

Modelling Scenarios for Carbon-aware Geographic Load Shifting of Compute Workloads

Wim Vanderbauwhede

Abstract—We present an analytical model to evaluate the reductions in emissions resulting from geographic load shifting. This model is optimistic as it ignores issues of grid capacity, demand and curtailment. In other words, real-world reductions will be smaller than the estimates. However, even with these assumptions, the presented scenarios show that the realistic reductions from carbon-aware geographic load shifting are small, of the order of 5%. This is not enough to compensate the growth in emissions from global data centre expansion.

Index Terms—Data centres, Geographic Load Shifting, Carbon Emissions

I. CONTEXT: EMISSIONS FROM AI DATA CENTRES

Greenhouse gas emissions from ICT have been estimated at 4% [1], [2] and rising steeply. The estimates are from before the rise in popularity of generative AI, which is currently the main driver for the growth in emissions from ICT. The global consulting firm McKinsey projects a global growth in AI data centres until 2030 of between 19% and 27% annually [3]. A growth in global data centre capacity of 22% until 2040 as per McKinsey's medium-range scenario corresponds to a growth of $20\times$ in data centre energy consumption, which is far larger than the projected $4\times$ in [2].

In 2023, the global electricity demand of data centres was estimated at 55 GW [3], [4]; that means the projected total power capacity for 2030 would be 1,100 GW and the energy consumption 9,640 TWh/y. For reference, the current world electricity production is about 30,000 TWh/y [5]. If this trend persists, AI data centres would become not only the main contributor to emissions from ICT but be responsible for a large fraction of the global carbon budget (see e.g. [6] for more details). It is therefore imperative to reduce emissions arising from data centre construction and activity.

In the face of exponential growth, any approach which reduces yearly data centre emissions by a fixed amount is essentially compensating for a certain time of continued growth. If the approach reduces emissions 10% per year, this reduction will be undone by the medium-range growth in less than a year. Fig. 1 shows the time it takes for the growth to undo the reduction in emissions (or put positively, the number of years of growth compensated).

What the figure shows is that for reductions of $< 20\%$, the growth will undo the reduction in less than a year. If we can achieve reductions of $> 50\%$, we can compensate several

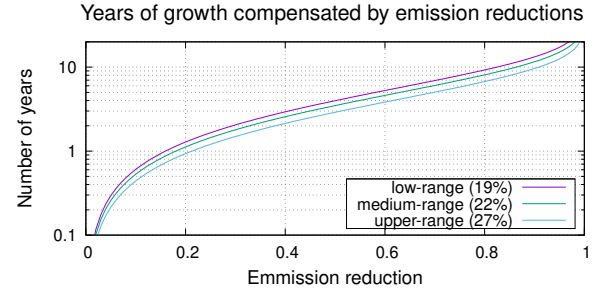


Fig. 1. Years of growth as per the McKinsey scenarios compensated by emission reductions

years, and that may be enough because economic exponential growth bubbles rarely last more than a decade.

II. RELATED WORK: CARBON-AWARE COMPUTING

Carbon-aware computing [7] refers to running compute workloads when and where the electricity grid is being powered by renewable energy. This requires shifting work in time and/or space. Time-based shifting is discussed e.g. in [8]; geographic shifting (also known as spatial load shifting or load migration) and spatio-temporal shifting are discussed in detail in this section. The purpose is to reduce emissions from data centre based compute activities in recognition of the steep growth in demand. In this paper we focus on geographic load shifting.

Lindberg *et al.* [9] present a model where data centres shift load independently of the independent grid system operators. They claim that, relative to previous models for data centre load shifting, their model improves accuracy and include more realistic assumptions regarding the operation of both data centres and the electricity market. They claim further that their approach is able to “reduce total system carbon emissions by roughly 33%”, for $0.01 \leq \epsilon \leq 0.2$ where ϵ is the fraction of the data centre capacity (load) that can be shifted per 5-minute time step. This does not mean only that percentage of load gets effectively shifted: from the graphs in the paper we see that the average percentage of load shifted is about 17%.

One major issue with the paper is that it does not list values of the electricity generation carbon intensity (CI) used in the model, nor the locations of the generators or data centres. It is therefore not possible to directly validate the results. The actual data are contained in the RTS-GMLC model used in the study, but this model [10] comes with a caveat: “While projecting the RTS model onto the southwestern United States enables the use of geospatially and temporally coincident

TBD

Manuscript received Month XX, Year; revised Month XX, Year.

0000-0000/00\$00.00 enables the use of geospatially and temporally coincident

weather-driven data, it would be inappropriate to use the projected RTS model to provide insights into the real-world power system in this location.” In other words, the results do not relate directly to real-world scenarios. Furthermore, the model does not consider data centre embodied carbon nor idle power consumption or compute load specific shifting constraints.

Acun *et al.* [11] presents the Carbon Explorer framework to analyse the potential for emission reduction in data centres, including geographic load shifting, and use this framework to balance trade-offs between operational and embodied based on geographic location and workload. The fundamental assumption of their paper is that a data centre operator such as Meta can buy renewable energy when available to achieve carbon-free operation. This ignores the issue that renewable generation is overall still a minority fraction of total generation, and therefore, if the data centre operator buys all renewable energy it is offloading its emissions to other users of the grid.

Lin *et al.* [12] propose a virtual queue algorithm for job scheduling across data centres in different locations, in other words a spatio-temporal model. The model uses electricity price of the and considers both wind and solar. There is no direct information on the CI values used; the paper states that simulations are based on real-world data from several states of Australia, and shows traces which indicate the CI lies between 300 and 415 gCO_2e/kWh . The data centre model does take into account idle power but not embodied carbon. It provides no details for active or idle server power consumption or PUE. It shows reductions of between 21% and 80% for a number of different scenarios, but these are reductions when compared to a grid without renewables, rather than reductions from spatio-temporal load shifting. The model does not consider compute load specific shifting constraints.

Bian *et al.* [13] use carbon-aware geographic load shifting for federated learning; the paper shows that they can achieve higher test accuracy for the same carbon budget as existing approaches. It does not show what reductions in emissions would be if the test accuracy was kept constant. The model does not consider data centre embodied carbon nor idle power consumption.

Riepin *et al.* [14] present an approach to spatio-temporal load shifting with the goal of achieving 24/7 Carbon-Free Energy (CFE) matching, which means that the data centre operator can claim to use 100% renewable energy. The paper is thorough in its analysis of the available wind and solar capacity and the correlation between sites. The final result is expressed in terms of cost decrease. It is not possible to tell how much the actual reductions in energy consumption are as a result of the proposed approach. The model does not consider data centre embodied carbon nor idle power consumption.

Coskun *et al.* [15] propose an alternative to workload migration: *Conductor* is a framework that coordinates the participation of multiple data centres in demand response, increasing their resilience to operate under power constraints without requiring any inter-data-centre workload migration. The aim is to provide greater flexibility in power consumption while improving the ability of the data centres to meet QoS

targets. We can view this as virtual geographic load shifting.

Finally, Zheng *et al.* [16] focuses on load migration from the Pennsylvania-New Jersey-Maryland interconnection (PMJ) geographical area to California (CAISO), specifically to make use of curtailed renewable energy, and find that geographic load shifting could absorb up to 62% of the total curtailment in the CAISO region. If no additional capacity is installed, shifting work to increase the load in the California data centres (for a total of 280 MW, 28 data centres of 10 MW) from 50% to 65% would result in a reduction 115 $ktCO_2e/y$. Calculating the baseline from their data gives about 720 $ktCO_2e/y$, assuming that the capacity and load of the data centres in PMJ and CAISO are the same. However, based on the discussion of additional capacity needed, this is not the case: they explain that in the ideal case (all load shifted from PMJ to CAISO), an additional 780 MW at 65% load would be needed. This means that the amount of load shifted is 550 MW. If we assume that is the entire load of all PMJ data centres, the baseline is 2.02 $MtCO_2e/y$ (if we assume it is less, then this baseline goes up). In other words, the best-case reduction without building additional data centres, 239 $ktCO_2e/y$, amounts to 12% assuming all work can be shifted. In practice it will be a little less (10%) because the paper uses a electricity generation carbon intensity of 421 $kgCO_2e/MWh$, whereas the most recent figure is 369 $kgCO_2e/MWh$.

This is the only work on geographic load shifting that takes into account the embodied carbon of the data centre. Their estimate is based on a figure from [17]: “The study shows that non-operational emissions account for 6.5% of the total life-cycle climate change impacts of a data centre.” However, that paper expresses all values in Eco-indicator points, and there is not sufficient information in the paper itself to separate the emissions from the other factors contributing to the points. Furthermore, they considerably underestimate the embodied carbon of storage (see e.g. [18] for a discussion on SSD and HDD embodied carbon). Therefore, it provides only an approximate estimate for the emissions. Based on this estimate, Zheng *et al.* conclude that “The embodied GHG emissions of a U.S. data centre therefore amounted to 0.20–0.18 $KtCO_2e/MW$ critical power per year during 2016–2019”. This should be contrasted with e.g. the more recent work by [19] which uses 5 $ktCO_2e/MW$.

III. RESEARCH QUESTION: WHAT IS THE POTENTIAL OF GEOGRAPHIC LOAD SHIFTING

There is an important question largely overlooked in the current research: even if there are no constraints on the grid, what is the expected level of emission reductions from deploying these geographic load shifting approaches, assuming realistic parameters for the data centres?

To answer this question, we present a model to assess the extent of reduction in emissions that can result from geographic load shifting and apply it to a number of scenarios, both for commercial AI data centres and HPC centres. This model improves on the state of the art in the following ways:

- Both embodied CO_2e emissions and operational CO_2e emissions are calculated based on up-to-date, detailed and rigorous Life Cycle Analysis models.
- We combine the calculated emission values into a novel, simple but expressive analytical model that allows to calculate the reduction in emissions arising from geographic load shifting.
- We consider not only the maximum theoretically possible gains but focus on realistic scenarios.

IV. MODEL FOR EMISSIONS REDUCTION OF GEOGRAPHIC LOAD SHIFTING

Our aim is to model the reduction in overall data centre emissions resulting from shifting work from a data centre in a high-emission region to one in a low-emission region. The model is not grid-aware: it assumes that there is sufficient power in the target region to power the data centre at full capacity. This is not a fundamental restriction as the available power is only one of the factors that determine the available capacity at the target site. As reduced power availability means less compute capacity, we can model power availability through the load of the data centre; reduced availability of renewables can be modelled through the CI.

The model is conceptually simple: there are only two different types of sites; we compute the embodied carbon emissions and operational emissions for each site (end-of-life emissions can be ignored, see e.g. [17], [19]) assuming that the sites are identical in specification (so we have twin data centres). The embodied carbon and the operational emissions are calculated using our LCA model [20]. Embodied carbon depends on the manufacturing process; operational emissions are based on the power consumption, and the electricity generation carbon intensity of the regions.

We further specify the load of the data centre and the idle power consumption factor, so that we can estimate the power consumption for less than full load. With these parameters, we can calculate the baseline emissions. We also specify which fraction of the work can be moved for which fraction of the time and finally add a factor for the overhead caused by moving the work.

A. Model parameters

The parameters for the baseline model are:

- n_{nodes} number of nodes in data centre. A "node" corresponds to a compute server in a rack, taking into account the overhead of the rack and top-of-rack networking equipment.
- C_{emb} embodied carbon of the data centre ($kgCO_2e/y$, expressed per node). This includes IT equipment, building, cooling and power equipment. It does not include the embodied carbon for the infrastructure outside of the data centre (roads, power lines, communication cables,...).
- C_{hi} operational carbon emissions of the high-emission site ($kgCO_2e/y$, per node, includes PUE). This

is based on the year-averaged electricity use and electricity generation carbon intensity.

- c_{lo} operational carbon emissions of the low-emission site ($kgCO_2e/y$, per node, includes PUE)
- λ_{hi} load of high-emission site, $0 \leq \lambda_{hi} \leq 1$. By load we mean the load on the server, i.e. the proportion of the time that it is not idle.
- λ_{lo} load of low-emission site, $0 \leq \lambda_{lo} \leq 1$
- γ idle power consumption as fraction of active power consumption, $0 \leq \gamma \leq 1$ (assumed to be the same for all sites)

The additional parameters for the geographic load shifting model are:

- α fraction of workload that can be moved, $0 \leq \alpha \leq 1$
- β fraction of the time that this workload can be moved, $0 \leq \beta \leq 1$
- η overhead factor for emissions incurred because of geographic, $0 \leq \eta \ll 1$ load shifting (network emissions, copying of data, ...), on a per-node basis

To simplify the equations, we further define

$$C_{\{emb,hi,lo\}} = c_{\{emb,hi,lo\}} \cdot n_{nodes} \quad (1)$$

B. Embodied emissions

The embodied carbon of the data centre C_{emb} is calculated separately using our model as discussed in [20]. The model is implemented in the functional programming language Haskell, the source code is available at [21]. We provide here a brief overview of the methodology and sources.

To compute the embodied carbon of server manufacturing, we have re-implemented the model by Boavizta [22]. This is a very comprehensive model but as it was published in 2021 we have updated the estimates for various parameters. The model includes estimates for the chips and contributions from packaging and assembly, power supplies, motherboard, server enclosure and rack enclosure.

For the various chips used in the server (CPU, GPU, RAM, SSD) we use the ACT methodology (Architectural Carbon modelling Tool) [23]. This methodology uses the electricity consumption of the manufacturing process, the embodied carbon for the materials, and the greenhouse gas potential for the various gases used in production. These parameters are combined with the die size to obtain an estimate for the embodied carbon of the chip. We have updated the parameters that were included without reference in the ACT paper using data from [18], [24] and [25]. We also extended the model to include non-integrated GPUs.

Our model also takes into account the embodied emissions of the facility (the data centre building and furnishings, power and cooling infrastructure) and of the networking equipment, based on estimates worked out in [19].

C. Emissions from use

The emissions from use (c_{hi} and c_{lo}) are on a per-node bases but do include the overhead of the network infrastructure and of the cooling etc., as expressed through the *PUE*.

We define E_{node} as the year-average energy consumption of a node, i.e. the average power consumption P_{node} multiplied by the time interval. If the node power consumption is expressed in W and the CI is expressed in $kgCO_2e/kWh$, then we have:

$$E_{node} = P_{node} \times 24 \times 365 / 1000 \text{ (kWh)} \quad (2)$$

The emissions from use $C_{\{hi,lo\}}$ are then given by:

$$C_{\{hi,lo\}} = E_{node} \cdot n_{nodes} \cdot (1 + \nu) \cdot PUE \cdot CI_{\{hi,lo\}} \text{ (kgCO}_2e\text{)} \quad (3)$$

These are the emissions if every node was always working. We combine these with the load λ and idle power consumption factor γ to get the actual operational emissions.

$$C_{op,\{hi,lo\}} = C_{\{hi,lo\}} \cdot (\lambda + \gamma \cdot (1 - \lambda)) \quad (4)$$

The total emissions per site are then

$$C_{\{hi,lo\}} = C_{op,\{hi,lo\}} + C_{emb} \quad (5)$$

D. Rationale for the model construction

Our model for geographic load shifting is linear in the sense that the equations do not contain non-linear terms for any of the parameters. The actual model equations are presented and explained in Section IV-E and following. To explain the rationale behind the model construction, it is sufficient to consider the per-site expressions. The embodied emissions are constant; the operational emissions are of the form

$$C_{tot}(t) = a \cdot \lambda(t) \cdot C(t) + b \cdot C(t) + c \quad (6)$$

This is simply a rewrite of Eqs. 4 and 5 with $a = 1 - \gamma$, $b = \gamma$, $c = C_{emb}$, to more clearly show the structure.

There is no correlation between the load $\lambda(t)$ and the carbon intensity $C(t)$ (the electricity generation is not influenced by the load and unless the data centre actively manages demand or uses temporal shifting, the load is not influenced by the carbon intensity either). The load and CI values are time series but there is no dependency between time steps, the calculation of the total emissions is instantaneous at every time step. We can therefore treat $\lambda(t)$ and $C(t)$ as independent random variables and use the means of the distributions, i.e. the time-integrated values of the load and emissions over one year. In general, for independent random variables, the following relations hold between the expected values, regardless of their distributions (see e.g. [26]):

$$E(XY) = E(X) \cdot E(Y) \quad (7)$$

and, with a , b and c constants:

$$E(aX + bY + c) = a \cdot E(X) + b \cdot E(Y) + c \quad (8)$$

Therefore, Eq. 6 becomes

$$E(C_{tot}(t)) = a \cdot E(\lambda(t)) \cdot E(C(t)) + b \cdot E(C(t)) \quad (9)$$

which we write as

$$C_{tot} = a \cdot \lambda \cdot C + b \cdot C + c \quad (10)$$

In other words, to obtain the year-averaged emissions for the data centre, there is no need to calculate the emissions at every individual time step, and instead of working with traces we can work with averages.

E. Baseline emissions model

We assume that the data centres have excess capacity that could be used for geographic load shifting when there are no peak loads. We take into account the excess capacity using the load parameter λ and the idle power consumption parameter γ . In other words, a load of 50% means that the data centre is dimensioned for twice its nominal load. The distinction between high-emission and low-emission sites is not a strictly geographical one: a given site could be low-emission or high-emission depending on the energy mix at a given point in time.

$$C_b = 2 \cdot C_{emb} + (\lambda_{hi} + (1 - \lambda_{hi}) \cdot \gamma) \cdot C_{hi} + (\lambda_{lo} + (1 - \lambda_{lo}) \cdot \gamma) \cdot C_{lo} \quad (11)$$

F. Emissions model including geographic load shifting

Now we assume that we can use some of the free capacity at the low-emission site by moving it from the high-emission site. We use α and β to express how much of the workload is shifted. The proportion α of the total load that can be shifted expresses constraints on the workload: not all workloads can be shifted, e.g. because they are too large or because of legal, privacy or security considerations. The proportion β expresses what proportion of the time we can shift load. The rest of the time the emissions are given by the baseline model. For example, it only makes sense to shift load to a region with solar energy generation when that region is receiving sunshine. As α , β and λ evolve over time, we use yearly averages without loss of generality.

We have at most $(1 - \lambda_{lo})$ free capacity on the low-emission site but the amount of work we can shift is also limited by the freed-up capacity on the high-emission side (because the latter can't exceed the total capacity). When the free load is smaller than the capacity we want to shift, we need to cap α to not exceed free load; when the free load is higher than the load of the high-emission data centre, we need to cap α to not shift more work than is available.

$$\alpha_{eff} = \begin{cases} \alpha & , \alpha \cdot \lambda_{hi} \leq (1 - \lambda_{lo}) \\ \frac{(1 - \lambda_{lo})}{\lambda_{hi}} & , \lambda_{hi} \geq (1 - \lambda_{lo}) \\ 1 & , \text{otherwise} \end{cases} \quad (12)$$

By shifting the work, the load on the low site increases from λ_{lo} to $\lambda_{lo} + \alpha \cdot \lambda_{hi}$ and reduces on the high site from λ_{hi} to $\lambda_{hi}(1 - \alpha)$.

For example, if the load on the high side is 0.9 and on the low side 0.7, and we move 20% of the work 50% of the time, then the average load on the high side becomes $0.9 \cdot (1.0 - 0.2) = 0.72$; the load on the low side becomes $0.7 + 0.9 \times 0.2 = 0.88$, so that the total remains 1.6. In other words,

$$\lambda_{hi}(1 - \alpha) + (\lambda_{lo} + \alpha \cdot \lambda_{hi}) = \lambda_{hi} + \lambda_{lo} \quad (13)$$

The overhead of geographic load shifting is assumed to be proportional to the amount of work moved and the emissions from use of the nodes:

$$overhead = n_{nodes} \cdot \eta \cdot \alpha \cdot (c_{hi} + c_{lo}) \quad (14)$$

The intuition for this is that moving work requires additional computations and storage. A detailed analysis for this overhead is presented in [27], which also assumes this proportionality. However, this paper uses Carbon Intensity of Data Transfer (essentially energy/bit transferred) which is a disputed metric, see e.g. [28] who show that power draw of wired network infrastructure is almost independent of the volume of data traffic, so moving data does not affect the emissions. We will therefore be conservative and use overhead estimates of no more than 1%, which is on the low side of the interval calculated in [27], $0.5\% < \text{migration overhead} < 100\%$.

In the above we simplified the discussion by leaving assuming there was a single data centre in the high and low emissions zones. In general, we can have several, and to take this into account we introduce n_{hi} and n_{lo} . With those additional parameters, the constraints become:

$$\alpha_{eff} = \begin{cases} \alpha & , \alpha \cdot n_{hi} \cdot \lambda_{hi} \leq n_{lo} \cdot (1 - \lambda_{lo}) \\ \frac{n_{lo} \cdot (1 - \lambda_{lo})}{n_{hi} \cdot \lambda_{hi}} & , n_{hi} \cdot \lambda_{hi} \geq n_{lo} \cdot (1 - \lambda_{lo}) \\ 1 & , \text{otherwise} \end{cases} \quad (15)$$

The final model equation for geographic load shifting becomes

$$\begin{aligned} C_{gls} = & (n_{hi} + n_{lo}) \cdot C_{emb} + \\ & n_{hi} \cdot (\lambda_{hi}(1 - \alpha_{eff}) + \\ & (1 - \lambda_{hi}(1 - \alpha_{eff})) \cdot \gamma) \cdot C_{hi} + \\ & n_{lo} \cdot (\lambda_{lo} + \alpha_{eff} \cdot \lambda_{hi} \cdot n_{hi} / n_{lo} + \\ & (1 - \lambda_{lo} - \alpha_{eff} \cdot \lambda_{hi} \cdot n_{hi} / n_{lo}) \cdot \gamma) \cdot C_{lo} + \\ & \eta \cdot \alpha_{eff} \cdot (n_{hi} C_{hi} + n_{lo} C_{lo}) \end{aligned} \quad (16)$$

Taking into account the factor β , the complete model becomes

$$C = \beta \cdot C_{gls} + (1 - \beta) \cdot C_b \quad (17)$$

If $\beta = 0$ we have the baseline model.

G. Emission reductions from geographic load shifting

The relative reduction in emissions is simply the relative difference between the baseline and geographic load shifting models:

$$r = (C_b - C_{ls}) / C_b \quad (18)$$

If the load is the same on both sites, the equation for the “ideal” case (no embedded carbon, zero idle, no overhead, move everything all the time) reduces to:

$$r = \begin{cases} \frac{C_{hi} - C_{lo}}{C_{hi} + C_{lo}} & , \lambda \leq 0.5 \\ 1 - \frac{(2\lambda - 1) \cdot C_{hi} + C_{lo}}{\lambda \cdot (C_{hi} + C_{lo})} & , \lambda > 0.5 \end{cases} \quad (19)$$

This equation provides the upper limit for achievable reductions in emissions under ideal conditions. Already this provides some insights: if the load is < 0.5 , the achievable reduction does not depend on the load; for the limiting case when the load is one, no reductions are achievable. For example, for a load of 0.5 or less, if we move all work from the US to the UK all the time, (electricity carbon intensity resp.

369 gCO_2e/kWh and 211 gCO_2e/kWh [29]), the reduction in emissions is $158/580 = 0.272$, in other words we can never reduce emissions by more than 27% by doing so.

As the ideal conditions can't exist in reality, the more detailed model, which takes into account embodied carbon, load, idle power consumption, availability and overhead, allows us to make specific predictions of the reductions for concrete scenarios. The model was implemented in Haskell, the source code is available on line [30].

V. SCENARIOS FOR GEOGRAPHICAL LOAD SHIFTING

In this section we apply the above models to a number of scenarios. We consider scenarios for a commercial data centre attempting to optimise emissions through use of solar energy and wind energy and for HPC centres offloading work to centres in low-emission regions. The locations used are shown in Fig. 2. The countries are used in the commercial data centre scenarios, the HPC sites in the HPC scenarios.

A. Scenarios for commercial data centres

In these scenarios, summarised in Table I, what we want to assess is the relative reduction in emissions resulting from geographic load shifting for AI cloud data centres. We base our scenarios on the same assumptions as in the work by Lindberg *et al.* [9], which is one of the papers proposing geographic load shifting to reduce carbon emissions, but which approaches the problem from a grid cost optimisation perspective, with the data centre grid load as an opaque parameter. Lindberg *et al.* assume four identical data centres in different locations, with different but unspecified electricity carbon intensity. The maximum load is 300 MW and the nominal load 250 MW. The absolute figure is not important as we want to investigate the relative reduction in emissions, but 250 MW is a realistic capacity for a very large hyperscale data centre (hyperscale data centres of more than twice this capacity are currently being proposed [31]).

There is no detail on the internals of the data centre in the work by Lindberg *et al.*. We assume an AI data centre with where the nodes are GPU servers similar to the Nvidia DGX-A100. Each node consumes 4,550 W when active [32], 30% of that when idle [33]–[36]; the data centre PUE is taken to be 1.16, representative for a hyperscale data centre (but probably on the low side: most data centres have a higher PUE). We set the number of server nodes so that in total the data centre consumes at most 300 MW. In other words, the load for 250 MW is $\lambda = 0.83$ (250/300) and we can move a fraction $\alpha = 0.2$ (50 MW out of 250 MW).

Our embodied carbon model yields an estimate for this configuration of 5,730 $kgCO_2e/y$ per node. This assumes a 4-year useful life and takes into account the data centre infrastructure. This corresponds to 0.45 $ktCO_2e/MW/y$, which is between the estimates of [16] and [19].

Table I shows the parameters used and the results.

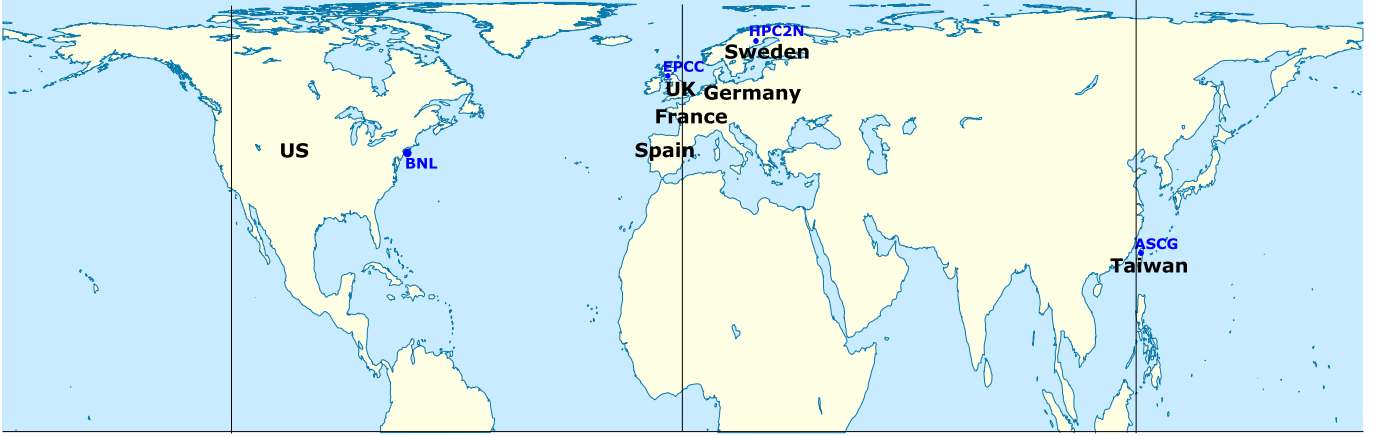


Fig. 2. Locations used in the scenarios. US, UK, Germany, France and Spain are used in the commercial data centre scenarios, the HPC sites (BNL, EPCC, HPC2N, ASCG) are used in the supercomputer scenarios.

TABLE I
PARAMETERS FOR GEOGRAPHIC LOAD SHIFTING TO OPTIMISE USE OF SOLAR ENERGY (SCENARIO 1)
AND WIND ENERGY (SCENARIO 2)

AI Data Centre Parameters	Scenario 1 (Solar)	Scenario 2 (Wind)
n_n	56,840	56,840
n_{hi}	2	2
n_{lo}	2	2
Embodied carbon c_{em} ($kgCO_2e/y$)	2,066	2,066
Operational emissions, high-CI c_{hi} ($kgCO_2e/y$)	18,978	17,451
Operational emissions, low-CI c_{lo} ($kgCO_2e/y$)	1,896	509
λ_{hi}	0.83	0.83
λ_{lo}	0.83	0.83
γ	0.30	0.30
α	0.20	0.20
β	0.52	0.54
η	0.00	0.00
overhead (tCO_2e/y)	0	0
Embodied (tCO_2e/y)	469,692	469,692
Baseline (tCO_2e/y)	2,565,724	2,273,088
Geographic load shifting (tCO_2e/y)	2,447,922	2,151,753
Emission reduction (%)	4.6%	5.3%

1) *Scenario 1: Geographic load shifting to optimise use of solar energy:* In this scenario, we start by assuming that for eight hours a day, three of the four data centres use predominantly solar power and for the rest of the day they are powered predominantly by fossil fuel generation. To make the example concrete, we assume that the other locations are the UK, the US and Germany, with average carbon intensities of resp. 211, 369 and 344 gCO_2e/kWh [29]. We assume the fourth data centre is located in France and uses nuclear when there is no sun. We assume for simplicity that the carbon intensity during solar powered operation is the same as the average CI. This is acceptable as according to the IPCC [37], solar power has an average CI of 41 and France has an average CI of 44 gCO_2e/kWh . According to [38], the CI varies between 6.5 and 108 gCO_2e/kWh depending on location and technology used, with most sites between 20 and 50 gCO_2e/kWh .

We further assume that the CI per country is the time-weighted average of the CI of predominantly solar and pre-

dominantly fossil fuel generation:

$$CI_{avg} = (8 \times CI_{sun} + 16 \times CI_{fossil})/24 \quad (20)$$

The "predominantly" is there because the energy mix typically includes non-fossil fuel sources such as nuclear and wind energy. We calculate the CI for the period with predominantly fossil fuel generation from the average CI for the country and the CI for solar energy generation. For example for the UK we have

$$\begin{aligned}
 CI_{fossil, UK} &= (24 \times CI_{avg, UK} - 8 \times CI_{sun})/16 \\
 &= (3 \times 211 - 1 \times 44)/2 \\
 &= 295 \text{ (} gCO_2e/kWh \text{)}
 \end{aligned} \quad (21)$$

The average of the thus calculated CI from predominantly fossil fuel generation over the three countries is 448 gCO_2e/kWh ; the predominantly solar generation is assumed to be 41 gCO_2e/kWh as per the IPCC [37].

As illustrated in Fig. 3, at any point in time there will always be two data centres with low CI and two with high CI. Fig.

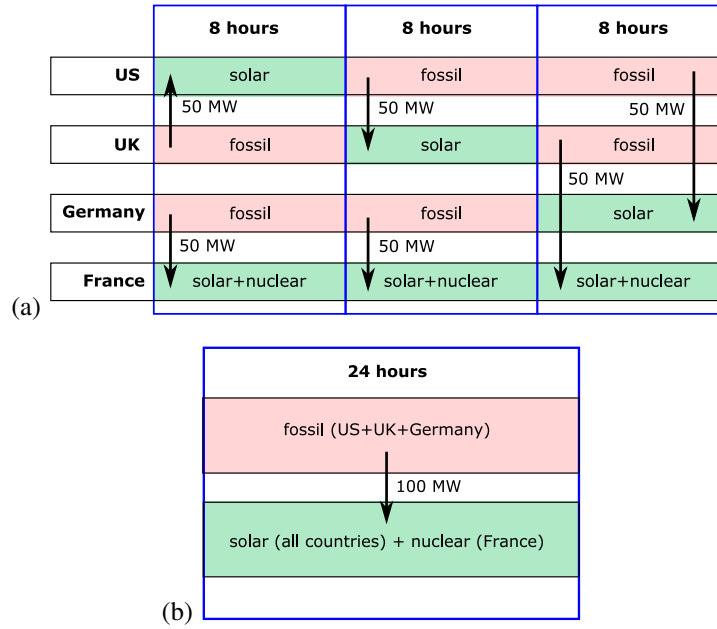


Fig. 3. Solar scenario, ideal assumptions (two sites are low-CI in any 8-hour period). 100 MW can be shifted.

3(a) shows the generation for every data centre in every 8-hour timespan; Fig. 3(b) shows that we can reorder this so that it corresponds to an equivalent load shifting model where 50 MW is moved from a high CI location to low CI one (so 100 W in total) all of the time, even though in reality we will move the load 2/3 of the time and to different sites. In other words, we can model this 4-site scenario using an equivalent 2-site scenario as used in our model. We use the average CI for the three high-emission regions for CI_{hi} . This is an optimal scenario as we can always move the maximum load.

As formulated, the scenario is overly optimistic as it assumes that in all three countries the sun will shine eight hours a day year round. This would be 2,920 hours of sunshine per year. For our example, in the UK, the average number of hours of sunshine per year is 1,524 hours [39]; in the US it is 2,627 hours (average over all states) [40] and in Germany it is 1,665 hours [41]. For France, there is of course no such correction, so the average number of hours is $(h_{US} + h_{UK} + h_D + 3 \cdot h_F)/6 = 0.83$. We correct for this by reducing the fraction of the time that the workload can be moved, β , by this factor. Furthermore, the scenario assumes that there is at all times a data centre in a region where the sun shines. If we have two data centres in Europe and one in the US, this is not the case. (We exclude the one in France as it is always low-CI because of the nuclear power.) Excluding Hawaii, US time ranges from UTC-8 to UTC-4. So we will assume the US data centre to be on UTC-6; Europe is on UTC+1.

For the purpose of illustration, in Fig. 4 we simplify this to no overlap between the US, the UK and Europe (i.e. we assume the entire US is on UTC-8 and Europe is on UTC). Then there is an 8-hour period where one out of three data centres is in a low emissions region, another with two and another with none, as illustrated in Fig. 4. In the first period,

we can move 100 MW; in the second and third period, we can only move 50 MW, so in total we can only move 200 MW rather than 300 MW. In practice, there will be 2 hours overlap between the US and the UK and 1 hour non-overlap between Germany and the UK, so that the final factor is not 2/3 but 5/8. We account for this further correcting β to 0.52.

2) *Scenario 2: Periodic load shifting to optimise use of wind energy* : In this scenario we assume that the four data centres are in the US, the UK, Germany and Spain (the choice of Spain is because after Germany, it is the EU country with the largest wind power capacity installed). We start by assuming that on average there is always enough wind in the regions of two out of four data centres (Fig. 5). This means that we can always move the load from the other two, so this amounts to the same situation as for Scenario 1, and we can move 100 MW between sites. For the wind generation CI we use 11 gCO_2e/kWh as per the IPCC [37]. According to [42]–[45], reported values vary between 10.3 and 45.2 gCO_2e/kWh . The average CI from predominantly fossil fuel generation over the four countries, calculated as above, is 531 gCO_2e/kWh .

As formulated, the scenario is again overly optimistic as it assumes there will always be enough wind to move 100 MW. In practice, load factors are not 50% but between 30% and 40% or lower [46], [47]; we will therefore reduce β to 0.7. Furthermore, there tends to be a correlation in the weather patterns across Europe, so that it will not always be possible to move all the load. Maps for the Pearson correlation for hourly wind capacity within Europe are presented in [14]. Based on those values and assuming no correlation with the US, using the average correlation between the UK, Germany and Spain reduces the effectiveness by 10%, so we finally set $\beta = 0.63$. This is illustrated in Fig. 6. In this example, in periods 1 and 4 there is no correlation, so load shifting is reduced from 50 MW to 35 MW by the wind load factor; the windy episodes

	8 hours	8 hours	8 hours
US	solar ↑ 50 MW	fossil ↓ 50 MW	fossil
UK	fossil	solar ↓ 50 MW	fossil
Germany	fossil ↓ 50 MW	solar	fossil ↓ 50 MW
France	solar+nuclear	solar+nuclear	solar+nuclear

Fig. 4. Solar scenario, more realistic assumptions (two, three or one low-CI sites in an 8-hour period). Only $(100+50+50)/3$ MW can be shifted.

	period 1	period 2	period 3	period 4
US	wind ↑ 50 MW	fossil ↓ 50 MW	wind ↑ 50 MW	fossil ↓ 50 MW
UK	fossil	wind ↓ 50 MW	fossil	wind ↓ 50 MW
Germany	fossil ↓ 50 MW	fossil	wind ↑ 50 MW	wind ↑ 50 MW
Spain	wind ↓ 50 MW	wind ↓ 50 MW	fossil	fossil

Fig. 5. Wind scenario, ideal assumptions (wind load factor 50% for all sites, no correlation). 100 MW can be shifted.

	period 1	period 2	period 3	period 4
US	wind ↑ 35 MW	fossil ↓ 28 MW	wind ↑ 28 MW	fossil ↓ 35 MW
UK	fossil	fossil ↓ 28 MW	fossil	wind ↓ 35 MW
Germany	fossil ↓ 35 MW	fossil	wind ↑ 28 MW	wind ↑ 35 MW
Spain	wind ↓ 35 MW	fossil ↓ 28 MW	fossil	fossil

Fig. 6. Wind scenario, more realistic assumptions (load factor 35% and correlation 10%). Only 63 MW can be shifted.

in periods 2 and 3 are overlapping by 20% (full correlation would mean complete overlap), leading to a further reduction to 28 MW. So in total, only 63 MW can be shifted instead of 100 MW.

3) *Discussion of Commercial Data Centre Scenarios:* As we can see from Table I, for both scenarios, which are still quite optimistic, the reduction in emissions is around 5% (4.6% and 5.3% respectively for the solar and wind scenarios). Several factors conspire to limit the gains:

- the carbon intensity of any current renewable technology is not zero.
- the CI of predominantly fossil fuel generation is lower than the worst case (100% coal generation is $820 \text{ gCO}_2\text{e/kWh}$; 100% gas is $490 \text{ gCO}_2\text{e/kWh}$, [37]; most data centres are in the EU and the US which both have CI lower than this ($207 \text{ gCO}_2\text{e/kWh}$ resp. $369 \text{ gCO}_2\text{e/kWh}$)
- the important contribution of embodied carbon in the low-

emission case,

- the idle power consumption is not zero and
- the fact that in realistic scenarios we can't move sufficient load to use up all excess capacity in the low-emission region all the time.

To explore these factors, in Fig. 7 we plot the emission reductions as a function of the load. We move as much work as possible. The set of curves shows the combined effect of each factor: "Idle power = 0" removes the effect of the idle power consumption; "Embodied carbon = 0" removes the embodied carbon contribution; "No time constraints" does both and ignores the time constraints on the renewables generation. The kink in the figure is the result of the constraint on α , which in these scenarios is triggered when the load exceeds 0.5. The intuition is that for loads < 0.5 , the amount of work that can be shifted is limited by the available work in the high-emission zone, whereas for loads > 0.5 , the limit is the free capacity in the low-emission zone. The reason why the reduction is

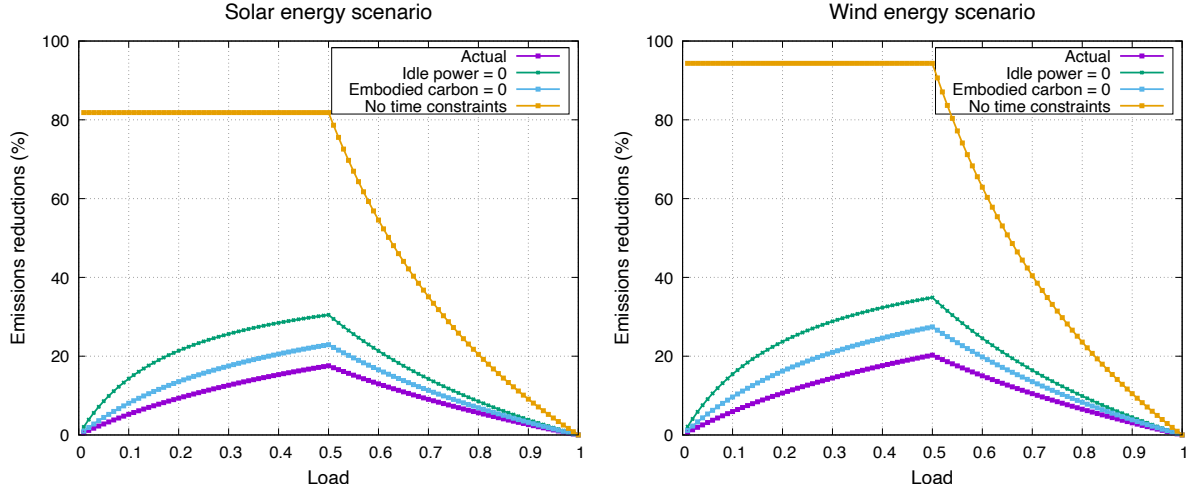


Fig. 7. Emissions reduction as a function of load for the AI data centre scenarios. The figure illustrates the effect of load λ , idle power γ , embodied emissions c_{em} and time constraints β .

independent of the load is that all work is moved to the low-emission zone, so that effectively the reduction is given by the second case in Eq. 19, $(1 - CI_{lo})/(CI_{hi} + CI_{lo})$

Our model also ignores the actual grid capacity. The low-emission location is only low emission if the additional energy required by the shifting of the workloads does not exceed the curtailed renewables generation. If data centre operators deploying carbon-aware computing move more work than the curtailed capacity then non-renewable generators will switch on and the result will be an increase in electricity generation carbon intensity. Using grid-aware computing, less work would be shifted to avoid exceeding the curtailed capacity [48]. In practice, curtailment is quite low. According to the IEA [49], it is between 1.5% and 4% in most large renewable energy markets, so there is not all that much scope for moving data centre workloads at scale.

B. Supercomputer centre scenarios

In this section we consider a different type of scenario, that of moving workloads between supercomputers. HPC centres are built for a notional capacity and their workloads do not depend strongly on market factors. In a network such as CERN's Worldwide LHC Computing Grid (WLCG), load could be moved between locations with high and low emissions if spare capacity would be available.

We model this using actual values for the ASGC HPC centre in Taiwan and the HPC2N centre in Sweden, CI 642 gCO_2e/kWh resp. 36 gCO_2e/kWh , as well as BNL in the US and EPCC in the UK, CI resp. 369 gCO_2e/kWh and 211 gCO_2e/kWh (CI values from [29]).

We assume a node consisting of 2x AMD EPYC 9754 HT with 1TB RAM and 0.5TB NVME SSD, with a node power consumption of 1.2 kW. Our embodied carbon model yields an estimate for this configuration of 444 $kgCO_2e/y$. This assumes a 4-year useful life and takes into account the data centre infrastructure. This corresponds to 0.37 $ktCO_2e/MW/y$ and is again between the existing estimates.

Table II shows the parameters used and the results.

1) *Scenario 1:* The first three scenarios model moving workloads from the ASGC HPC centre in Taiwan, which has a very high electricity carbon intensity, to the HPC2N centre in Sweden, which has a very low electricity carbon intensity. This is the most optimistic case. From discussion with HPC facility operators, we assume an average load of 80%. However, for the first scenario we assume HPC2N has 50% free capacity for the entire year and ASGC has a load of 100%. We further assume that we would like to move as much of the work for as much of the time as possible from the ASGC cluster to HPC2N. These are very optimistic assumptions. We assume no overhead in moving the work. In this case, the reduction in emissions is 30.5% or about 365 tCO_2e/y . We start with this scenario to demonstrate that only highly unrealistic scenarios result in high reductions.

2) *Scenario 2:* In this scenario we assume that both ASGC and HPC2N have an average load of 80%, which is a realistic load for HPC centres, and that we can move as much work as possible (i.e. so that HPC2N is 100% loaded). We also assume an overhead of 1% for moving the work. There are no data for this overhead, but even if it was zero, this would not change the figures much. This results in a reduction in emissions of 13.6%.

3) *Scenario 3:* In practice, it will not be possible to move all of the work all of the time. Many HPC workloads are limited in particular by large data sets and large volumes of data produced, which would make the workload movement impractical, or by software or data set licensing or data privacy issues which prohibit work movement. In this scenario, we assume that we can move a quarter of the workload for half of the time. This amounts to moving 12.5% instead of 20%. The resulting reduction in emissions is 6.8%.

4) *Scenario 4:* Moving workloads from sites with very high CI to ones with very low CI is of course optimal, but most HPC centres are neither in very high nor very low emission regions. We therefore look at moving workloads between less

TABLE II

SCENARIO 1: EXCESS CAPACITY IS 50%, MAXIMUM POSSIBLE AMOUNT OF WORK IS MOVED, NO OVERHEAD;
 SCENARIO 2: AS SCENARIO 1 BUT EXCESS CAPACITY IS 20%, OVERHEAD 1%;
 SCENARIO 3: AS SCENARIO 2 BUT ONLY MOVE 12.5% OF THE WORK;
 SCENARIO 4: AS SCENARIO 2 BUT MOVE BETWEEN BNL (US) AND EPCC (UK);
 SCENARIO 5: SCENARIO 3 WITH CI OF SCENARIO 4;
 VALUES FOR EMISSIONS ARE IN $kgCO_2e/y$

HPC Centre Parameters	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
n_n	100	100	100	100	100
n_{hi}	1	1	1	1	1
n_{lo}	1	1	1	1	1
Embodied carbon c_{em} ($kgCO_2e/y$)	444	444	444	444	444
Operational emissions, high-CI c_{hi} ($kgCO_2e/y$)	10,831	10,831	10,831	3,879	3,879
Operational emissions, low-CI c_{lo} ($kgCO_2e/y$)	390	390	390	1,304	1,304
λ_{hi}	1.00	0.80	0.80	0.80	0.80
λ_{lo}	0.50	0.80	0.80	0.80	0.80
γ	0.30	0.30	0.30	0.30	0.30
α	1.00	1.00	0.25	1.00	0.25
β	1.00	1.00	0.50	1.00	0.50
η	0.00	0.01	0.01	0.01	0.01
overhead (tCO_2e/y)	0	3	1	1	1
Embodied (tCO_2e/y)	89	89	89	89	89
Baseline (tCO_2e/y)	1,197	1,054	1,054	534	534
Geographic load shifting (tCO_2e/y)	832	910	982	500	517
Emission reduction (%)	30.5%	13.6%	6.8%	6.5%	3.3%

extreme high and low emission sites, BNL in the US and EPCC in the UK. Using the other parameters from Scenario 2, this results in a reduction in emissions of 6.5%.

5) *Scenario 5*: Finally, in what is probably the most realistic scenario, we use Scenario 3 with the CI of Scenario 4. In other words, we move a realistic amount of work between sites with a realistic difference in CI. This results in a reduction in emissions of 3.3%.

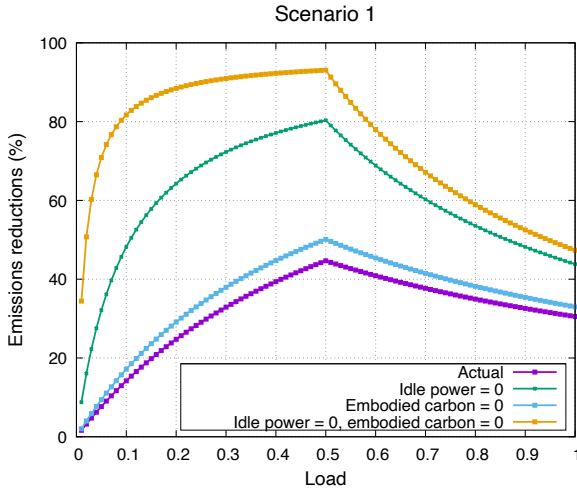


Fig. 8. Emissions reduction as a function of load for HPC centre Scenario 1. The figure illustrates the effect of load λ , idle power γ and embodied emissions c_{em} .

6) *Discussion of HPC Scenarios*: For Scenario 1, Fig. 8 shows the effect of load λ , idle power γ , embodied emissions c_{em} . In this scenario, $\lambda_{lo} = 0.5$ and we sweep λ_{hi} . Fig. 9 explores the effect of load λ , idle power γ , embodied emissions c_{em} and load flexibility β . For Scenarios 2 and 4, $\beta = 1$ and therefore the "No time constraints" case is

redundant. Comparing Scenario 2/4 to Scenario 3/5 shows the impact of the flexibility in moving work; comparing Scenario 2/3 to Scenario 4/5 shows the impact of the CI differential.

These supercomputer scenarios show that, even if the data centres don't provision excess capacity and therefore don't incur excess embodied carbon, geographic load shifting only results in large reductions if the free capacity at the low-emission HPC centre is very high, the high-emission site has very high emissions and the low emissions site very low ones, and the overhead for moving the workloads is very small (Scenario 1). Unless there is complete flexibility in moving workloads, i.e. we can move the work all the time, the most likely obtainable emission reductions are around five percent. But many scientific computing workloads are limited in particular by large data sets and large volumes of data produced, which would make the workload movement impractical. Other techniques such as frequency downscaling for I/O-limited workloads [50] or heat reuse [51] can yield much higher reductions, and should therefore be prioritised. Reducing embodied carbon through server life extension will become increasingly important as the grid decarbonises.

VI. CONCLUSION

Based on realistic scenarios for commercial AI data centres and HPC centres, and a linear analytical model that takes into account key server and data centre parameters, we have simulated the reduction in emissions achievable through for geographic load shifting. We conclude that the reductions in emissions from geographic load shifting are small (typically less than five percent) for realistic scenarios for both commercial data centres and HPC centres. Much larger reductions in emissions are needed to counter the current growth in global data centre capacity.

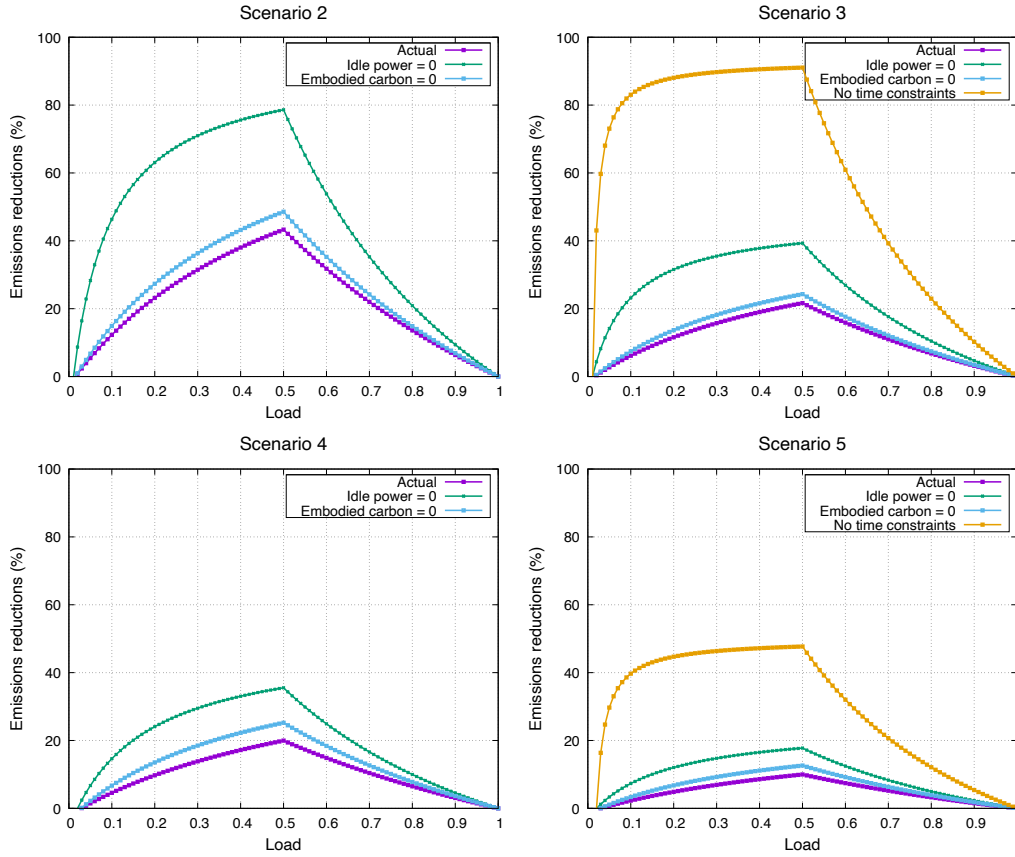


Fig. 9. Emissions reduction as a function of load for the HPC centre scenarios. The figure illustrates the effect of load λ , idle power γ , embodied emissions c_{em} and time constraints β .

ACKNOWLEDGEMENT

We thank Dr. Lauritz Thamsen for his thorough review and helpful suggestions which led to a much improved manuscript.

REFERENCES

- [1] B. Knowles, K. Widdicks, G. Blair, M. Berners-Lee, and A. Friday, "Our house is on fire: The climate emergency and computing's responsibility," *Communications of the ACM*, vol. 65, no. 6, pp. 38–40, 2022.
- [2] L. Belkhir and A. Elmeligi, "Assessing ict global emissions footprint: Trends to 2040 & recommendations," *Journal of cleaner production*, vol. 177, pp. 448–463, 2018.
- [3] McKinsey & Company, "AI power: Expanding data center capacity to meet growing demand," Oct 2024. [Online]. Available: <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand>
- [4] A. Shehabi, A. Hubbard, A. Newkirk, N. Lei, M. A. B. Siddik, B. Holec, J. Koomey, E. Masanet, D. Sartor *et al.*, "2024 United States Data Center Energy Usage Report," Lawrence Berkeley National Laboratory, Tech. Rep. LBNL-2001637, dec 2024.
- [5] Ember, "Electricity generation from other renewables, excluding bioenergy – Ember," 2025, energy Institute - Statistical Review of World Energy (2024) - with major processing by Our World in Data. [Online]. Available: <https://ourworldindata.org/grapher/electricity-prod-source-stacked>
- [6] W. Vanderbauwhede, "The real problem with the AI hype," 01 2025. [Online]. Available: <https://limited.systems/articles/the-real-problem-with-AI/>
- [7] A. Radovanović, R. Koningstein, I. Schneider, B. Chen, A. Duarte, B. Roy, D. Xiao, M. Haridasan, P. Hung, N. Care *et al.*, "Carbon-aware computing for datacenters," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1270–1280, 2022.
- [8] P. Wiesner, I. Behnke, D. Scheinert, K. Gontarska, and L. Thamsen, "Let's wait awhile: How temporal workload shifting can reduce carbon emissions in the cloud," in *Proceedings of the 22nd International Middleware Conference*, 2021, pp. 260–272.
- [9] J. Lindberg, B. C. Lesieutre, and L. A. Roald, "Using geographic load shifting to reduce carbon emissions," *Electric Power Systems Research*, vol. 212, p. 108586, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0378779622006757>
- [10] C. Barrows, A. Bloom, A. Ehlen, J. Ikäheimo, J. Jorgenson, D. Krishnamurthy, J. Lau, B. McBenett, M. O'Connell, E. Preston, A. Staid, G. Stephen, and J.-P. Watson, "The iee reliability test system: A proposed 2019 update," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 119–127, 2020.
- [11] B. Acun, B. Lee, F. Kazhamiaka, K. Maeng, U. Gupta, M. Chakkaravarthy, D. Brooks, and C.-J. Wu, "Carbon explorer: A holistic framework for designing carbon aware datacenters," in *Proc. ASPLOS 2023*, ser. ASPLOS 2023. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3575693.3575754>
- [12] W.-T. Lin, G. Chen, and H. Li, "Carbon-aware load balance control of data centers with renewable generations," *IEEE Transactions on Cloud Computing*, vol. 11, no. 2, pp. 1111–1121, 2023.
- [13] J. Bian, L. Wang, S. Ren, and J. Xu, "CAFE: Carbon-Aware Federated Learning in Geographically Distributed Data Centers," in *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems*, 2024, pp. 347–360.
- [14] I. Riepin, T. Brown, and V. M. Zavala, "Spatio-temporal load shifting for truly clean computing," *Advances in Applied Energy*, vol. 17, p. 100202, 2025.
- [15] F. Acun, I. C. Paschalidis, and A. K. Coskun, "Conductor: A collaboration framework for multi-data-center demand response," in *2024 IEEE 15th International Green and Sustainable Computing Conference (IGSC)*, 2024, pp. 22–28.

- [16] J. Zheng, A. A. Chien, and S. Suh, "Mitigating curtailment and carbon emissions through load migration between data centers," *Joule*, vol. 4, no. 10, pp. 2208–2222, 2020.
- [17] B. Whitehead, D. Andrews, and A. Shah, "The life cycle assessment of a uk data centre," *The International Journal of Life Cycle Assessment*, vol. 20, pp. 332–349, 2015.
- [18] S. Tannu and P. J. Nair, "The dirty secret of ssds: Embodied carbon," *SIGENERGY Energy Inform. Rev.*, vol. 3, no. 3, p. 4–9, Oct. 2023. [Online]. Available: <https://doi.org/10.1145/3630614.3630616>
- [19] P. Lin, R. Bunker, and V. Avelar, "Quantifying Data Center Scope 3 GHG Emissions to Prioritize Reduction Efforts," may 2023. [Online]. Available: https://www.se.com/ww/en/download/document/SPD_WP99_EN/
- [20] M. Wadenstein and W. Vanderbauwhede, "Life cycle analysis for emissions of scientific computing centres," *The European Physical Journal C*, vol. 85, no. 8, p. 913, Aug. 2025.
- [21] —, "Model for life cycle analysis for emissions of scientific computing centres," Aug. 2025. [Online]. Available: <https://doi.org/10.5281/zenodo.16894210>
- [22] R. Lorenzini, "Digital & environment: How to evaluate server manufacturing footprint, beyond greenhouse gas emissions?" Nov 2021. [Online]. Available: <https://boavizta.org/en/blog/empreinte-de-la-fabrication-d-un-serveur>
- [23] U. Gupta, M. Elgamal, G. Hills, G.-Y. Wei, H.-H. S. Lee, D. Brooks, and C.-J. Wu, "Act: designing sustainable computer systems with an architectural carbon modeling tool," in *Proceedings of the 49th Annual International Symposium on Computer Architecture*, ser. ISCA '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 784–799.
- [24] M. Garcia Bardon, P. Wuytens, L.-A. Ragnarsson, G. Mirabelli, D. Jang, G. Willems, A. Mallik, A. Spessot, J. Ryckaert, and B. Parvais, "Dtco including sustainability: Power-performance-area-cost-environmental score (ppace) analysis for logic technologies," in *2020 IEEE International Electron Devices Meeting (IEDM)*, 2020, pp. 41.4.1–41.4.4.
- [25] A. Shilov, "SK Hynix Details DDR5-6400," Feb 2019. [Online]. Available: <https://www.anandtech.com/show/13999/sk-hynix-details-its-ddr56400-dram-chip>
- [26] J. Kermond, "An introduction to the algebra of random variables," in *Proceedings of the 47th Annual Conference of the Mathematical Association of Victoria*, 2010, pp. 1–16.
- [27] Y. Guo, A. Tomlinson, R. Su, and G. Porter, "The effect of the network in cutting carbon for geo-shifted workloads," *arXiv preprint arXiv:2504.14022*, 2025.
- [28] D. Schien, P. Shabajee, L. Krug, G. McSorley, and C. Preist, "Causal allocation of fixed impacts in product systems: Assessing the effect of data demand on network energy consumption," *Journal of Industrial Ecology*, vol. n/a, no. n/a, Jul 2025. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jiec.70057>
- [29] Ember, "Carbon intensity of electricity generation - Ember and Energy Institute," 2025, energy Institute - Statistical Review of World Energy (2024) - with major processing by Our World in Data. [Online]. Available: <https://ourworldindata.org/grapher/carbon-intensity-electricity>
- [30] W. Vanderbauwhede, "A model for the achievable reduction in emissions from carbon-aware geographic load shifting," 2025. [Online]. Available: <https://codeberg.org/wimvanderbauwhede/carbon-aware-geographic-load-shifting-model>
- [31] Mark Allen Group, "Scotland's Ravenscraig proposed for £4bn hyperscale data centre," Jun 2025. [Online]. Available: <https://www.theengineer.co.uk/content/news/ravenscraig-site-proposed-for-4bn-hyperscale-data-centre>
- [32] A. Gray, "Energy and Power Efficiency for Applications on the Latest NVIDIA Technology," Mar 2024. [Online]. Available: <https://developer-blogs.nvidia.com/wp-content/uploads/2024/10/Energy-Efficiency-GTC.pdf>
- [33] P. Kennedy, "AMD EPYC 7742 Benchmarks and Review Simply Peerless," dec 2019. [Online]. Available: <https://www.servethehome.com/amd-epyc-7742-benchmarks-and-review-simply-peerless/>
- [34] Standard Performance Evaluation Corporation, "SPEC CPU 2017 Integer Speed Result," 2019. [Online]. Available: <http://spec.org/cpu2017/results/res2020q2/cpu2017-20200413-21925.pdf>
- [35] Z. Yang, K. Adamek, and W. Armour, "Accurate and convenient energy measurements for gpus: A detailed study of nvidia gpu's built-in power sensor," in *SC24: International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE, Nov. 2024, pp. 1–17. [Online]. Available: <http://dx.doi.org/10.1109/SC41406.2024.00028>
- [36] Y. Zhang, Q. Wang, Z. Lin, P. Xu, and B. Wang, "Improving GPU energy efficiency through an application-transparent frequency scaling policy with performance assurance," in *Proceedings of the Nineteenth European Conference on Computer Systems*, 2024, pp. 769–785.
- [37] S. S., T. Bruckner, L. Fulton, E. Hertwich, A. McKinnon, D. Perczyk, J. Roy, R. Schaeffer, R. Sims, P. Smith, and R. Wiser, "Annex III: Technology-specific cost and performance parameters." 2014, Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.).
- [38] A. Ziemińska-Stolarska, M. Pietrzak, and I. Zbiciński, "Application of lca to determine environmental impact of concentrated photovoltaic solar panels - state-of-the-art," *Energies*, vol. 14, no. 11, 2021. [Online]. Available: <https://www.mdpi.com/1996-1073/14/11/3143>
- [39] UK Met Office, "UK temperature, rainfall and sunshine time series," 2025. [Online]. Available: <https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-temperature-rainfall-and-sunshine-time-series>
- [40] NOAA National Centers for Environmental Information, "WMO Climate Normals," 2025. [Online]. Available: <https://www.ncei.noaa.gov/products/wmo-climate-normals>
- [41] Deutsche Wetterdienst, "Deutschlandwetter im Jahr 2022," 2022. [Online]. Available: https://www.dwd.de/DE/presse/pressemitteilungen/DE/2022/20221230_deutschlandwetter_jahr2022_news.html
- [42] A. Alsaleh and M. Sattler, "Comprehensive life cycle assessment of large wind turbines in the us," *Clean Technologies and Environmental Policy*, vol. 21, pp. 887–903, 2019.
- [43] A. Garcia-Teruel, G. Rinaldi, P. R. Thies, L. Johanning, and H. Jeffrey, "Life cycle assessment of floating offshore wind farms: An evaluation of operation and maintenance," *Applied Energy*, vol. 307, p. 118067, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0306261921013520>
- [44] Q. Li, H. Duan, M. Xie, P. Kang, Y. Ma, R. Zhong, T. Gao, W. Zhong, B. Wen, F. Bai, and A. K. Vuppalladiyam, "Life cycle assessment and life cycle cost analysis of a 40 mw wind farm with consideration of the infrastructure," *Renewable and Sustainable Energy Reviews*, vol. 138, p. 110499, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032120307851>
- [45] M. Ozoemena, W. M. Cheung, and R. Hasan, "Comparative lca of technology improvement opportunities for a 1.5-mw wind turbine in the context of an onshore wind farm," *Clean Technologies and Environmental Policy*, vol. 20, no. 1, pp. 173–190, 2018.
- [46] A. Albatayneh, R. AbuAlRous, M. Kay, R. Abdallah, A. Juaidi, A. Garcia-Cruz, and F. Manzano-Agugliaro, "Wind farm capacity factor forecasting: An australian case study," *Energy Nexus*, vol. 18, p. 100422, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2772427125000634>
- [47] I. Staffell and R. Green, "How does wind farm performance decline with age?" *Renewable Energy*, vol. 66, pp. 775–786, 2014. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0960148113005727>
- [48] I. Velasco, "Carbon Aware Computing: Next Green Breakthrough or New Greenwashing?" 01 2024. [Online]. Available: <https://hackernoon.com/carbon-aware-computing-next-green-breakthrough-or-new-greenwashing>
- [49] International Energy Agency, "Renewable Energy Market Update – Outlook for 2023 and 2024," Jun 2023. [Online]. Available: <https://www.iea.org/reports/renewable-energy-market-update-june-2023/will-more-wind-and-solar-pv-capacity-lead-to-more-generation-curtailment>
- [50] A. Jackson, A. Simpson, and A. Turner, "Emissions and energy efficiency on large-scale high performance computing facilities: Archer2 uk national supercomputing service case study," in *Proceedings of the SC23 Workshops of the International Conference on High Performance Computing, Network, Storage, and Analysis*, ser. SC-W 23. New York, NY, USA: Association for Computing Machinery, 2023, pp. 1866–1870. [Online]. Available: <https://doi.org/10.1145/3624062.3624269>
- [51] V. Ljungdahl, M. Jradi, and C. Veje, "A decision support model for waste heat recovery systems design in data center and high-performance computing clusters utilizing liquid cooling and phase change materials," *Applied Thermal Engineering*, vol. 201, p. 117671, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1359431121010966>