

Team Control Number

15803

Problem Chosen

B

2024

HiMCM

Summary Sheet

Summary

Contents

1	Introduction	3
1.1	Background information	3
1.2	Restatement of the problems	3
1.3	Our Work	3
2	HPC energy consumption model	3
2.1	Assumptions	3
2.2	Model overview	5
2.2.1	Energy consumption of data centers	5
2.2.2	Energy consumption of cryptocurrency mining	5
2.2.3	Total annual energy consumption	6
2.3	Assessing the model using current data	6
2.3.1	Data centers	6
2.3.2	Cryptocurrency mining	7
2.3.3	Evaluation	7
3	HPC carbon emission model	7
3.1	Assumptions	8
3.2	Carbon emissions at a fixed time	8
3.2.1	Carbon emission of data centers	9
3.2.2	Carbon emission of cryptocurrency mining	9
3.2.3	Results and evaluation	10

1 Introduction

1.1 Background information

Under the current global trend of technology advancement, high-powered computing (HPC) is gaining rising attention for the increasing demand of computationally intensive tasks such as data science, artificial intelligence (AI) training, and cryptocurrency mining.

It is very common to use a massive number of dedicated hardware with strong computing power for these tasks. Data centers are the specific spaces to hold these hardware. The operation of data centers is usually extremely energy-consuming, with a tremendous demand of electricity to keep the computer systems running and to cool them down. Cryptocurrency mining machine clusters also work similarly to data centers, with all resources dedicated for mining. The intensive use of electricity, contributing heavily to global warming, as well as a lot of other problems, e.g. water usage, e-waste, etc., poses the environmental concerns of HPC.

As sustainability arising as one of the most important development goals in the twenty-first century, it is high time we take action to assess the environmental consequences of HPC and to minimize its negative effects. We proposed an approach to evaluate the environmental impacts due to HPC through mathematical models.

[**TODO** add more about past researches on energy demands; cites needed]

[**TODO** dataset and used tools]

1.2 Restatement of the problems

[**TODO** adjust this section later]

[**TODO**]

1.3 Our Work

[**TODO**]

2 HPC energy consumption model

In this section, we model the annual worldwide energy consumption of *high-powered computing* (HPC). HPC includes various computationally intensive tasks, which are mostly carried out by computer systems in data centers and cryptocurrency mining machines. Therefore, it is feasible to estimate the total energy consumption of HPC from these two sectors.

2.1 Assumptions

1. The total power data offered by data center providers reflects the maximum power.
Justification: The data of total power in a data center is typically defined by the machines' design specifications. Therefore, we interpret this as the hard limit of the possible power allocated on the machines.
2. Possible non-HPC data centers are also included.
Justification: Since most data centers are hybrid, it is hard to exactly distinguish between HPC and non-HPC ones. Non-HPC hardware such as storage may also contribute to the

Table 1: Definition of symbols. Units: TWh: terawatt hours; MW: megawatts; TH: terahashes.

Symbol	Unit	Description
E	TWh	Total annual energy consumption of all HPC activities.
E_D	TWh	Total annual energy consumption of data centers.
E_M	TWh	Total annual energy consumption of cryptocurrency mining.
T	h	The time duration (in this case, = 1yr = 8760h).
P_i	MW	Total maximum power of data center i .
$\bar{\rho}$	–	Average utilization rate of data centers.
H	TH/s	Total hash rate performed by the network.
κ	–	Multiplier between the average and the optimal mining efficiency.
ε_{opt}	J/TH	Efficiency of the best mining machine (power per hash rate).
$\bar{\varepsilon}$	J/TH	Average efficiency among mining machines ($\bar{\varepsilon} = \kappa \varepsilon_{\text{opt}}$).

process of HPC. Moreover, they share most of the infrastructure with HPC consuming a more significant amount of energy. Therefore, including some possible non-HPC facilities is not likely to create a large impact on the results.

3. All data are from a same, *short-run* scenario where there are no differences in conditions.

Justification: Most of the data we found are statistics or information within or throughout the year 2022 to 2024. We tried to use the latest statistics, but this is not always available since the year of 2024 is not over yet by the time we work on this model. The assumption of the *short-run* data means that we regard all data as taken from a same set of conditions with no development or changes between them; delving into the detailed differences would make the model way too complex and beyond the intended scope.

4. The utilization rates across different data centers are the same.

Justification: The actual utilization rates across different data centers are not varying very much. Since the emphasis of our model is on the estimation of total energy, the detailed utilization rates on each data center are not of significant interest and neglected.

5. Bitcoin (BTC) mining is considered as the whole cryptocurrency mining industry.

Justification: By the time we work on this model, Bitcoin (BTC) is the cryptocurrency with the largest capitalization and mining energy consumption. Cryptocurrencies other than BTC, namely *altcoins*, consumes way less energy than BTC mining, especially with the second largest Ethereum (ETH) switched to proof of stake (PoS) instead of the traditional proof of work (PoW) mining. **[TODO cite]**

6. Cryptocurrency mining machines are always operating at their *full capacity rate*.

Justification: As a highly-competitive industry, miners are profit-driven and tend to maximize their profits. Mining at the machines' full capacity ensures the maximum output which lowers the averaged fixed costs, such as the costs of mining machines and related infrastructures, and thus maximizes the profits.

7. All other unlisted source of HPC are not considered.

Justification: We believe the two sectors are enough to give a representation of all HPC activities. Unlisted sources, such as altcoin mining and small-scale server clusters, contributes very little on the total energy consumption.

2.2 Model overview

We define a set of symbols as in Table 1. The total annual power consumption of HPC is

$$E = E_D + E_M, \quad (1)$$

which is the sum of energy consumed by data centers and cryptocurrency mining. In this model, some proportional relationships will be intentionally introduced to facilitate the subsequent models.

2.2.1 Energy consumption of data centers

The annual energy consumption of data centers *at full capacity* is straightforwardly defined as

$$E_{D\max} = T \sum_i P_i, \quad (2)$$

which is the product of the time of a year and the sum of maximum powers of all data centers. To get the case at *average utilization rate*, we simply multiply it by the average utilization rate $\bar{\rho}$:

$$E_{D\text{avg}} = \bar{\rho} E_{D\max} = \bar{\rho} T \sum_i P_i. \quad (3)$$

2.2.2 Energy consumption of cryptocurrency mining

During cryptocurrency mining, many miners join a network to solve blocks. The *hash rate* is a measure of the computational power of a machine, defined as the number of hash calculations performed per unit time. The total hash rate H of a network is the sum of the hash rates of all machines connected to the network, with a higher value representing a higher competition among miners.

A mining machine works at a specific power and holds a specific computational power, or hash rate. The ratio between the working power and its computational power gives its *efficiency* ε , the power required to achieve unit hash rate (in the unit of power per hash rate, or energy per hash).

To calculate the total power P in a network, we multiply the average efficiency $\bar{\varepsilon}$ among all mining machines with the network's total hash rate H (de Vries, 2020):

$$P = \bar{\varepsilon} H. \quad (4)$$

Over time, new mining machines with more advanced technologies are produced and slowly take up the market whilst old machines are still being used (economically speaking, they will be used until their outputs are lower than their operating costs). Therefore, the replacement process takes time; the average efficiency is always lower than the efficiency of the latest machine.

We define κ as the ratio between the average efficiency $\bar{\varepsilon}$ (among all machines) and the optimal efficiency ε_{opt} (of the current best model). Equation 4 can be rewritten as

$$P = \bar{\varepsilon} H = \kappa \varepsilon_{\text{opt}} H. \quad (5)$$

Finally, we multiply the power with the one-year time T to get the annual energy consumption:

$$E_M = PT = \kappa \varepsilon_{\text{opt}} HT; \quad (6)$$

this is both the *full capacity* level and the *average utilization* level since cryptocurrency mining is assumed to be always operating at the maximum utilization possible.

Table 2: Some records of data center information from *datacenters.com*. Only columns including data related to the model are displayed; names and addresses that are too long are truncated.

ID	Name	Address	Total space (sqft)	Total power (MW)
0	Equinix:, Tokyo, Japan	58992	6.75
1	KAO Data:, UK	40000	40
2	NorthC: Eindhoven, Netherlands	43056	4.5
3	DC2Scale:, France	15000	1.5
4	DataBank:, TX, USA	23000	0.675
5	Digital Realty:, Opfikon, Switzerland	79800	5
6	Ascenty,, São Paulo, Brazil	22604.21	34
7	EdgeConnex:, Calle Larga, Chile	116519.33	14
8	Centersquare:, MA, USA	201590	15
9	Digital Realty:, Tamil Nadu, India	196000	20.4
⋮				
4341	Equinix:, São Paulo, Brazil	52743	1.8

2.2.3 Total annual energy consumption

Combining the above equations, the total annual energy consumption at *full capacity* is expressed as

$$E_{\max} = E_{D\max} + E_M = T \sum_i P_i + \kappa \varepsilon_{\text{opt}} HT, \quad (7)$$

and the total annual energy consumption at *average utilization rate* is

$$E_{\text{avg}} = E_{D\text{avg}} + E_M = \bar{p}T \sum_i P_i + \kappa \varepsilon_{\text{opt}} HT. \quad (8)$$

2.3 Assessing the model using current data

To assess the accuracy of the energy consumption model, we substitute existing data into our model and compare the model’s output with the related existing aggregated statistics.

2.3.1 Data centers

We acquired data from *datacenters.com*, a website offering detailed information of data centers worldwide. 4342 data centers (Datacenters.com, 2024) are included, with information of each’s name, location, provider, space, total power, etc. Table 2 shows some of the results from the website.

We plotted the distribution of the data as in Figure 1; it is worth mentioning that there are two very significant outlier values (denoted yellow on the top) exceeding 10^6 megawatts. Obviously, it is improbable for those two data centers to consume about 13 times more power than the sum of all other ones (more intuitively, taking up almost 5% of the total human power consumption); therefore, the data is very possibly incorrectly recorded or in a wrong unit. We thereby removed the two, and now all data are in a reasonable range, as shown in Figure 2.

We then calculated the sum of the total powers of all data centers without the outliers, yielding $\sum_i P_i = 309,160\text{MW}$. Therefore, the annual energy consumption of data centers at *full capacity* is

$$E_{D\max} = T \sum_i P_i = 8,760\text{h} \times 309,160\text{MW} \approx \mathbf{2,708TWh}. \quad (9)$$

The 2024 Kubernetes Cost Benchmark Report (Cast.ai, 2024) reveals that the average utilization rate for Kubernetes server clusters is 13%. Since Kubernetes is widely used in data centers, this figure

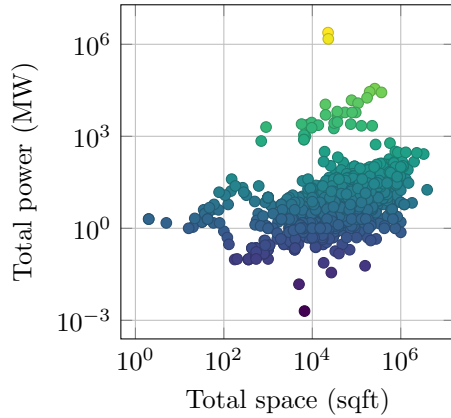


Figure 1: The original dataset.

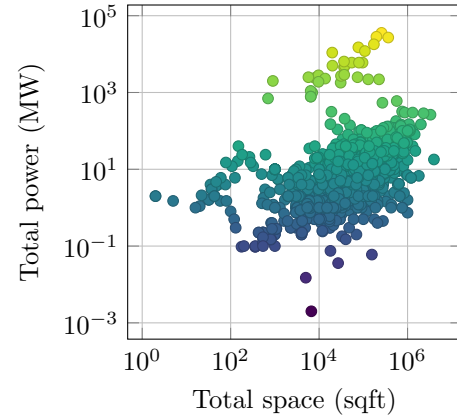


Figure 2: The dataset with outliers removed.

is highly representative for general data center utilization. Therefore, we get the annual energy consumption of data centers at *average utilization rate*:

$$E_{D_{\text{avg}}} = \bar{\rho} E_{D_{\text{max}}} = 13\% \times 2,708 \text{ TWh} \approx \mathbf{352 \text{ TWh}}. \quad (10)$$

This falls in the range of world data centers stats [**TODO** add stats here].

2.3.2 Cryptocurrency mining

Using the recent (November 2024) Bitcoin network data (Blockchain.com, 2024), the total hash rate $H \approx 750,000,000 \text{ TH/s}$. The currently most advanced mining machine is Antminer S21 (Neeti and Selvasundar, 2024), with required power of 3531W and a hash rate of 300TH/s, giving the efficiency

$$\varepsilon_{\text{opt}} = \frac{3531 \text{ W}}{300 \text{ TH/s}} = 11.77 \text{ J/TH}. \quad (11)$$

According to an examination of a real-world mine (Digiconomist, 2017), the actual power consumption of the mine is about 70% higher than the theoretical optimum, thus the multiplier $\kappa = 1.7$. This gives the total annual energy consumption of cryptocurrency mining:

$$E_M = \kappa \varepsilon_{\text{opt}} H T = 1.7 \times 11.77 \text{ J/TH} \times 750,000,000 \text{ TH/s} \times 8,760 \text{ h} \approx \mathbf{131 \text{ TWh}}, \quad (12)$$

which also falls in the range [**TODO** range].

2.3.3 Evaluation

[**TODO** whether necessary?]

3 HPC carbon emission model

In this section, we model the carbon emissions of all HPC activities based on the previous energy consumption model. [**TODO** write more]

We additionally defined several symbols in this model, as shown in Table 3.

Table 3: Definition of new symbols. The symbol $x \in \{F, N, R\}$ stands for the method of electricity production, respectively fossil fuels (F), nuclear energy (N), and renewable energy (R).

Symbol	Unit	Description
Ψ	t	Total annual carbon emission of all HPC activities.
Ψ_D	t	Total annual carbon emission of data centers.
Ψ_M	t	Total annual carbon emission of cryptocurrency mining.
$\Xi_{i,T}$	TWh	Total annual electricity production of country i .
$\Xi_{i,x}$	TWh	Annual electricity production of country i using method x .
M_x	t/TWh	Mass of CO ₂ created per unit of electricity produced using method x .
C_i	t/TWh	Carbon emission per unit electricity produced in country i .
$\zeta_{i,x}$	–	Proportion of electricity production of country i using method x .
$\zeta_{M,x}$	–	Proportion of electricity for crypto. mining produced using method x .
t	–	A particular year in A.D.

3.1 Assumptions

Assumptions in Section 2.1 are still used in this section. Here are the additional assumptions:

1. The energy mix of a data center is exactly the same as the energy mix of the country it is in.
Justification: Most data centers are directly connected to the national electrical grid, which's electricity comes from a mixed source. The detailed investigation of the energy mix at each location is far beyond the scope of this model; we hence simplify this to a uniform national averaged distribution of electricity source.
2. All data, such as total power and energy mix, are discretized for each year, i.e., within one year, the data is uniform and does not change.
Justification: This assumption is to match most of the data that comes in the form of annual statistics. Since the model aims to provide insights on the current state and long-term future predictions of carbon emissions, ignoring the temporal variations within one year simplifies the model and makes it easier to be implemented from data.
3. Each type of energy source has a same carbon emission per unit energy everywhere.
Justification: In reality, a same type of energy source may have different carbon emissions in different places (e.g., the difference in the effectiveness of wind power due to different natural conditions at different sites). To accurately model this, we would need emissions data at province level or even city level; such differences are ignored because they would make the model too complex while not creating any significant change.

3.2 Carbon emissions at a fixed time

Similarly, we calculate the total emissions through adding up the two sectors. We categorize all energy sources that are used to produce electricity into three types: fossil fuels (F), nuclear energy (N), and renewable energy (R).

3.2.1 Carbon emission of data centers

Suppose the electricity production of country i (with a total of n countries) from energy type x is $\Xi_{i,x}$, we have the mix and total energy for the production of electricity of all countries:

$$\mathbf{\Xi} = \begin{bmatrix} \Xi_{1,F} & \Xi_{1,N} & \Xi_{1,R} \\ \Xi_{2,F} & \Xi_{2,N} & \Xi_{2,R} \\ \vdots & \vdots & \vdots \\ \Xi_{n,F} & \Xi_{n,N} & \Xi_{n,R} \end{bmatrix}; \quad \mathbf{\Xi}_T = \begin{bmatrix} \Xi_{1,T} \\ \Xi_{2,T} \\ \vdots \\ \Xi_{n,T} \end{bmatrix} = \begin{bmatrix} \Xi_{1,F} + \Xi_{1,N} + \Xi_{1,R} \\ \Xi_{2,F} + \Xi_{2,N} + \Xi_{2,R} \\ \vdots \\ \Xi_{n,F} + \Xi_{n,N} + \Xi_{n,R} \end{bmatrix} = \mathbf{\Xi} \mathbf{1}. \quad (13)$$

Therefore, we define the proportions of the electricity produced from each energy source in all electricity production for each country, $\zeta \in \mathbb{R}^{n \times 3}$, as

$$\zeta_{i,x} = \frac{\Xi_{i,x}}{\Xi_{i,T}} = \frac{\Xi_{i,x}}{\Xi_{i,F} + \Xi_{i,N} + \Xi_{i,R}}, \quad (x \in \{F, N, R\}). \quad (14)$$

For the mass of CO₂ created per unit of electricity

$$\mathbf{M} = \begin{bmatrix} M_F \\ M_N \\ M_R \end{bmatrix}, \quad (15)$$

the energy mix proportions ζ are the weights on each energy type; through linear combination, we calculate the weighted average carbon emission per unit electricity in each country as

$$\mathbf{C} = \zeta \mathbf{M}. \quad (16)$$

Since we already have the detailed information of all data centers, we can calculate the total capacity power of data centers in each country as

$$\mathbf{P} = \begin{bmatrix} \sum P_i \text{ in country 1} & P_i \\ \sum P_i \text{ in country 2} & P_i \\ \vdots & \\ \sum P_i \text{ in country n} & P_i \end{bmatrix}, \quad (17)$$

giving the total *real* energy consumption due to data centers in each country

$$\mathbf{E}_D = \bar{\rho} T \mathbf{P}, \quad (18)$$

where $\bar{\rho}$ is the average utilization rate and T is the time period (one year).

Through linear combination, we deduce the total annual carbon emissions worldwide due to data centers by calculating the multiplication of annual energy and carbon emission per energy:

$$\Psi_D = \mathbf{E}_D^\top \mathbf{C} = \bar{\rho} T \mathbf{P}^\top \zeta \mathbf{M}. \quad (19)$$

3.2.2 Carbon emission of cryptocurrency mining

The industry of cryptocurrency mining is separated mainly for the difficulty in identifying country-wise energy mix. However, the energy mix of the industry as a whole is feasible. We could model the industry in a way similar to modeling another country.

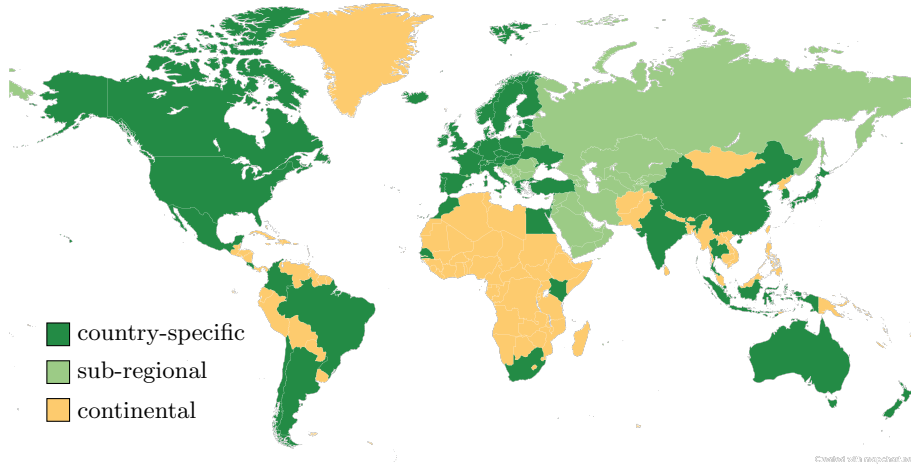


Figure 3: Detailed energy mix data availability in each country.

Previously, we have calculated that the annual energy consumption because of cryptocurrency mining is $E_M = \kappa \varepsilon_{\text{opt}} HT$. Similarly, we can calculate the carbon emission per unit of electricity as

$$\mathbf{C}_M = \boldsymbol{\zeta}_M \mathbf{M}, \quad \left(\boldsymbol{\zeta}_M = \begin{bmatrix} \zeta_{M,F} & \zeta_{M,N} & \zeta_{M,R} \end{bmatrix} \right), \quad (20)$$

and therefore, the total carbon emission due to cryptocurrency mining can be defined:

$$\Psi_M = E_M \mathbf{C}_M = \kappa \varepsilon_{\text{opt}} HT \boldsymbol{\zeta}_M \mathbf{M}. \quad (21)$$

3.2.3 Results and evaluation

We add up the two sectors to find the total annual carbon emission of HPC:

$$\Psi = \Psi_D + \Psi_M = \bar{\rho} T \mathbf{P}^\top \boldsymbol{\zeta}_M + \kappa \varepsilon_{\text{opt}} HT \boldsymbol{\zeta}_M \mathbf{M}. \quad (22)$$

We use the World Energy Balances 2024 Highlights (free extract) dataset (IEA, 2024) offered by the International Energy Agency (IEA). The dataset includes comprehensive energy mix data over the past fifty years (since 1971) in many countries and broader regions, including total electricity production and production from fossil fuels, nuclear energy, and renewable energy. Figure 3 shows the most accurate data available from the dataset for each country. Countries having the most data centers, such as the U.S., western European countries, China, Japan, Singapore, etc., all have their own country-specific data available. Therefore, the dataset is highly reliable for modeling the energy structure used by data centers worldwide.

References

- Blockchain.com (2024). Bitcoin total hash rate.
- Cast.ai (2024). *2024 Kubernetes Cost Benchmark Report*. Cast.ai.
- Datacenters.com (2024). Data center locations: Top cities, states, countries and regions.
- de Vries, A. (2020). Bitcoin’s energy consumption is underestimated: A market dynamics approach. *Energy Research & Social Science*, 70.
- Digiconomist (2017). *A Deep Dive in a Real-World Bitcoin Mine*. Digiconomist.

Table 4: An extract from the IEA World Energy Balances 2024 Highlights dataset.

Entry	Country	Energy Type	Electricity output (GWh)				
			1971	1972	...	2022	2023 [†]
1	Australia	Fossil fuels	41201	43730	...	187536	181363
2	Australia	Nuclear	0	0	...	0	0
3	Australia	Renewable sources	11844	11853	...	83282	92216
4	Australia	Total	53045	55583	...	270818	273579
5	Austria	Fossil fuels	11752	11939	...	13509	10205
6	Austria	Nuclear	0	0	...	0	0
7	Austria	Renewable sources	16450	16986	...	50432	59397
8	Austria	Total	28202	28925	...	64713	70301
⋮							
244	World	Total	5256947	5698376	...	29143442	–

IEA (2024). World Energy Balances 2024 Highlights (free extract).

Neeti and Selvasundar, V. (2024). *Best Bitcoin Miners & Crypto Mining Rigs: Ultimate Guide for ASIC Miners In 2024*.