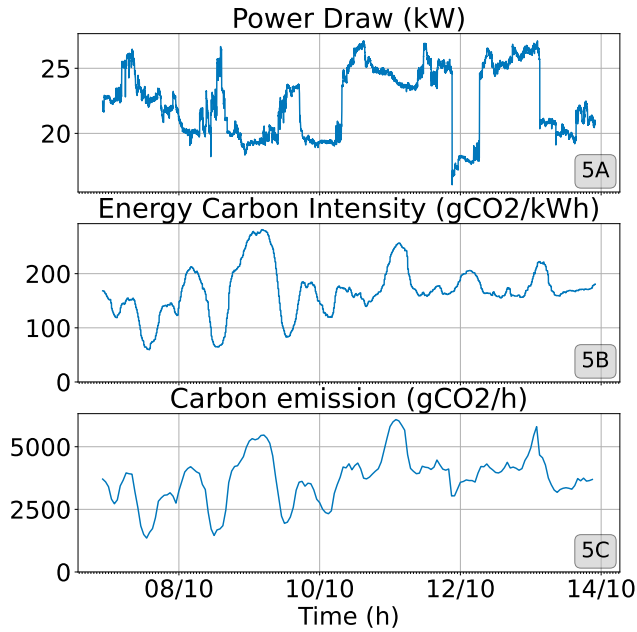[8]https://transparency.entsoe.eu/dashboard/show

Figure 5: The Carbon emission of a workload over time, determined using FootPrinter. Graph 5A shows the power draw over time. Graph 5B shows the carbon intensity of the grid during the workload. Graph 5C combines the two other graphs, showing the carbon emission during the workload.

5 EXPERIMENTS

This section demonstrates how FootPrinter can be used in different use cases from section 3. The accuracy of FootPrinter is validated by comparing it to an empirically measured energy usage trace.

5.1 Operational Carbon Footprint

We use FootPrinter to determine the operational carbon footprint of a data center (*UC-Footprint*). To illustrate the process, we simulate a workload trace gathered from the SURF Lisa[9] cluster, an HPC data center in the Netherlands. The workload consists of 7,850 jobs executed over seven days. The duration of the jobs ranges from less than an hour to several days. The CPU demand is sampled at a 30-second interval for each job in the trace. The workload is run on a data center comprising 277 physical machines. FootPrinter replays this trace on a mid-range laptop (Intel Core I7-8750H Processor[10]) in 10 seconds. This allows for rapid experimentation mentioned in section 3.

Figure 5 depicts the process of determining operational carbon footprint using FootPrinter. Figure 5A shows the simulator-determined power draw of the data center during the workload, sampled every 30 seconds. The graph depicts the power draw of the entire data center. However, FootPrinter can also provide similar graphs for specific nodes or jobs. The aggregate power draw varies in the range of 16 to 28 kW. The energy usage at a sample can be

---

[9]https://www.surf.nl/en/lisa-computing-cluster-extra-computing-power-for-research

[10]https://ark.intel.com/content/www/us/en/ark/products/134906/intel-core-i7-8750h-processor-9m-cache-up-to-4-10-ghz.html
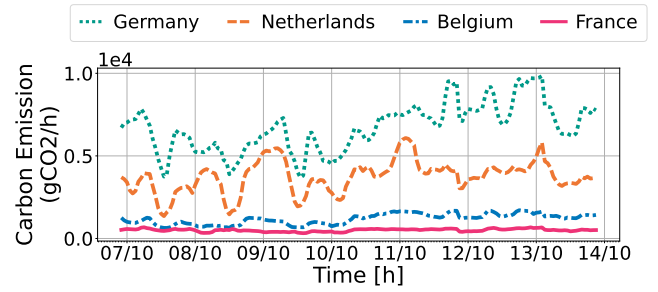

Figure 6: The carbon emission during the same workload simulated executed on the same data center located in four different locations.

determined by multiplying the power draw and the time since the previous sample. Figure 5B depicts the carbon intensity of the grid sampled from ENTSO-E. The difference in carbon intensity during the chosen period is significant, ranging between 100 and 400 gCO2/h. Figure 5C depicts the carbon emission during the workload. Carbon emission at a sample can be calculated by multiplying the energy usage at a sample with the carbon intensity. The carbon emission is primarily influenced by the carbon intensity, due to the much higher variability in the carbon intensity compared to the power draw. This demonstrates the importance of measuring the carbon footprint directly, instead of just energy usage.

5.2 Selecting location

FootPrinter can be used to compare the impact of building or expanding the data center infrastructure in multiple locations (*UC-Location*). Figure 6 depicts the effect of the data center location on its carbon emission. The workload introduced in subsection 5.1 is replayed on the same data center in different locations. France and Belgium perform much better than the Netherlands and Germany. This is because France and Belgium source around half of their energy from nuclear power plants emitting almost no carbon. The Netherlands and Germany, however, rely more on energy sources such as coal, which is very carbon intensive.

5.3 Validation

To quantify the accuracy of our simulator, we compare the power draw of a workload determined by the simulator, to the real-world power draw of the same workload. We use the same workload as used in subsection 5.1. Figure 7 shows the simulated power draw determined by FootPrinter and the real-world power draw. We determine the accuracy of the estimation using three different metrics. Each metric is calculated separately for all points, the points in which FootPrinter *underestimates* (underestimation error), and the points in which FootPrinter *overestimates* the power draw (overestimation error).

The first metric of estimation accuracy is the Mean Absolute Percentage Error (MAPE), a popular measure of the accuracy of forecasting methods. MAPE is commonly used to determine forecast accuracy because of its intuitive interpretation in terms of relative error [16]. MAPE is a relative error measure that uses absolute values to keep the positive and negative errors from canceling one another out [33] and is calculated using Equation 4:
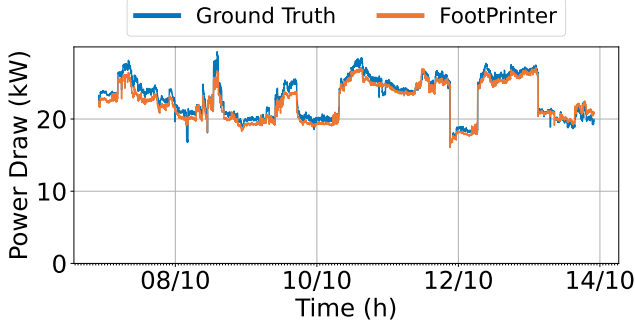
Figure 7: The power draw of a data center during a given workload simulated by the FootPrinter tool compared to the actual Power Draw of the data center.

$$MAPE[\%] = \frac{1}{n} \sum_{t=0}^{n} |\frac{P_t - P'_t}{P_t}| \times 100 \qquad (4)$$

In which P_t and P'_t are the actual and simulated power draw at sample t and n is the number of samples. Comparing FootPrinter to the ground truth results in a MAPE total error of 3.15%, underestimation error of 3.19%, and overestimation error of 2.93%.

The second metric of prediction accuracy is the Normalized Absolute Differences (NAD). NAD describes the total error of the prediction divided by the sum of the ground truth and is calculated using Equation 5:

$$NAD[\%] = \frac{\sum_{t=0}^{n} |P_t - P'_t|}{\sum_{t=0}^{n} P_t} \times 100 \qquad (5)$$

In which P_t and P'_t are the actual and simulated power draw at sample t and n is the number of samples. Comparing FootPrinter to the ground truth results in a NAD total error of 3.17%, underestimation error of 3.22%, and overestimation error of 2.83%.

Finally, we look at the distribution of the errors. Figure 8 shows the percentage of time points with an error less than the given threshold. Over half of the points have an error less than 3%, and 93% an error less than 6%.

6 RELATED WORK

The research community has built many high-quality simulators that provide a rich set of features to build upon [6, 8]. CloudSim [9] is the closest to OpenDC, the simulator used in this paper. CloudSim offers a number of single-feature simulators such as CloudAnalyst [39], iFogSim [21], and WorkflowSim [11]. However, the single focus of these simulators makes it challenging to combine without extensive engineering. In contrast, OpenDC is a flexible general purpose simulator that supports various different features. Building FootPrinter op OpenDC guaranties support for a varied applictions.

Extending simulators to estimate the carbon footprint of a data centers is not a novel idea. In their paper from 2022, Song et al. discuss over 100 papers working on data center carbon footprint in the last ten years, in which 75% used simulators in their experiment [37]. Most of the works discussed extend third party simulators to estimate carbon footprint. The Most popular simulator for this purpose is Cloudsim [15, 25, 42], but other simulators, such as SimGrid [5],
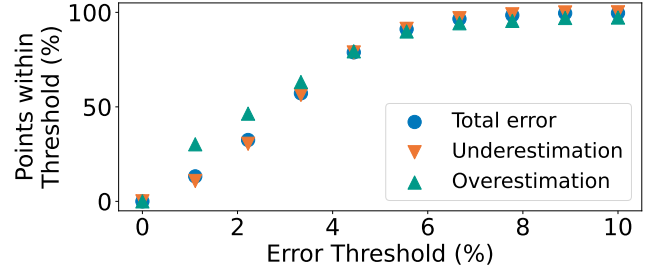

Figure 8: The distribution of the error of samples. Each point represents the percentage of samples with an error less than the specified threshold.

EcoMultiCloud [2], and iFogSim [41], are also used. Because of the single-feature nature of the simulators used, most of these tools are very specialized for their specific purpose. In contrast, Foot-Printer is more general purpose. Another distinction is that many tools focus on single green energy sources, such as solar [26, 27], or wind [17]. FootPrinter is not dependent on any specific type of energy source.

7 CONCLUSION

This work introduces FootPrinter, a first-of-its-kind tool that uses simulation to determine the operational carbon footprint of a data center. FootPrinter replays workload traces to determine the energy usage and carbon emission during the workload execution. FootPrinter is designed to work with any trace granularity to make it accessible to all data center operators. We have validated Foot-Printer by comparing the simulated energy usage to the real-world energy usage. FootPrinter can simulate energy usage with a Mean Average Percentage Error of less than 3.15%.

We discussed three use cases highlighting challenges for data center designers and operators who want to evaluate the sustainability impact of their actions. In this paper, we showed how FootPrinter can be used to determine operational carbon footprint and compare data center locations. FootPrinter is an open-source tool and can be extended to support more use cases and provide more insights. We are already actively working on supporting hardware upgrades and their impact on performance and carbon footprint. Additionally, we are working on adding support for more elements that can influence the energy usage of a data center, such as temperature and humidity. Finally, while FootPrinter currently quantifies the operational carbon emissions of a data center, we believe it can be easily extended to also incorporate embodied carbon emissions.

REFERENCES
[1] Bilge Acun, Benjamin Lee, Fiodar Kazhamiaka, Kiwan Maeng, Udit Gupta, Manoj Chakkaravarthy, David Brooks, and Carole-Jean Wu. 2023. Carbon Explorer: A

Holistic Framework for Designing Carbon Aware Datacenters. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2* (Vancouver, BC, Canada) *(ASPLOS 2023)*. Association for Computing Machinery, New York, NY, USA, 118–132. https://doi.org/10.1145/3575693.3575754

[2] Abada Ahmed and Marc St-Hilaire. 2018. *Renewable Energy Curtailment via Incentivized Inter-datacenter Workload Migration*. 143–157. https://doi.org/10.1007/978-3-319-94295-7_10

[3] Anders S. G. Andrae and Tomas Edler. 2015. On Global Electricity Usage of Communication Technology: Trends to 2030. *Challenges* 6, 1 (2015), 117–157. https://doi.org/10.3390/challe6010117

[4] Georgios Andreadis, Fabian Mastenbroek, Vincent van Beek, and Alexandru Iosup. 2021. Capelin: Data-Driven Capacity Procurement for Cloud Datacenters using Portfolios of Scenarios – Extended Technical Report. arXiv:2103.02060 [cs.DC]

[5] Gaganjeet Singh Aujla and Neeraj Kumar. 2018. SDN-based energy management scheme for sustainability of data centers: An analysis on renewable energy sources and electric vehicles participation. *J. Parallel Distrib. Comput.* 117, C (jul 2018), 228–245. https://doi.org/10.1016/j.jpdc.2017.07.002

[6] Ilyas Bambrik. 2020. A Survey on Cloud Computing Simulation and Modeling. *SN Computer Science* 1, 5 (2020), 249.

[7] Noman Bashir, David Irwin, Prashant Shenoy, and Abel Souza. 2022. Sustainable Computing – Without the Hot Air. arXiv:2207.00081 [cs.CY]

[8] James Byrne, Sergej Svorobej, Konstantinos M. Giannoutakis, Dimitrios Tzovaras, Peter J. Byrne, Per-Olov Östberg, Anna Gourinovitch, and Theo Lynn. 2017. A Review of Cloud Computing Simulation Platforms and Related Environments. In *CLOSER 2017 - Proceedings of the 7th International Conference on Cloud Computing and Services Science, Porto, Portugal, April 24-26, 2017*, Donald Ferguson, Víctor Méndez Muñoz, Jorge S. Cardoso, Markus Helfert, and Claus Pahl (Eds.). SciTePress, 651–663.

[9] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, César A. F. De Rose, and Rajkumar Buyya. 2011. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience* 41, 1 (2011), 23–50.

[10] Henri Casanova, Arnaud Giersch, Arnaud Legrand, Martin Quinson, and Frédéric Suter. 2014. Versatile, scalable, and accurate simulation of distributed applications and platforms. *J. Parallel and Distrib. Comput.* 74, 10 (2014), 2899–2917. https://doi.org/10.1016/j.jpdc.2014.06.008

[11] Weiwei Chen and Ewa Deelman. 2012. WorkflowSim: A toolkit for simulating scientific workflows in distributed environments. In *8th IEEE International Conference on E-Science, e-Science 2012, Chicago, IL, USA, October 8-12, 2012*. IEEE Computer Society, 1–8.

[12] European Commission. 2020. Stepping up Europe's 2030 climate ambition Investing in a climate-neutral future for the benefit of our people. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52020DC0562

[13] ACM Technology Council. 2021. Computing and climate change. https://doi.org/doi/pdf/10.1145/3483410

[14] J. Davis, D. Bizo, A. Lawrence, O. Rogers, M. Smolaks, L. Simon, and D. Donnellan. 2022. Uptime Institute Global Data Center Survey 2022. Uptime Institute.

[15] Inès De Courchelle, Tom Guérout, Georges Da Costa, Thierry Monteil, and Yann Labit. 2019. Green energy efficient scheduling management. *Simulation Modelling Practice and Theory* 93 (2019), 208–232. https://doi.org/10.1016/j.simpat.2018.09.011 Modeling and Simulation of Cloud Computing and Big Data.

[16] Arnaud de Myttenaere, Boris Golden, Bénédicte Le Grand, and Fabrice Rossi. 2016. Mean Absolute Percentage Error for regression models. *Neurocomputing* 192 (2016), 38–48. https://doi.org/10.1016/j.neucom.2015.12.114 Advances in artificial neural networks, machine learning and computational intelligence.

[17] Xiaowen Dong, Taisir El-Gorashi, and Jaafar Elmirghani. 2011. Green IP over WDM Networks: Solar and Wind Renewable Sources and Data Centres. 1–6. https://doi.org/10.1109/GLOCOM.2011.6134175

[18] Mariam Elgamal, Doug Carmean, Elnaz Ansari, Okay Zed, Ramesh Peri, Srilatha Manne, Udit Gupta, Gu-Yeon Wei, David Brooks, Gage Hills, and Carole-Jean Wu. 2023. Carbon-Efficient Design Optimization for Computing Systems. In *Proceedings of the 2nd Workshop on Sustainable Computer Systems* (Boston, MA, USA) *(HotCarbon '23)*. Association for Computing Machinery, New York, NY, USA, Article 16, 7 pages. https://doi.org/10.1145/3604930.3605712

[19] Anshul Gandhi, Dongyoon Lee, Zhenhua Liu, Shuai Mu, Erez Zadok, Kanad Ghose, Kartik Gopalan, Yu David Liu, Syed Rafiul Hussain, and Patrick Mcdaniel. 2023. Metrics for Sustainability in Data Centers. *SIGENERGY Energy Inform. Rev.* 3, 3 (oct 2023), 40–46. https://doi.org/10.1145/3630614.3630622

[20] James Glanz. 2012. *Power, Pollution and the Internet*. https://www.nytimes.com/2012/09/23/technology/data-centers-waste-vast-amounts-of-energy-belying-industry-image.html

[21] Harshit Gupta, Amir Vahid Dastjerdi, Soumya K. Ghosh, and Rajkumar Buyya. 2017. iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things, Edge and Fog computing environments. *Software: Practice and Experience* 47, 9 (2017), 1275–1296.

[22] Udit Gupta, Mariam Elgamal, Gage Hills, Gu-Yeon Wei, Hsien-Hsin S. Lee, David Brooks, and Carole-Jean Wu. 2022. ACT: designing sustainable computer systems with an architectural carbon modeling tool. In *Proceedings of the 49th Annual International Symposium on Computer Architecture* (New York, New York) *(ISCA '22)*. Association for Computing Machinery, New York, NY, USA, 784–799. https://doi.org/10.1145/3470496.3527408

[23] IEA. 2017. *Digitalisation and Energy*. https://www.iea.org/reports/digitalisation-and-energy

[24] IEA. 2022. *Data Centres and Data Transmission Networks*.

[25] Chao Li, Rui Wang, Depei Qian, and Tao Li. 2016. Managing Server Clusters on Renewable Energy Mix. *ACM Trans. Auton. Adapt. Syst.* 11, 1, Article 1 (feb 2016), 24 pages. https://doi.org/10.1145/2845085

[26] Chao Li, Wangyuan Zhang, Chang-Burm Cho, and Tao Li. 2011. SolarCore: Solar energy driven multi-core architecture power management. In *2011 IEEE 17th International Symposium on High Performance Computer Architecture*. 205–216. https://doi.org/10.1109/HPCA.2011.5749729

[27] Longjun Liu, Hongbin Sun, Yang Hu, Jingmin Xin, Nanning Zheng, and Tao Li. 2015. Leveraging distributed UPS energy for managing solar energy powered data centers. (02 2015). https://doi.org/10.1109/IGCC.2014.7039150

[28] Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. 2020. Big Tech Gets Caught Up in Europe's Energy Politics. https://www-science-org.vu-nl.idm.oclc.org/doi/10.1126/science.aba3758

[29] Fabian Mastenbroek, Georgios Andreadis, Soufiane Jounaid, Wenchen Lai, Jacob Burley, Jaro Bosch, Erwin van Eyk, Laurens Versluis, Vincent van Beek, and Alexandru Iosup. 2021. OpenDC 2.0: Convenient Modeling and Simulation of Emerging Technologies in Cloud Datacenters. In *2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*. 455–464. https://doi.org/10.1109/CCGrid51090.2021.00055

[30] Rich Miller. 2021. The Sustainability Imperative: Green Data Centers and Our Cloudy Future. https://www.datacenterfrontier.com/special-reports/article/11428454/the-sustainability-imperative-green-data-centers-and-our-cloudy-future

[31] United Nations. 2015. *The Paris Agreement*. https://www.un.org/en/climatechange/paris-agreement

[32] Alberto Núñez, Jose Vázquez-Poletti, Agustín Caminero, Gabriel Castañé, Jesus Carretero, and Ignacio Llorente. 2012. ICanCloud: A Flexible and Scalable Cloud Infrastructure Simulator. *Journal of Grid Computing* 10 (03 2012), 185–209. https://doi.org/10.1007/s10723-012-9208-5

[33] Oracle. 2014. *Working with Planning: MAPE*. https://docs.oracle.com/en/cloud/saas/planning-budgeting-cloud/pfusu/insights_metrics_MAPE.html#GUID-C33B0F01-83E9-468B-B96C-413A12882334

[34] Interfovernmental panel on climate change. 2022. *IPCC PRESS RELEASE*. https://www.ipcc.ch/report/ar6/wg2/downloads/press/IPCC_AR6_WGII_PressRelease-English.pdf

[35] V. Reddy, B. Setz, G. K. Rao, G. Gangadharan, and M. Aiello. 2017. Metrics for Sustainable Data Centers. *IEEE Transactions on Sustainable Computing* 2, 03 (jul 2017), 290–303. https://doi.org/10.1109/TSUSC.2017.2701883

[36] April Roach and Ewa Krukowska. 2022. Recalibrating global data center energy-use estimates. https://www.bnnbloomberg.ca/big-tech-gets-caught-up-in-europe-s-energy-politics-1.1782670

[37] Jie Song, Peimeng Zhu, Yanfeng Zhang, and Ge Yu. 2022. Versatility or validity: A comprehensive review on simulation of Datacenters powered by Renewable Energy mix. *Future Generation Computer Systems* 136 (11 2022), 326–341. https://doi.org/10.1016/j.future.2022.06.008

[38] Petroc Taylor. 2023. *Data center average annual power usage effectiveness (PUE) worldwide 2007-2023*. https://www.statista.com/statistics/1229367/data-center-average-annual-pue-worldwide/

[39] Bhathiya Wickremasinghe, Rodrigo N. Calheiros, and Rajkumar Buyya. 2010. CloudAnalyst: A CloudSim-Based Visual Modeller for Analysing Cloud Computing Environments and Applications. In *24th IEEE International Conference on Advanced Information Networking and Applications, AINA 2010, Perth, Australia, 20-13 April 2010*. IEEE Computer Society, 446–452.

[40] Jackson Woodruff, David Schall, Michael F.P. O'Boyle, and Christopher Woodruff. 2023. When Does Saving Power Save the Planet?. In *Proceedings of the 2nd Workshop on Sustainable Computer Systems* (Boston, MA, USA) *(HotCarbon '23)*. Association for Computing Machinery, New York, NY, USA, Article 20, 6 pages. https://doi.org/10.1145/3604930.3605719

[41] Chenhan Xu, Kun Wang, Peng Li, Rui Xia, Song Guo, and Minyi Guo. 2020. Renewable Energy-Aware Big Data Analytics in Geo-Distributed Data Centers with Reinforcement Learning. *IEEE Transactions on Network Science and Engineering* 7, 1 (2020), 205–215. https://doi.org/10.1109/TNSE.2018.2813333

[42] Minxian Xu and Rajkumar Buyya. 2020. Managing renewable energy and carbon footprint in multi-cloud computing environments. *J. Parallel and Distrib. Comput.* 135 (2020), 191–202. https://doi.org/10.1016/j.jpdc.2019.09.015

[43] Runlin Zhou, Yingjie Shi, and Chunge Zhu. 2013. AxPUE: Application level metrics for power usage effectiveness in data centers. In *2013 IEEE International Conference on Big Data*. 110–117. https://doi.org/10.1109/BigData.2013.6691705