Team Control Number

**15248**

Problem Chosen

**B**

**2024**

**HiMCM**

**Summary Sheet**

---


## Summary

With great "power" comes great responsibility. In recent years, the world has seen a massive increase and progression of high-powered computing (HPC). Now a critical component in AI industries, the rise of HPCs signifies a new area of mass energy consumption, which can lead to serious environmental impacts. In this paper, we introduce a multi-step environmental impact model to evaluate the effects of HPC on the environment, apply and adapt our model to various scenarios, and develop a set of recommendations to mitigate these impacts along with a proposal letter to the United Nations Advisory Board to urge for more emphasis on their consideration towards high-powered computing.

First, we collect data from two different datasets, one with a more narrowed scope (TOP500), and one with an inclusive and broad scope (a global data center energy consumption dataset). We preprocess our data using methods such as linear interpolation and grouping by average. We utilize Holt's linear trend method to forecast values for TOP500 which considers the futuristic growth of HPC. We leverage the geographical region consideration, including North America, Asia, and Europe, from TOP500 and scale the quantities by a dataset ratio derived with our second dataset to obtain more accurate global HPC environmental impacts. From our data, we notice that HPC capabilities will utilize around 362 TWh of electricity in 2024, and a forecast of 1593 TWh in 2030.

To build our HPC environmental impact model, we first account for a dynamical energy mix by developing a **recursive function model** based on our logistic function for renewable energy growth. We **distinguish energy mix ratios between each geographical region** to provide a more realistic approach to these regions' varying energy source priorities. Next, we develop a **multivariate carbon emissions model** and leverage our dataset ratio to consider all data centers worldwide. Finally, to compose a clearer definition of environmental impact, we include two environmental factors in our model: temperature impact and air quality impact. We build our temperature impact model based on the theory of **radiative forcing,** and our air quality impact model based off pollutant mechanisms. After constructing our model, we analyze the strengths and limitations. We later expand on our HPC environmental impact analysis by developing a new model for HPC data center land usage using a U.S. data center dataset, developing a relationship between power usage and land usage. We test **polynomial, linear, and logarithmic regression** and use $R^2$ metrics to choose our best-fit.

Next, we apply our models to a variety of different scenarios. We use our model to analyze the **increase in energy demand** from other sectors by adding a **cost parameter to model the increase in energy costs.** We notice that higher costs for certain energy sources due to demand slow their growth in energy mixes, eventually reaching a dropping threshold where our model emphasizes priority over other energy sources. We then use our model to forecast the impacts of HPC on the environment in 2030, creating a **best-case and worst-case scenario to provide realistic bounds**. Furthermore, we analyze the case of **increasing the portion of renewable energy composition**, measuring the carbon emissions reductions from 25% to 50% to 75% to 99% renewable energy composition along with addressing the possibility of a 100% composition. Lastly, we use our findings to create a set of recommendations to mitigate the environmental impact of HPC and draft a proposal letter to the United Nations Advisory Board to urge for more emphasis on their consideration towards high-powered computing.

# Contents

# 1 Introduction

## 1.1 Background

In recent years, the need for high-powered computing (HPC) has increased dramatically. The main takeaway from HPC is its capabilities on processing and solving complex calculations at a high speed. HPC computing allows for advanced algorithms such as generative AI and data analytics by using parallel processing. This technology employs the processing power of numerous supercomputers to complete complex tasks. As fields such as AI, data science, and cryptocurrency mining continue to grow in the 21$^{st}$ century, the demand on high-powered computing becomes more prominent each day. However, the expanding usage of HPC capabilities requires massive amounts of energy. In the AI industry alone, there has been serious concern about the future shortage of energy due to a rapid need for energy consumption. According to forecasts of the global tech research company Gartner, in the next two years there will be a 160% growth in energy consumption from major AI data centers [1].

AI data centers are only one use of HPCs. Numerous other industries are now quickly increasing their utilization of these supercomputers, all of which require high energy consumption. Consequently, high energy consumption takes a toll on the environment. Energy sources such as fossil fuels can emit concerning amounts of pollution and greenhouse gas emissions. But how much exactly does global high-power computing impact the environment? In this paper, we utilize mathematical modeling to evaluate the environmental impacts of HPC capabilities both regionally and worldwide, get a glimpse of the issue within the near future considering multiple forecast scenarios, and propose a set of recommendations to mitigate the environmental impacts of HPC based on our results.

## 1.2 Problem Restatement

High-powered computing reaps a range of resources from the environment. From mass consumption of energy leading to carbon emissions, to depleting natural resources such as silicon and taking up massive areas of land. Thus, it is evident that the environmental impacts of HPC within recent years must be studied and analyzed. In this paper, we split the prompt into 5 problems:

**Problem 1:** Describe the scope of this problem in terms of the annual energy consumption of the HPC capabilities worldwide considering both full capacity and average utilization rates.

**Problem 2:** Develop a comprehensive model that describes the environmental impacts of HPC use. It should encompass the major energy mixes during the production of energy for HPCs. Based on this, the model should then fully describe the main areas of the environment that are affected e.g. thermal impact, air quality, etc.

**Problem 3:**

- **3a.** Address the potential effects that technological growth of HPC, increasing energy demand from other sectors, and varying energy mixes have on our model.
- **3b.** Apply our model to forecast realistic bounds of the issue in 2030, analyzing the best and worse-case scenarios.

**Problem 4:**

- **4a.** Use a model to predict how carbon emissions would decrease in response to rises in the use of renewable energy. Discuss the feasibility and possible effects there would be if all HPCs used a 100% renewable energy source.
- **4b.** Develop a new model or refine the current model to explore how would other areas of the environment (water usage, E-waste, resource depletion, land use, air quality, etc) would be impacted by the growth in HPC usage. Explain in detail why you choose a specific area for analysis.

**Problem 5:**

- **5a.** By providing both technical and policy-oriented solutions, provide a set of recommendations on mitigating the environmental impacts of HPC.

**5b.** Take one of your recommendations and incorporate it into your models to show its feasibility and benefit.

**5c.** Urge the UN to include a more comprehensive section on the impacts of HPC in their scheduled goals for 2026. Use your findings and recommendations to write a two-page letter to the UN advisory board.

## 1.3 Assumptions and Justification

**Assumption 1:** The data collected from the data sets used in this work are both true and valid.
**Justification:** Our model utilizes datasets to both forecast and calculate quantities such as carbon emissions. These data from these datasets were collected from large surveying and extensive research, thus we assume that our data given is true and valid.
**Assumption 2:** There will be no unexpected sudden changes regarding HPC energy consumption in the near future.
**Justification:** Our model uses historical data; thus it may not guarantee a reasonable accuracy in the future if sudden events or turning points occur in the future, which is nearly impossible to foresee.
**Assumption 3:** Renewable energy is expected to continue growing in the near future.
**Justification:** Given the current circumstances of renewable energy advocation via governments, policies, awareness programs, we assume that renewable energy will continue to grow in the near future, as our model utilizes a logistic growth function for renewable energy composition.
**Assumption 4:** Power generation methods other than fossil fuels do not produce $SO_x$ or $NO_x$ emissions.
**Justification:** Fossil fuels produce a very high proportion of $SO_x$ or $NO_x$ emissions used in our air quality impact model, thus it is reasonable to assume that only fossil fuels produce these pollutants.

# 2 Model Preparation

## 2.1 Understanding the Issue and Scope

We develop a concrete understanding of the problem and its scope, and accurately prepare and process our data, serving as a foundation and preliminary analysis for constructing complex models later on. To understand the scope of the issue, we first introduce a few inherent concepts. To measure the energy consumption of HPC capabilities, we acknowledge the following physics principle:

$$W = P\Delta t \tag{1}$$

Where W = work, P = power, and $\Delta t$ is total time elapsed. Thus, in the context of HPC energy consumption, the total hours in a year can be used to calculate the total energy consumption (in terms of kWh) at full capacity annually:

$$E = P_f \cdot 8760 \tag{2}$$

Where E is the total energy consumption, $P_f$ is the full capacity, and 8760 is the total number of hours per year. The total capacity represents the maximum possible energy consumption by HPC capabilities annually. However, many HPC data centers do not operate at full capacity throughout the year due to various reasons such as maintenance or redundancy. Taking this into consideration, a more realistic approach is to scale the full capacity by the utilization rate, which is found by dividing the actual output by the full capacity over the same period. Thus, the total energy consumption by HPC capabilities annually at average utilization can be represented as:

$$E = U \cdot P_f \cdot 8760 \tag{3}$$

Where U is utilization rate. To obtain the total energy consumption of HPC data centers globally, **we utilize a comprehensive global dataset** which takes estimates from numerous different works to provide estimates [2]. In addition, **we attempt to break the problem down more precisely**. To calculate carbon emissions and environmental impacts, **geographical region plays a crucial role** in determining energy mixes, and using global data directly may be inaccurate. In addition to the global dataset, we condense our scope and also utilize data from TOP500, which provides comprehensive datasets for the top 500 most powerful commercially available HPC computer systems in the world and represents the peak of high-powered computing. Its growth over the years represents the face of the industry and its data is **categorized by continent**. After modeling by continent and summing for a global approximate, we can **multiply this by a ratio between the two datasets** to **capture quantities for all data centers around the world** which will be discussed later. In this paper, our main scope for environmental impact is **temperature change, air quality, and land use.** We support our TOP500 dataset choice based on other research regarding this topic which consider TOP500 to identify the trends of the whole domain [3].

## 2.2 Data Preprocessing and Visualization

In this section, we preprocess and visualize both datasets utilized in this paper. Every June and November, TOP500 releases an updated dataset, so we first look at data cutoffs during June and November for each year and then calculate an average for each year. Due to the difficulty of global data collection, some computer systems' power usage was not submitted to the database, imposing a limitation on our data given. However, the inconsistency is not as severe, as data centers who choose not to submit power usage data usually maintain this behavior over multiple periods of time. To calculate total energy consumption per year, we multiply the power usage by the total amount of hours in a year, as illustrated in the previous equations.

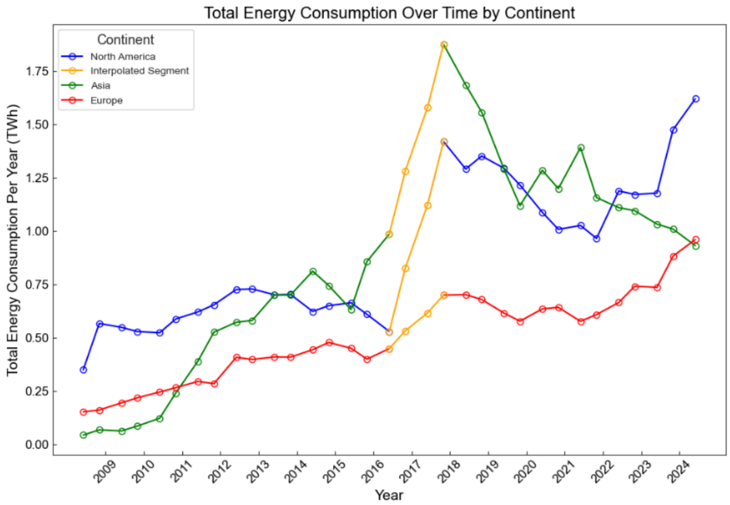


*Figure 1: TOP500 Energy Consumption by Continent with Interpolated Values*

After analyzing the TOP500 dataset and visualizing the total power usage within the last 17 years, we chose to intentionally omit the power usage of November 2016 and June 2017 because we noticed abnormal quantities which drastically skew the data. To achieve a better understanding of TOP500 energy consumption, we grouped the data by geographical region, specifically by continent. TOP500 provided data for North America, South America, Oceania, Europe, Asia, and Africa. After analyzing the data, we removed the power usage and energy consumption of South America, Oceania, and Africa from our consideration because they were statistically insignificant compared to other continents. Our decision to group data by geographical region will be highly relevant when constructing our model later due to considerations for energy mixes.

Since we omitted data for November 2016 and June 2017, there is a gap in our data. To handle the missing values within our data, we perform linear interpolation for each individual continent to achieve a continuous dataset. Linear interpolation is a mathematical method of deriving new estimated values by using two or more pre-existing data points and curve-fitting via linear polynomials [4]. From the graph, we can see that over the span of 15+ years, TOP500 energy consumption in all three continents has risen by a fine amount.

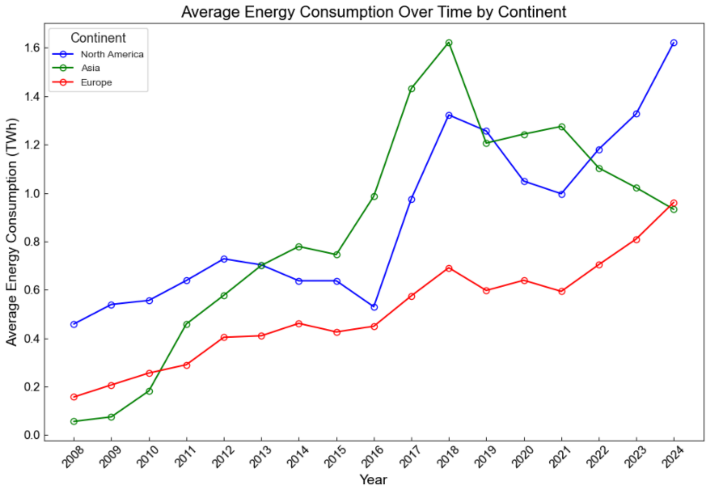


*Figure 2: TOP500 Energy Consumption by Continent After Averaging*

After we address the missing values in 2016 and 2017, since data is given for November and June for each year (excluding 2024), we organize our data by taking the average for November and June during each year to represent the total energy consumption that year. In 2024, our TOP500 data indicates that North America, Asia, and Europe will consume 1.62TWh, 0.93TWh, 0.96TWh of energy within TOP500's scope, and 4.41TWh globally.

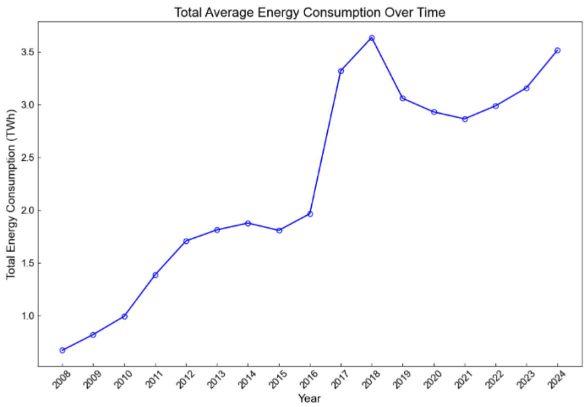


*Figure 3: TOP500 Total Average Energy Consumption by Year*

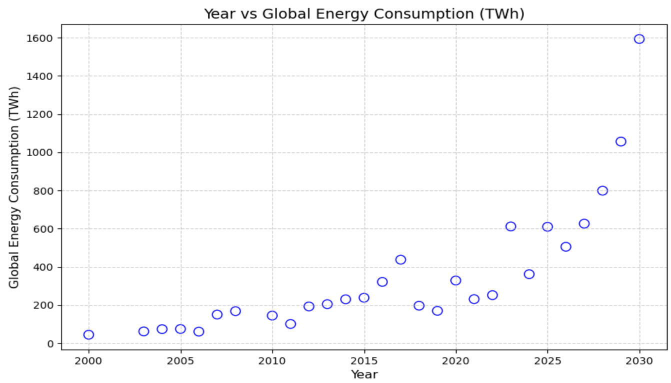


*Figure 4: Global Datacenter Dataset Average Energy Consumption by Year*

The global data center dataset is our second data collection. In Figure 4, we visualize the global dataset and average the values for years that consist of more than one energy consumption estimate since multiple estimates were taken from different works. It can be seen in both figures that the overall usage within recent years is much higher than over a decade ago, highlighting current concerns of HPC energy consumption. It is estimated that **HPC data centers globally will consume 362TWh in 2024**. Additionally, from our global dataset, it is estimated that in 2030 HPC data centers will consume 1593TWh, which is around a 340% increase.

We use TOP500 as it is advantageous for regional analysis of energy consumption, allowing us to consider varying energy mixes. On the other hand, the global HPC dataset is a much better representation for all data centers around the world as TOP500 only considers the "tip of the iceberg". To leverage the advantage for both datasets, we will calculate ratios between the global dataset and the TOP500 dataset in the next section which will be applied in our model. We will first forecast TOP500 energy consumption up to 2030, as the global dataset already consists of estimated forecasts up to 2030.

## 2.3 Forecasting HPC Energy Consumption

Before introducing our model, we first take a glance at what TOP500's energy consumption will look like in the near future, which considers the future growth of HPC addressed in **problem 3a** and will later be used to expand upon our model by forecasting future scenarios. Instead of considering HPC growth after building our model, we do this first as it is a critical component to forecasting HPC environmental impact.

To explore what our model will look like in the future taking in account of the future growth of HPC, we utilize time-series forecasting, specifically Holt's Linear Trend Method to forecast TOP500 energy consumption up to 2030. Holt's Linear Trend Method further extends simple exponential smoothing. Exponential smoothing accounts for the "level" of the time series, making it effective for data without a clear upward or downward trend. However, when the data exists a trend over time, such as an increasing pattern as seen in the three continents, Holt's model is a more suitable choice. In addition to our model selection process, we tested the Autoregressive Integrated Moving Average (ARIMA) model, which showed no variation (e.g., flat or nearly constant forecasts) for forecasts, indicating that the model doesn't detect enough trend or seasonality in the data to generate forecasts with sufficient variation. Holt's model consists of three main components: the level equation, the trend equation, and the forecasting equation [5] .

**Level Equation:**

$$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \tag{4}$$

In the level equation, the level at time $t$ is represented as $l_t$, where $\alpha$, the smoothing parameter ($0 \le \alpha \le 1$), is accounted for along with the previous trend estimate $b_{t-1}$ and the actual observed value at time $t$.

**Trend Equation:**

$$b_t = \beta * (l_t - l_{t-1}) + (1 - \beta *) b_{t-1} \tag{5}$$

In the trend equation, $b_t$ is used to represent the trend at time $t$, and a smoothing parameter $\beta *$ is applied with bounds $0 \le \beta * \le 1$.

**Forecasting Equation:**

$$\hat{y}_{t+k} = l_t + k b_t \tag{6}$$

In the forecasting equation, both the level and trend are combined to forecast $k$ steps ahead. To optimize parameters $\alpha$ and $\beta *$ in our model, we utilized the Python module statsmodels, which automatically selects optimal smoothing parameter values by minimizing the sum of squared errors (or SSE) of the fit, which is a method for parameter optimization.

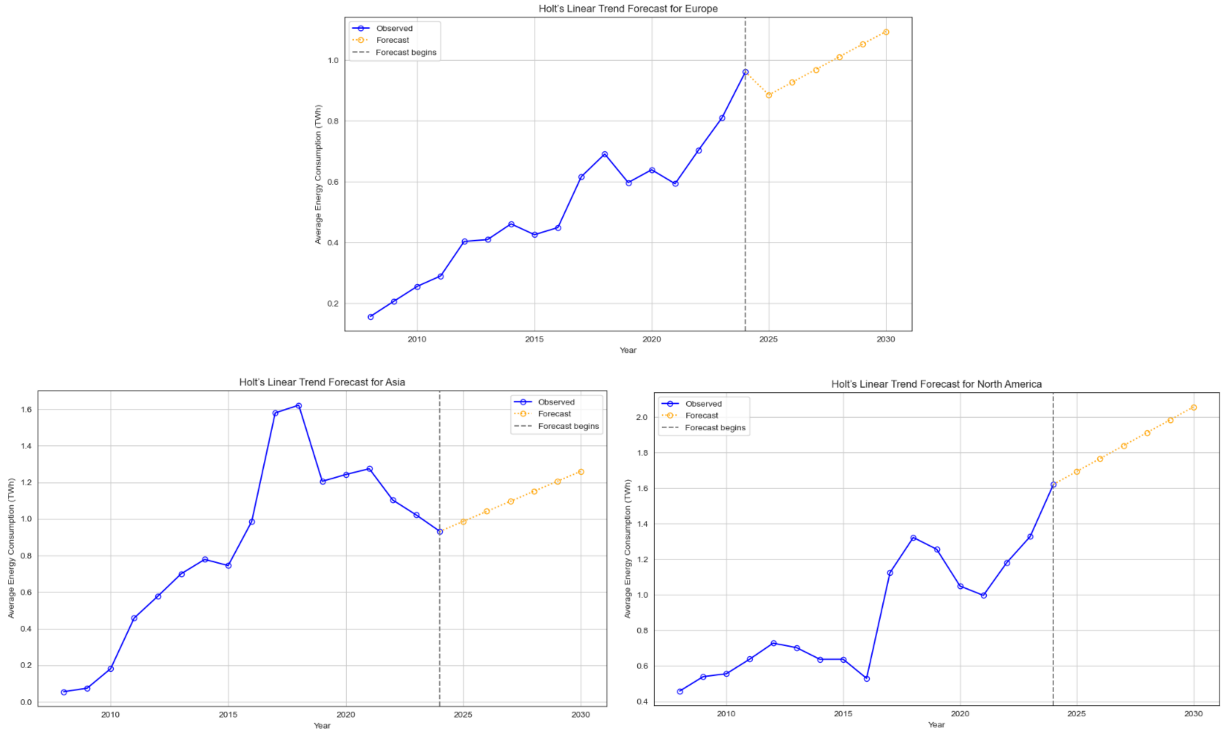


*Figure 5: Forecasts of TOP500 Energy Consumption by Continent to 2030*

Our forecast model predicts that in 2030, Asia's, North America's, and Europe's top-performing HPC capabilities will consume approximately **1.26 TWh, 2.06 TWh, 1.09 TWh** respectively. To develop a global TOP500 forecast, we aggregate the data from the three continents and use Holt's Linear Trend model for the global data.

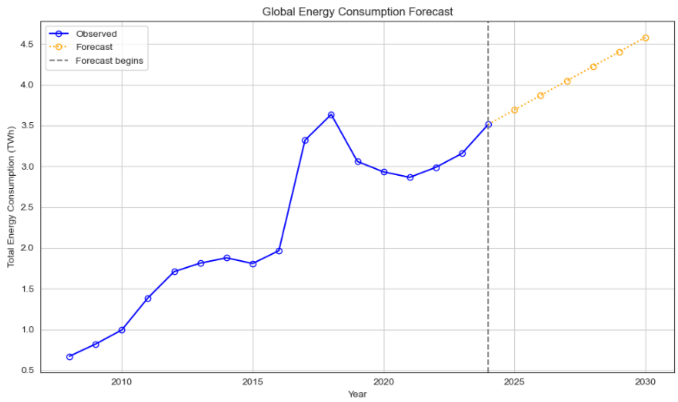


*Figure 6: Forecasts of TOP500 Global Energy Consumption to 2030*

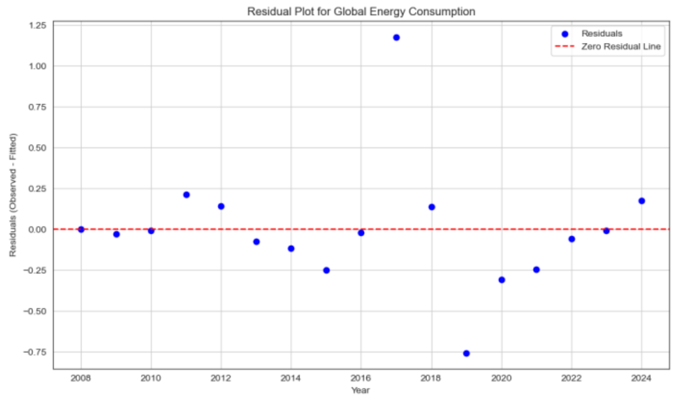


*Figure 7: Residual Plot for Holt's Linear Trend Method on TOP500*

In 2030, Our model predicts that global TOP500 HPC capabilities will consume approximately 4.41 TWh, which is a 25.6% increase from 2024. To evaluate how well our model fits the observed data, we assess our model with the Mean Squared Error (MSE) and R-squared value. Our model obtained an MSE of 0.14 and a $R^2$ value of 0.85. Applying time series forecasting on our dataset allows us to forecast future TOP500 energy consumption, however one limitation of our forecasting model is a relatively small dataset size.

## 2.4 The Dataset Ratio

To leverage the advantages of both datasets, we will calculate ratios between the global dataset and the TOP500 dataset. Specifically, this dataset ratio will be useful since we will use TOP500 data for modeling regional energy consumption (to better address energy mixes) and scale the quantities by our ratio to obtain quantities relative to all data center energy consumption.

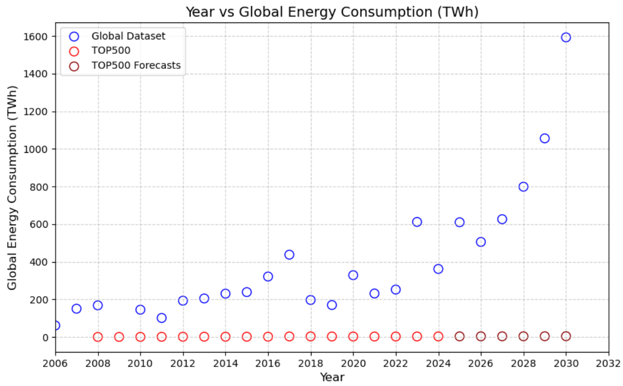


*Figure 8: Global Energy Consumption for Both Datasets*

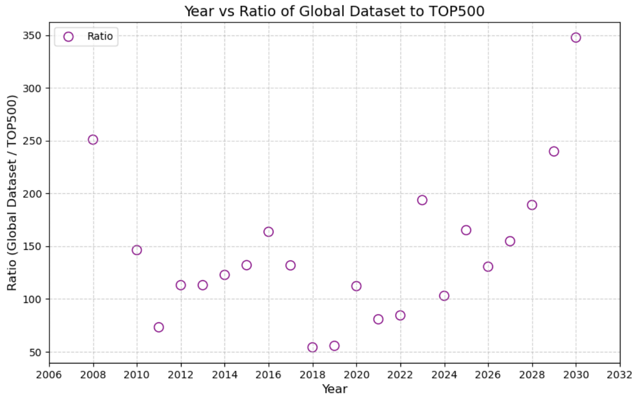


*Figure 9: Dataset Ratios from 2008 to 2030*

As shown above, Figure 8 depicts both datasets which also includes our TOP500 forecasts from earlier. In Figure 9, the ratio of our global data center data to TOP500 data is plotted. Since this ratio changes over time, we will use $D(t)$ to represent our dataset ratio during year $t$.

# 3 Environmental Impact Model

## 3.1 Model Architecture

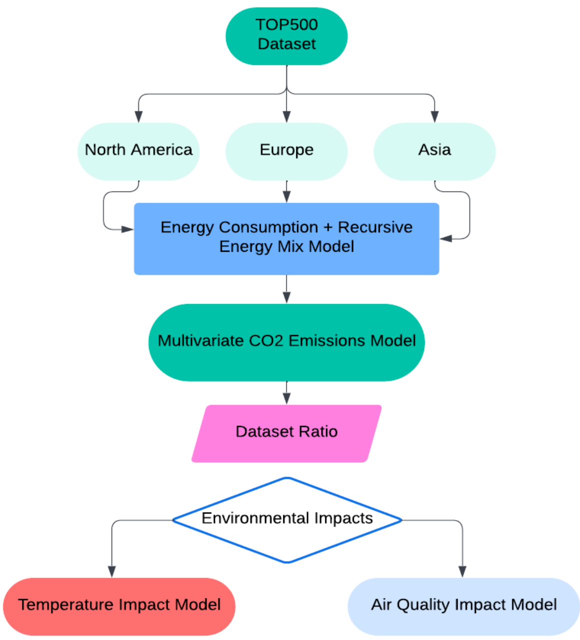


*Figure 10: Model Architecture*

In this section, we propose a mathematical model to determine the environmental impact of HPC capabilities. Since different energy sources that fuel HPC capabilities can have varying effects on the environment, we account for dynamical energy mixes by developing a recursive function model based on renewable energy logistic growth. We distinguish energy mix ratios between each geographical region to provide a more realistic approach to the varying energy source priorities in these regions. Moreover, we develop a multivariate carbon emissions model and leverage our dataset ratio to consider all data centers across the world. Finally, to compose a clearer definition of environmental impact, we include two factors in our model: temperature impact and air quality impact.

## 3.2 Modeling Dynamical Energy Mixes

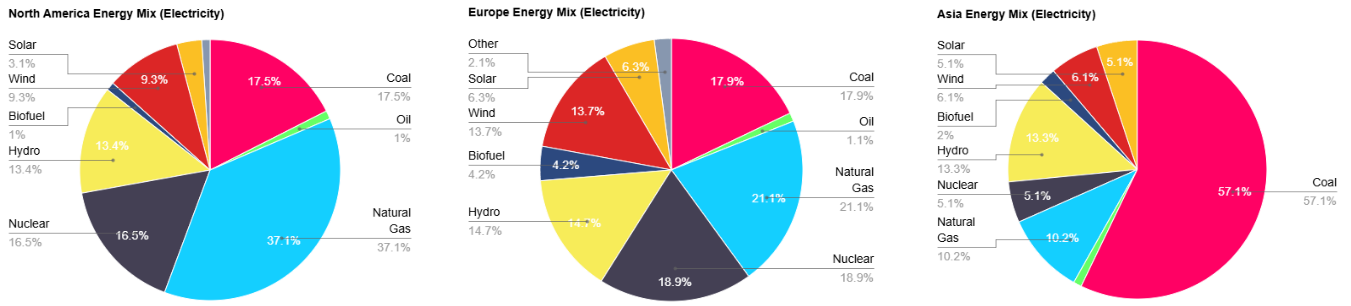


*Figure 11: 2024 Energy Mix of Electricity Generation from North America, Europe, and Asia*

To obtain the 2024 energy mix of electricity generation of the three continents, we utilize data collected by the International Energy Agency (IEA) [6]. The core energy sources that we consider in our model are: coal, oil, natural gas, nuclear, hydropower, solar, wind, and biofuel. Other sources of energy were in extremely low amounts relative to these eight and are thus omitted from our consideration. This section considers the potential for varying energy sources and mixes as addressed in **problem 3a** instead of doing this after building our model, to provide a robust and accurate model and estimations.

However, energy mixes are dynamic and can vary with respect to time due to a variety of factors. One important factor is the push for a more renewable energy mix via government policies. One example of this is the Renewable Portfolio Standards in the United States, which advocates for a growth in renewable energy composition [7]. Another example is the European Green Deal, setting a goal of becoming the first climate-neutral continent by targeting for renewable energy adoption across its member states [8]. In previous literature, energy mixes have been modeled based on a logistic relationship [9]. Thus, to account for energy mix dynamics and the increase in renewable energy composition throughout recent years, we propose our own logistic function for each continent to model the proportions for renewable energy sources:

$$\psi_{NA}(t) = \frac{1}{1 + e^{-0.16746558(t-22)}} \tag{7}$$

$$\psi_{Asia}(t) = \frac{1}{1 + e^{-0.17002344(t-22)}} \tag{8}$$

$$\psi_{Europe}(t) = \frac{1}{1 + e^{-0.07525293(t-22)}} \tag{9}$$

For each continent, $\psi(t)$ represents the renewable energy proportion within the total energy mix $t$ years after 2008 (2008 is used as an initial benchmark given our TOP500 dataset's time frame). The parameters of the logistic functions were calculated with energy mix data from 2024 (Figure 11). To emphasize the regional differences in energy mix, we preserve the 2024 ratios for each continent between the renewable energy sources and between the non-renewable energy sources, while using logistic functions (7) (8) (9) to account for the growth in renewable energy as a whole. We consider solar, wind, biofuel, and hydropower as renewable energy sources. Our choice for the logistic function maintains both a degree of simplicity and applicability as it captures the rapid initial adoption followed by gradual stabilization as renewables attempt to reach towards 100%. Additionally, we let 2030 (or t = 22) be the inflection point, which is the time at which the renewable energy proportion is expected to reach 50% [10]. In 2024, North America, Asia, and Europe have renewable energy proportions of 26.8%, 26.5%, and 38.9% respectively. The ratio for renewable energy source $i$ for continent $j$ can be represented by a recursive function:

$$R_{i,j}(t) = \frac{\psi_j(t)}{\psi_j(t-1)} R_{i,j}(t-1) \tag{10}$$

As renewable energy proportions increase, nonrenewable energy proportions must decrease. Since the sum of all energy mix ratios should be equal to 1, the ratio for nonrenewable energy source $i$ for continent $j$ can be represented by a separate recursive function:

$$R_{i,j}(t) = \frac{1 - (R_{hydro,j}(t) + R_{solar,j}(t) + R_{wind,j}(t) + R_{biofuel,j}(t))}{R_{coal}(t-1) + R_{oil}(t-1) + R_{natural}(t-1) + R_{nuclear}(t-1)} R_{i,j}(t-1) \tag{11}$$

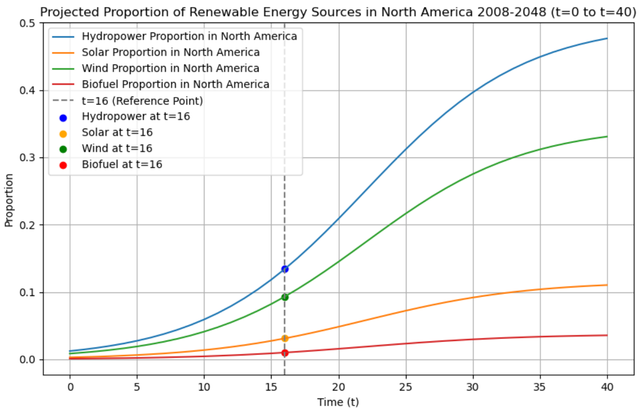


*Figure 12: Projected Proportions of Renewable Energy Sources in North America*

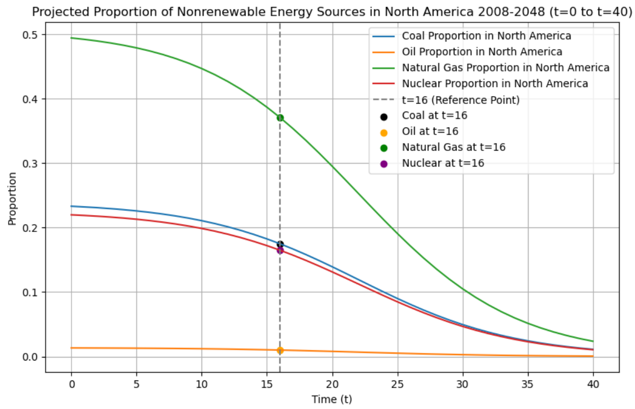


*Figure 13: Projected Proportions of Nonrenewable Energy Sources in North America*

Shown above, Figure 12 represents our recursive function for renewable energy source ratios applied onto North America. Figure 13 represents our recursive function for nonrenewable energy source ratios applied onto North America, and the dotted line on both figures represents North America's 2024 energy mix.

### 3.3 Modeling Carbon Emissions

To model the total carbon emissions throughout the year, we construct a multivariate model with respect to the time variable that considers four main factors: the energy consumption of each continent in TOP500, the energy mix ratios for each energy source with respect to their continent, and the carbon intensity for each source, and our dataset ratio. To calculate the total carbon emission from TOP500 HPC capabilities for each continent, we construct the following multivariate equation generalized for continent $j$:

$$\begin{aligned} C_j(t) = E_j(t) \cdot (&I_{coal}R_{coal,j}(t) + I_{oil}R_{oil,j}(t) + I_{natural}R_{natural,j}(t) \\ &+ I_{nuclear}R_{nuclear,j}(t) + I_{hydro}R_{hydro,j}(t) + I_{solar}R_{solar,j}(t) \\ &+ I_{wind}R_{wind,j}(t) + I_{biofuel}R_{biofuel,j}(t)) \end{aligned} \tag{12}$$

Where $C_j(t)$ is the TOP500 carbon emissions for a continent, $E_j(t)$ is the TOP500 energy consumption for a continent, $I$ is the carbon intensity for each energy source, and $R(t)$ is the energy mix ratio with respect to the energy source for a continent. Now, utilizing our **dataset ratio** between our global dataset and TOP500 for each year $D(t)$ and combining the equation above for each continent, we obtain:

$$CO_2(t) = D(t) \cdot (C_{NA}(t) + C_{Asia}(t) + C_{Europe}(t)) \tag{13}$$

Where $CO_2$ represents global total HPC energy consumption. Condensing the entire equation, we obtain:

$$CO_2(t) = D(t) \cdot \sum_j E_j(t) \cdot (\sum_i I_i \cdot R_{i,j}(t)) \tag{14}$$

Where j is the index that represents each continent and i is the index that represents each energy source.

| **Energy Source** | **Carbon Intensity ($gCO_2$ per kWh)** |
|---|---|
| Coal | 1040 |
| Oil | 1080 |
| Natural Gas | 490 |
| Nuclear | 12 |
| Hydropower | 24 |
| Solar | 48 |
| Wind | 12 |
| Biofuel | 230 |

*Table 1: Carbon Intensity of Each Energy Source*

Since each energy source has different magnitudes of contribution on carbon emissions, we utilize the carbon intensity values of each energy source listed in Table 1 obtained from external sources [11][12]. From the table, we can see that nonrenewable energy sources such as coal and oil have significantly higher carbon intensities than renewable energy sources, as they are considered fossil fuels.

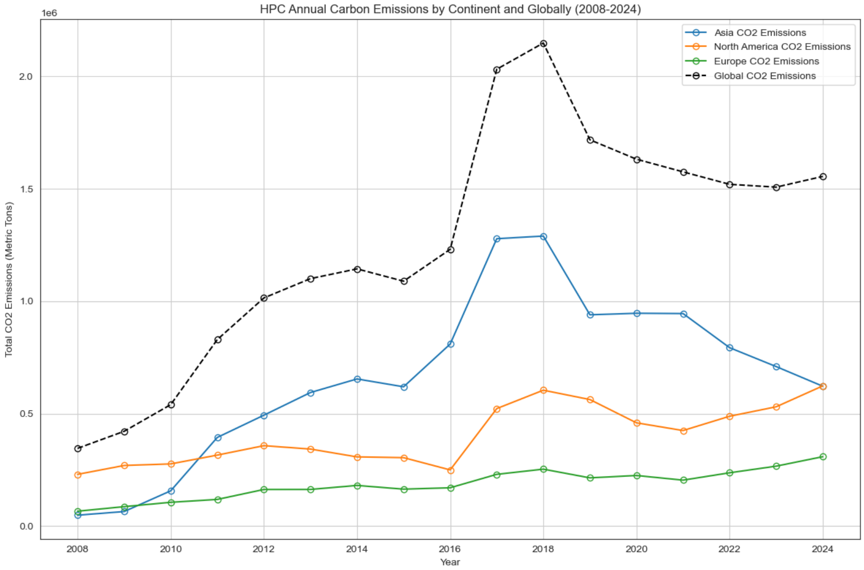


*Figure 14: TOP500 $CO_2$ Emissions Regionally and Globally*

With our carbon emissions model, we estimate that TOP500 HPC capabilities will emit approximately 1,555,134 metric tons CO2 in 2024. Scaling this with our dataset ratio $D(t)$ for 2024, we obtain that approximately 1,555,134*102.979855 = **160,147,474 metric tons $CO_2$ will be emitted by HPC data centers globally in 2024**.

## 3.4 Temperature Impact Model

The first environmental factor we consider in our model is temperature impact. Despite not being included in the prompt's additional environmental impacts of HPC, we strongly believe that assessing the thermal impact of HPC is important to evaluating its environmental impact because it is a core variable to many environmental issues. For example, the increase of global temperature, or global warming, can have lethal biological effects such as on coral reefs. Global warming also leads to an increase in precipitation and humidity, increasing natural disasters such as flash floods. Furthermore, the increase in global temperature can have drastic effects on sea levels and sea ice decline.

Since the relationships between carbon emissions and temperature are largely complicated, our model assumes a positive correlation between carbon emissions and temperature for the short-term. Since climate sensitivity ($\lambda$) is defined as the equilibrium global surface temperature change ($\Delta T$) in response to a specified unit forcing after the planet has come back to energy balance [13], we can utilize the following relationship from literature [13]:

$$\Delta T(t) = \lambda \cdot \Delta F(t) \tag{15}$$

Where $\Delta T$ is the change in temperature, $\lambda$ is climate sensitivity, $\Delta F$ is radiative forcing. Radiative forcing is a measure of how the energy balance of the Earth–atmosphere system is influenced [14]. To derive a relationship for radiative forcing, we use the simplified expression of radiative forcing for $CO_2$ derived by the Intergovernmental Panel on Climate Change (IPCC) [15].

$$\Delta F(t) = 5.35 \cdot \ln\left(\frac{CO_2(t) + C_0}{C_0}\right) \tag{16}$$

In this expression, $CO_2(t)$ is our HPC carbon emissions respective to the year, $C_0$ is preindustrial carbon emission levels. We use a magnitude of 4.8 as climate sensitivity suggested based on literature [16] and we let $C_0$ = 280ppm [17]. Since preindustrial carbon emissions is measured in ppm, we convert our metric tons $CO_2$ to ppm using its conversion factor. Preindustrial carbon emissions is added on numerator because we can add the preindustrial level (280 ppm) to each year's $CO_2$ concentration to assess the total $CO_2$ concentration each year relative to preindustrial levels. This allows us to obtain a dynamic radiative forcing with respect to the time in year, deriving temperature changes from global HPC capabilities.

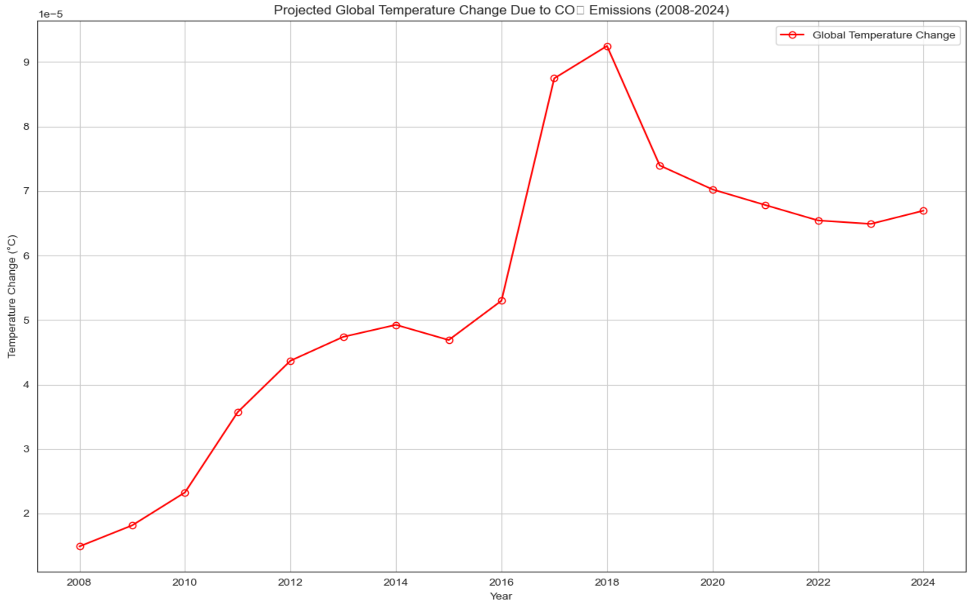


*Figure 15: TOP500 Impact on Global Temperature Change*

In the Figure 15, our model estimates that in 2024, global temperature rises by $6.696 \times 10^{-8}$ °C from TOP500 HPC capabilities, and a global temperature rise by $6.896 \times 10^{-6}$ °C from HPC data centers globally after scaling with the dataset ratio.

## 3.5 Air Quality Impact Model

Considering that fossil fuels—such as coal, oil, and natural gas—remain the dominant energy sources for power generation in most regions, including North America, Europe, and Asia, it is essential to examine the environmental pollution caused by emissions from their combustion. Fossil fuels contain not only carbon and hydrogen, the primary elements for generating energy, but also impurities like sulfur and nitrogen. When burned, these impurities lead to the formation of gaseous compounds such as $SO_x$ and $NO_x$.

These unwanted byproducts stem from the origin of fossil fuels. The resulting gaseous compounds are often toxic and can react with atmospheric water to form acids, contributing to acid rain. For example, in China, approximately 1.4 million tons of $SO_2$ are emitted annually, about 90% of which comes from coal-fired power generation [18]. This significant contribution is why sulfur dioxide is included as a key metric in the Air Quality Index (AQI).

For $SO_x$ emissions**,** to account for the variation in impurity proportions in fossil fuels across different regions, the concept of an emission factor can be introduced, forming the basis for a comprehensive $SO_x$ emission model. The emission factor is expressed in a composite unit, representing the mass of gas emitted per unit of energy produced. To align with the requirements of subsequent AQI calculations, the unit for $SO_x$ is specified as micrograms per terawatt-hour (μg/TWh). The formula for calculating the $SO_x$ emission factor is as follows:

$$EF_{SO_x} = (1 - r) \cdot \frac{S}{H} \tag{17}$$

Where $EF_{\mathrm{SO_x}}$ represents $SO_x$ emission factor, $S$ represents sulfur content of the fuel, $H$ represents thermal content of the fuel, and $r$ represents mass fraction of sulfur retained in the ash. The mass fraction of sulfur retained in ash varies significantly across fuel types. For anthracite, this value fluctuates around 5%, which for lignite, it can range from 34% to 45% [19]. In contrast, for liquid or gaseous fuels, the sulfur retention in ash is nearly zero. Below is a list of the sulfur content and thermal content of various energy sources collected from literature [20]:

| Energy source | Sulfur content (in % of fuel weight) | Thermal content (in MJ/kg) |
|---|---|---|
| **High-sulfur coal** | 3.000 | 28 |
| **Low-sulfur coal** | 1.000 | 28 |
| **Residual fuel oil** | 3.000 | 43 |
| **Distillate fuel oil** | 0.300 | 45 |
| **Natural gas** | 0.002 | 51 |

*Table 2: Sulfur Content and Thermal Content of Various Energy Sources*

In the power generation industry, thermal coal is the most used type of coal for electricity production. The decision to use Chinese standards as a basis for assumptions is driven by global coal consumption patterns. Since 2000, China's consumption of thermal coal has consistently accounted for approximately 50% of global coal usage during the same period [21]. Therefore, adopting Chinese standards provides a more accurate foundation than other standards for analysis. In the context of this paper, it is reasonable to assume that $r$ for coal used in thermal power is 40%, while the value for liquid or gaseous fuels is assumed to be 0%. Below is an example of the calculation process for some of the data presented in the table:

$$EF_{high} = \left(1 - r_{high}\right) \cdot \frac{S_{high}}{H_{high}} = (1 - 40\%) \cdot \frac{3\%}{28MJ \cdot kg^{-1}} \cdot 10^9 \mu g \cdot kg^{-1} \cdot 3.6 \times 10^9 MJ \cdot TWh^{-1}$$
$$\approx 2.3143 \times 10^{15} \mu g \cdot TWh^{-1}$$

$$EF_{dist} = (1 - r_{dist}) \cdot \frac{S_{dist}}{H_{dist}} = (1 - 0\%) \cdot \frac{0.3\%}{45MJ \cdot kg^{-1}} \cdot 10^9 \mu g \cdot kg^{-1} \cdot 3.6 \times 10^9 MJ \cdot TWh^{-1}$$
$$= 2.4000 \times 10^{14} \mu g \cdot TWh^{-1}$$

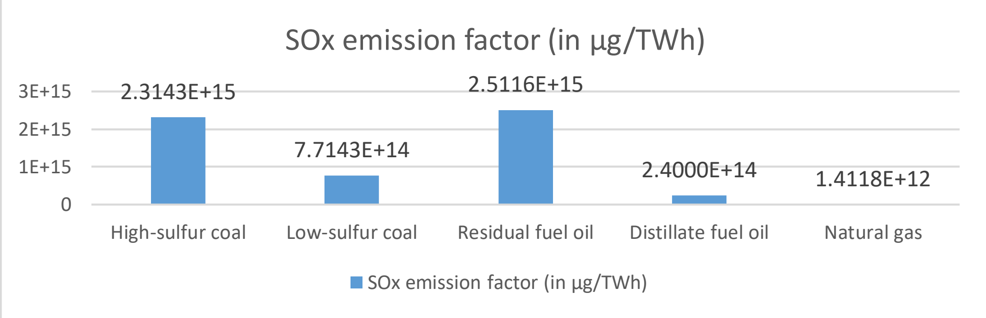


*Figure 16: $SO_x$ Emission Factors*

Thus, emission factors can be found as shown in Figure 16. After emission factors are found, the mass of total $SO_x$ emitted can be calculated by using the following equation (all sulfide masses here are calculated as sulfur dioxide equivalents):

$$m_{emission} = \sum EF_i \cdot E_i \tag{18}$$

Where $E$ represents HPC energy consumption (in TWh) and $i$ represents fuel type. To facilitate subsequent AQI calculations, total emissions are converted into emissions per unit volume, also known as concentration values. The formula is as follows:

$$C_p = \frac{m_{emission}}{S_{cont} \cdot h_{diff}} \tag{19}$$

Where $C_p$ represents concentration of the pollutant, $S_{cont}$ represents surface area of the continent under analysis, and $h_{diff}$ maximum dispersion height of the pollutant. For $NO_x$, despite differences in mechanisms from $SO_x$, $NO_x$ emissions can be calculated using the same method applied to $SO_x$ emissions, with the emission factor (EF) derived from databases. The formulas that are used are also (18) and (19).The following table summarizes the range of $NO_x$ emission factors for different duel types (in g/GJ) collected from literature [22]:

| **Type of fuel** | **Coal** | **Oil** | **Gas** |
|---|---|---|---|
| **Industry** | 150 – 540 | 85 - 180 | 50 - 98 |
| **Electric utilities** | 156 – 558 | 112 - 200 | 50 - 170 |

*Table 3: Range of $NO_x$ Emission Factors for Different Fuel Types*

The calculation of $NO_x$ requires careful attention to technology type, though, as existing technologies can significantly reduce emissions. For instance, the use of low-$NO_x$ burners can lower $NO_x$ emissions from thermal power plants by 30-50% [23]. However, due to limitations in tracking the specific technologies used by power plants, this factor is not included in the current model. Consequently, the calculated $NO_x$ emission estimates may have limited applicability for certain assessments.

Next, we apply our model to evaluate the air quality impacts that HPC contributes to in 2024 and we make a few assumptions first. Sulfur does not affect combustion itself but causes severe corrosion to equipment. Therefore, the sulfur content of fossil fuels used in power plants is generally below 2.5%. For calculations, the data for low-sulfur coal, distillate fuel oil, and natural gas in Table 2 are used as the fuel assumptions. It is assumed that the emission factors of the fuels used by power plants are uniformly distributed within the given range. Thus, the $NO_x$ emission factor is processed as the average of the upper and lower bounds of the data in the table. Considering that the primary environmental impact of $SO_x$ and $NO_x$ is acid rain, which occurs mainly in the troposphere where most clouds are located, and given that the troposphere contains approximately 75% of the atmosphere's total mass and the majority of its impurities, the maximum dispersion height of pollutants is set at the average height of the troposphere's upper boundary, 12 km [24][25].

Using energy mix ratios also used previously in our carbon emissions model, we can estimate the amount of energy consumed under different fossil fuels. For example, $E_{coal} = 29.477\% \times 362TWh \approx$ 106.705TWh, $E_{oil} \approx 3.823TWh, E_{gas} \approx 89.928TWh.$

Thus, mass of $SO_x$ and $NO_x$ would be:

$$m_{SO_x} = \sum EF_i \cdot E_i = EF_{coal} \cdot E_{coal} + EF_{oil} \cdot E_{oil} + EF_{gas} \cdot E_{gas}$$
$$= 7.7143 \times 10^{14} \mu g \cdot TWh^{-1} \cdot 106.705TWh + 2.4 \times 10^{14} \mu g \cdot TWh^{-1} \cdot 3.823TWh$$
$$+ 1.4118 \times 10^{12} \mu g \cdot TWh^{-1} \cdot 89.928TWh \approx 8.336 \times 10^{16} \mu g$$

the same,

$$m_{NO_x} \approx 1.749 \times 10^{11} \mu g$$

Subsequently, the environmental impact can be assessed by calculating the proportion of pollutant mass in total emissions or by calculating the IAQI. Here, the $C_P$ calculation of $SO_x$ as a part of IAQI is chosen:

$$C_{SO_x} = \frac{m_{SO_x}}{S_{cont} \cdot h_{diff}} = \frac{8.336 \times 10^{16} \mu g}{1.49 \times 10^{14} m^2 \cdot 12000m} = 4.662 \times 10^{-2} \mu g \cdot m^{-3}$$

The other parameters except $C_{SO_x}$ should be treated as functions of pollutant concentrations when doing IAQI calculations, denoted as $f : \mathbb{R} \to \mathbb{R}$. These mappings are derived from the AQI breakpoint table.

### 3.6 Strength and Limitations

**Strengths:** Our model first utilizes TOP500 data and our forecasts to derive energy consumption which is effective at considering energy mixes by geographical region for a more realistic and in-depth analysis. We then consider a dynamic energy mix of 8 different which is effective at capturing renewable energy growth, deriving a recursive model. Additionally, our multivariate $CO_2$ emissions model considers the carbon intensity of each source and is scaled by our dataset ratio to obtain an estimate of global data center carbon emissions. From here, our model can then evaluate two environmental impacts of HPC: temperature impact and air quality impact.

**Limitations:** Our model uses historical data, thus it may not guarantee high accuracy in the future as they are limited in predicting sudden events or turning points in the future. Since our model primarily focuses on carbon emissions, it does not extend beyond temperature and air quality impact and may not demonstrate the same degree of concern for other environmental factors. Moreover, our model is limited by both dataset's sizes, as a more extensive dataset may more accurately depict trends and forecasts that extend beyond our model's current consideration.

# 4 Model Application to Future Scenarios

### 4.1 Analyzing the Increase in Energy Demand

In recent years, renewable energy has become a cost-effective energy solution, with many renewable source becoming cheaper and cost-efficient compared to non-renewables. The increase in demand results in more deployment, in which prices fall, introducing increased competition in the market [26]. However, when considering the increasing demand for energy in other sectors such transportation or residential, the cost dynamics are complex when factoring in economies of scale and technological advancement. Currently, our model does not specifically address the increasing competition in other sectors, so in this section, we consider the scenario of increasing demand for energy in other sectors by assuming that an increase in demand for renewable energy will cause a short-term increase in equilibrium price.

From the IEA's 2024 energy mix data, it is clear that hydropower is a leading renewable energy source in all three continents. With the increase in demand for hydropower, the average installation cost for hydropower has shown to have an increasing tendency within the last decade [27]. Thus, we take hydropower as a case study and simulate the effects of increasing demand from other sectors to hydropower proportion in HPC, specifically in North America. To consider the growth in cost, we must make changes to our model. Utilizing our recursive energy mix growth model, we add another parameter $C$, which is the % of cost increase of an energy source relative to 2024.

$$R_{i,j}(t) = \frac{\psi_j(t)}{\psi_j(t-1)} R_{i,j}(t-1) \cdot \frac{1}{1+C} \tag{20}$$

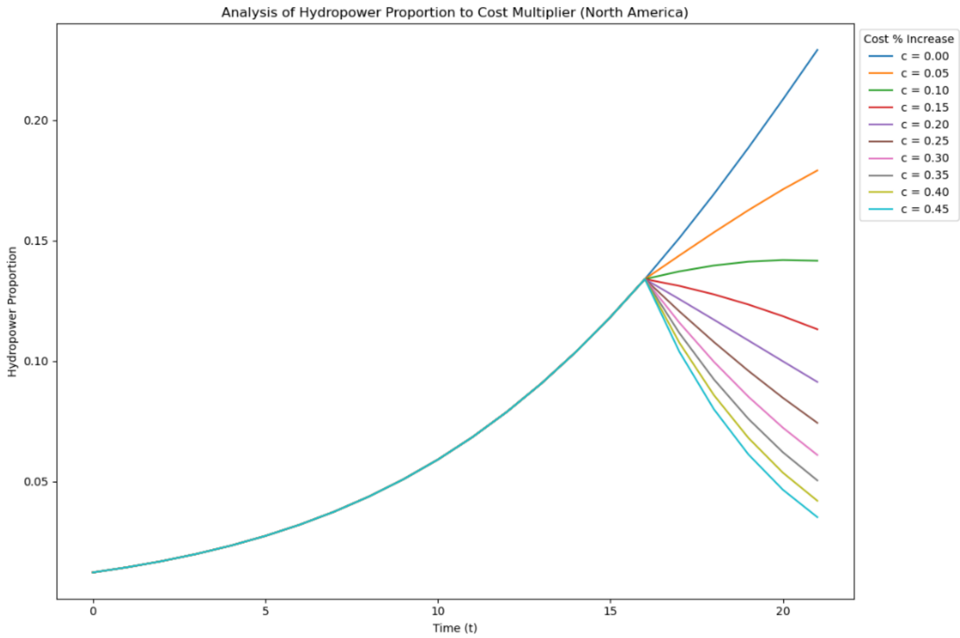


*Figure 17: Cost Multiplier Impact on Hydropower Proportion After 2024*

Analyzing hydropower proportion with different cost increase margins starting from 2024, we identify that initially hydropower proportion slows down in growth and levels off a bit earlier. Interestingly, after ~10% cost increase, we notice that hydropower proportion begins to fall, reaching a threshold where our model emphasizes priority over other renewable energy sources to take hydropower's place when our model deems a higher hydropower proportion to be unreasonable given its cost increase. From our analysis, we model increasing demand for renewable energy in other sectors besides HPC to correlate to cost increases in a short-term period which has shown us that the growth of energy sources will plateau and flip after a specific threshold.

## 4.2 Forecasting the Scope of the Issue in 2030

To provide realistic bounds that can provide insight into the scope of the problem in the year 2030, we will address the bounds by exploring the **worst-case scenario** and the **best-case scenario**. Mainly, we look at how the energy mixes of each continent will change. From our global dataset, it is estimated that in 2030 HPC data centers will consume 1593TWh. For the best-case scenario, we assume that renewable energy will continue its logistic growth within energy mix as discussed in our energy mix model because of factors such as government initiatives and advocation for cleaner energy. For the worst-case scenario, we make the underlying assumption that the renewable energy composition of each continent peaks in 2024 and does not continue its assumed growth. Additionally, our HPC energy consumption forecast model on TOP500 data will be implemented and then scaled by our dataset ratio to provide a forecast of all HPC data center energy consumption in 2030. This way we can provide realistic bounds for carbon emissions and its environmental impacts from HPC by addressing optimal and un-optimal conditions of renewable energy composition.

**Best-case scenario:** In our best-case scenario, we utilize our energy mix model to derive the following energy mix ratios for each energy source within each continent. Additionally, we use our energy consumption forecast model to predict that in 2030, Asia's, North America's, and Europe's TOP500 HPC capabilities will consume 1.26 TWh, 2.06 TWh, 1.09 TWh respectively. Values in Table 4 below are truncated to the thousandths.

| 2030 | Hydropower | Solar | Wind | Biofuel | Coal | Oil | Natural Gas | Nuclear |
|---|---|---|---|---|---|---|---|---|
| Asia | 0.250 | 0.096 | 0.115 | 0.037 | 0.388 | 0.007 | 0.069 | 0.034 |
| NA | 0.249 | 0.057 | 0.173 | 0.018 | 0.119 | 0.006 | 0.253 | 0.112 |
| Europe | 0.188 | 0.080 | 0.176 | 0.053 | 0.146 | 0.009 | 0.172 | 0.154 |

*Table 4: 2030 Best-Case Scenario Energy Mixes*

**Worst-case scenario:** Many sources suggest that renewable energy will indeed continue to grow. For example, the IEA anticipates that total renewable energy capacity is expected to double in the next five years [28]. To keep our lower bound more realistic, in our worst-case scenario we will assume that renewable energy growth peaks during 2024 and does not grow any further prior to 2030. We chose not to assume the scenario where renewable energy proportions begin to drop up to 2030, as this scenario is less likely to occur and is questionable in terms of plausibility. Thus, we assume that energy mix proportions stay constant relative to 2024. However, we will still consider the technological progression of HPC, utilizing our HPC energy consumption model to forecast in 2030.

| 2030 | Hydropower | Solar | Wind | Biofuel | Coal | Oil | Natural Gas | Nuclear |
|---|---|---|---|---|---|---|---|---|
| Asia | 0.133 | 0.051 | 0.061 | 0.020 | 0.571 | 0.011 | 0.102 | 0.051 |
| NA | 0.134 | 0.031 | 0.093 | 0.010 | 0.175 | 0.010 | 0.371 | 0.165 |
| Europe | 0.147 | 0.063 | 0.137 | 0.042 | 0.179 | 0.011 | 0.211 | 0.189 |

*Table 5: 2030 Worst-Case Scenario Energy Mixes*

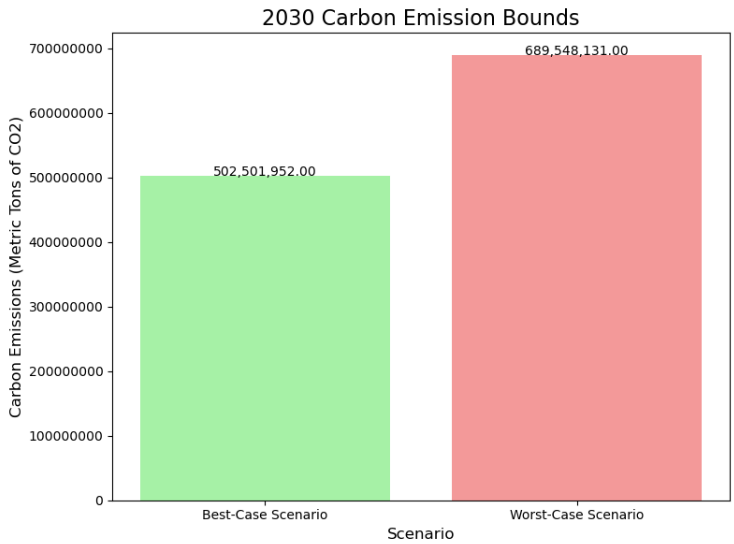


*Figure 18: 2030 Carbon Emissions Bounds*

To calculate the total carbon emissions in both scenarios, we utilize our carbon emissions model which takes in carbon intensity, energy mix ratios, energy consumption, and our dataset ratio. For the **best-case scenario**, our model (scaling with dataset ratio) estimates **that 502,501,952 metric tons of $CO_2$ will be emitted globally from HPC data centers around the world in 2030**. For the **worst-case scenario**, our model estimates that **689,548,131 metric tons of $CO_2$ will be emitted from HPC data centers globally in 2030.**

# 5 Case Analysis: Increasing Renewable Energy Proportion

In our energy mix model, we discussed the implications of an increasing renewable energy composition. In addition, in this section we seek to explore the reductions in carbon emissions in a year when renewable energy composition varies as an independent variable, while keeping energy demand/consumption constant. To do so, we will take a look at 2024 specifically. In 2024, our data indicates that for TOP500 North America, Asia, and Europe will consume 1.62TWh, 0.93TWh, 0.96TWh of energy. For 2024, we use our carbon emissions model to explore the respective carbon emissions for 25%, 50%, 75%, and 100% renewable energy composition. However, our model cannot output a 100% renewable energy composition due to its logistic nature, which is a realistic characteristic that will be discussed later, so we use 99% instead. Additionally, we keep the same energy source ratios in 2024 for each continent initially. We find the respective time inputs where our model outputs our desired renewable energy composition and collect the respective energy source ratios. With our energy source ratios, we can then utilize our carbon emissions model which also converts TOP500 emissions to global data center emissions via dataset ratio.

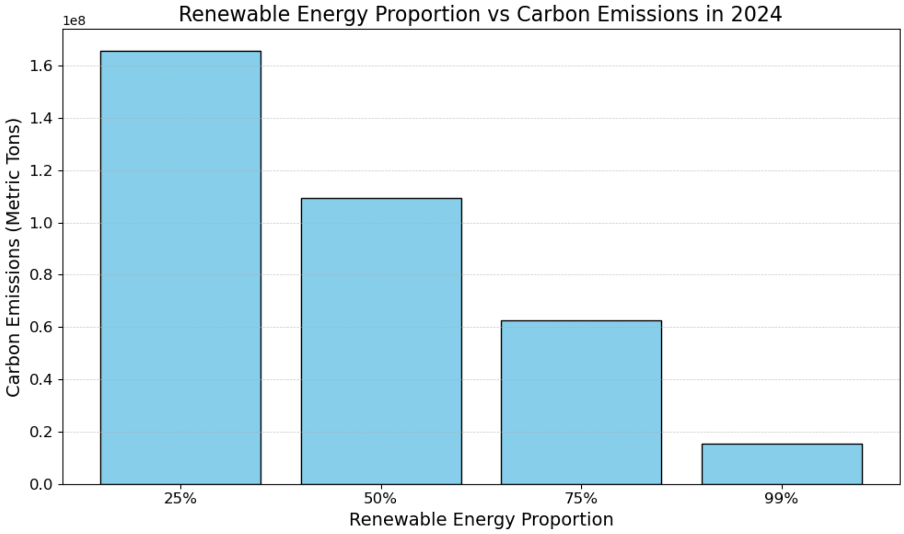


*Figure 19: 2024 Renewable Energy Proportion (25%, 50%, 75%, 99%) vs Carbon Emissions*

With 25% proportion, there is an initial 165,690,907 metric tons of $CO_2$ emitted. Referring to our results, the 25% to 50% jump demonstrates a reduction of 56,342,665 metric tons of $CO_2$, the 50% to 75% jump demonstrates a reduction of 46,772,688 metric tons of $CO_2$, and the 75% to 99% jump demonstrates a reduction of 47,173,826 metric tons of $CO_2$. This reflects the general trend of a reduction of carbon emissions as renewable energy proportion increases, however the change in reductions as % increases begin level off. In terms of seeking to switch to a 100% renewable energy source, it is critical to understand that renewable energy growth faces challenges in the long-term as it seeks to peak towards perfection. Financially, renewable energy is generally more expensive than its counterpart, posing as a challenge for many developing countries in the world. Switching to a 100% renewable energy composition would not only affect the HPC industry, but also external industries. For example, gas-powered cars are still dominant in the automotive industry, and to globally or regionally adapt to a 100% renewable energy mix would be highly challenging, as that would mean in this example that all gas-powered cars would cease to be in use. Additionally, renewable energy sources such as wind and solar are intermittent and depend on weather conditions and time of day which could result in reliability issues. Of course, a 100% renewable energy composition would be the most optimal to reducing carbon emissions and other environmental impacts, but the feasibility of this outcome is currently unrealistic which our energy mix model takes into account of.

# 6 Modeling HPC Land Usage

## 6.1 Modeling HPC Land Usage

In this section, we develop a model to consider land use as an environmental impact from HPC data centers. Our work from earlier focuses on temperature change and air quality, so additionally we construct a model for data center land use to further understand the impact of HPC on the environment. Since our datasets do not address land use, we utilize a new dataset which contains the power usage (MW) and land usage (sq ft) for data centers in the United States [29]. We specifically choose these two variables to develop a relationship as more land usage generally indicates larger amounts of hardware and computing, which means more power and energy consumption. Other than its connection to power and energy consumption, land usage has a critical impact on the wilderness and biodiversity, as allocating massive areas of land for data centers strips the natural habitats in many ecosystems.

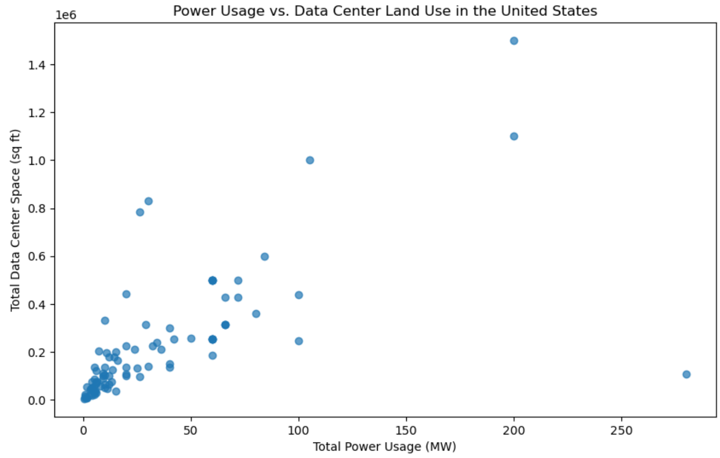


*Figure 20: U.S. Datacenter Dataset Power vs. Land Usage*

However, for some data centers, despite having equivalent land usage, some may have more advanced technology than others, resulting in larger power usage. Thus, to consider this relationship, for power usage data with differing land usage, we average the land use values to obtain a general estimate. To develop a relationship, we utilize regression techniques to find the best curve-fit for the dataset. First, since we notice that our data is a bit noisy, we eliminate possible outliers from the dataset by analyzing z-score and eliminating points with a z-score magnitude of 2 or greater and we scale MW to kW for consistency. Next, we apply three regression techniques, namely: polynomial, linear and logarithmic regression.

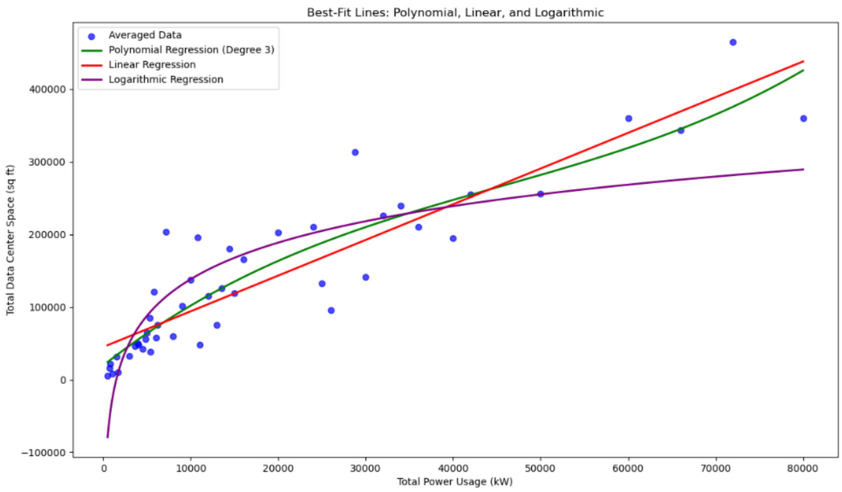


*Figure 21: Polynomial, Linear and Logarithmic Regression*

To evaluate the performance of our regression techniques, we compare the $R^2$ values. Specifically, polynomial regression (degree 3) obtained an $R^2$ value of 0.84, linear regression obtained an $R^2$ value of 0.82, and logarithmic regression obtained an $R^2$ value of 0.72. Based on these metrics, we choose to use polynomial regression as our primary fit. With a general relationship between a data center's power usage and land usage, we develop the following model for all data centers:

$$L(P,t) = (19587.466 + 9.493P - 0.0001345P^2 + 0.0000000009916P^3) * n_{centers}(\mathrm{t}) \tag{21}$$

Where $L(P,t)$ represents the total land use in square feet from data centers around the world with respect to $P$, the average power usage of a data center and $t$, the time variable or the respective year. Additionally, we consider $n_{centers}(\mathrm{t})$, the number of data centers around the world during year $t$ to calculate the total land usage. Simply put, we model the total land usage of HPC data centers by finding the average land use per data center with our polynomial regression given the average power of a data center and scale this quantity by the total number of data centers in the world.

To apply our model and forecast HPC land usage, we use our model to find the approximate HPC land usage in 2024 and in 2030. In 2024 it is estimated that there are 5,709 data centers around the world. In 2030 it is forecasted that there will be approximately 8,378 data centers around the world [30]. Now, to find the average power usage, we take energy consumption quantities from our HPC global dataset and divide by the number data centers in that year and the total number of hours. Using this approach, in 2024, we calculate the average power per data center to be approximately 362TWh / (5,709 * 8,760h) = 7,238.4 kW. In 2030, we calculate the average power per data center to be approximately 1,593TWh / (8,378 * 8,760h) = 21,705.6 kW. Integrating these quantities into our model with (21) we obtain:

$$L(7{,}238.4, 2024) \approx 466{,}029{,}163 \text{ square feet of land usage}$$

$$L(21{,}705.6, 2030) \approx 1{,}444{,}464{,}841 \text{ square feet of land usage}$$

Thus, our model forecasts that by 2030, there will be an approximate **210% increase in HPC data center land usage** relative to 2024.

### 6.2 Strengths and Limitations

**Strengths:** Our HPC land usage model utilizes regression to develop a general relationship between power usage and land usage. Since power usage is relatively easier to obtain than many other metrics regarding data centers, our model is widely and more easily applicable.

**Limitations:** Since our model is based on historical data, it may not accurately forecast sudden turns within the future and can only base its forecasts and relationship from previous data. Note that our model is weak when considering the possibility of massive technological progression, where higher usages of HPC would not use nearly as much land as our model predicts. Due to the size and nature of the dataset used, a larger dataset size and data inclusive of other regions besides the U.S. would improve the accuracy and realism of our model.

## 7 Mitigating the Environmental Impacts of HPC

### 7.1 Recommendations

In recent years, global warming and other environmental problems have become one of the most urgent issues concerning the whole of mankind. Now, with the rapid rise of high-powered computing, it is necessary to take action to mitigate the environmental impacts of HPC.

Thus, we suggest governments adopt measures that:

1. Provide subsidies for HPC data centers to enforce a renewable energy composition from electricity usage equal to or higher than a certain proportion, such as 50%.
2. Increase government spending on innovative technology development related to decreasing greenhouse gases emitted by fossil fuels to improve energy efficiency.

3. Levy taxes on energy sources with high carbon emission factors, especially coal and oil, to decrease the amount of fossil fuels used and slow down resource depletion.
4. Inform the public more about the damage of HPC data centers' carbon emissions to the environment and people's health to raise awareness of the public and encourage spontaneous

We suggest data centers to:

1. Use hardware with higher sustainability, to reduce the total number of racks required by the data center and therefore decrease the great amount of carbon dioxide generated by manufacturing the hardware racks.
2. Choose hardware of higher energy efficiency. For example, use solid-state drives (SSDs) instead of traditional hard disk drives (HDDs), to maximize the output of certain energy consumption and reduce the amount of energy sources needed to be consumed.
3. Choose power-saving server models that can minimize the amount of energy use when the hardware is idle to decrease futile energy consumption and reduce energy sources consumption, as hardware in data centers must be on all the time.

## 7.2 Demonstrating Recommendation with Our Model

If our recommendation for enforcing a renewable energy composition of at least 50% is considered and action is taken upon by governments around the world, our models can be used to calculate and demonstrate the effectiveness of this recommendation by inputting the respective energy mix ratios.

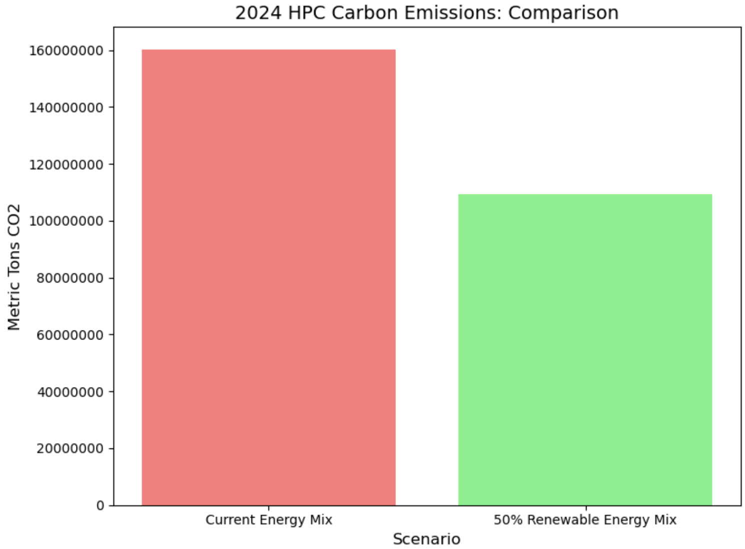


*Figure 22: 2024 $CO_2$ Emissions of Current Energy Mix vs 50% Renewable Energy Mix*

Take 2024 for example, if we keep the current energy mix, then our model estimates that approximately 160,147,474 metric tons of $CO_2$ will be emitted from HPC data centers worldwide. However, if our recommendation is adopted and enforced, our model estimates that around 109,348,242 metric tons of $CO_2$ will be emitted instead. This is a near 32% reduction in $CO_2$ emissions from HPC capabilities.

# 8 Conclusion

The use of mathematical modeling to assess the environmental impacts of high-powered computing is not only effective, but also necessary in today's world. The models constructed within our work not only provide a concrete estimation of quantities but are also capable of forecasting future values. Our multivariate carbon emissions model estimates that in 2024, HPC capabilities worldwide emit around 160,147,474 metric tons $CO_2$, and 502,501,952 metric tons of $CO_2$. Our model considers the environmental impacts of temperature change (estimating a 6.896*10-6 °C increase in 2024 from HPC), air quality ($4.662 \times 10^{-2} \mu g \cdot m^{-3}$ in 2024), and land usage later (a near 210% increase within the next 6 years).

# 9 A Proposal to the United Nations Advisory Board

Dear United Nations Advisory Board,

First and foremost, we would like to express our gratitude for the work being done towards a cleaner and more sustainable future. However, after reading the "Governing AI for Humanity" report issued in September 2024, we realized that a crucial and very far-reaching aspect concerning AI was not fully addressed: high-powered computing. With due respect, we are writing to call for increased attention on the environmental impacts of HPC data centers and capabilities, and a more detailed section on this important topic in your scheduled developmental goals for 2030.

With the growth of various fields such as data science and cryptocurrency mining, high-powered computing data centers are developing rapidly both in number and in size. So far, in 2024, there are already about 5,709 data centers all around the world, and in 2030 estimates indicate there will be around 8,378. The development of technology brings a major environmental problem and arouses the concerns of many people. According to the report issued by International Energy Agency in 2023, the energy consumption of data centers comprised approximately 1-1.5% of the world's total electricity use. Although it may seem that HPC does not comprise a large amount, the rapid growth of high-powered computing is inevitable and will flourish until it is too late if we do not take action. From the data used in our work, it is estimated that HPC will use 362 TWh in 2024 and 1593 TWh of energy in 2030, a **340% increase in just 6 years.**

Using mathematical models, we estimate that there will be 160,147,474 metric tons of carbon emissions from HPC in 2024. Maybe this number isn't significant yet. Maybe this number doesn't contribute to the global carbon emissions per year that much. But what we would like to emphasis is that **this is not about comparison**. This is about the ever-growing carbon emissions in the atmosphere. This is about the future of this planet, whether we can pass along a green world to future generations. **Every metric ton counts.** With each ton of $CO_2$ emitted into the atmosphere, our planet will continue to rise in temperature. In fact, our models estimate that in 2024, HPC energy consumption correlates to an approximate 6.896*10-6 °C rise in temperature. Our point is, that **environmental impact is accumulative**. This isn't some calculation where we compare ratios to value what's more urgent. This is a situation where every bit counts. We recognize the outstanding actions currently being taken to save our planet such as the Paris Agreement, and a net-zero coalition by 2050. However, carbon neutrality by 2050, for example, won't be met if we do not consider all the sectors contributing to carbon emissions.

We fully understand the current prioritization on larger sectors. But these **figures won't stop growing**. The potential that high-powered computing has within the next few decades must not be underestimated. By 2030, we estimate that HPC capabilities will emit 502,501,952 metric tons of $CO_2$ per year, **a 214% increase in just 6 years.** The future holds unpredictable. With merely a few steps forward, HPC carbon emissions already double themselves. At such a rate, if the environmental impacts of HPC are not addressed within the next couple decades**, such numbers could be lethal**.

However, it is not too late to act, but if we don't soon it will be too late. In fact, one recommendation we proposed from our work is for governments to provide subsidies to enforce a renewable energy composition of at least 50%. Our work estimates that by doing so, there will be a near 32% reduction in CO2 emissions from HPC capabilities in 2024, just from one policy. Nowadays, many data centers are gradually shifting their energy reliance from fossil fuels to renewable energies, but the current average percentage of fossil fuels, predominantly coal and natural gas, remain very high, with North America 73.2%, Asia 73.5%, and Europe 61.1%.

Although we cannot reverse the trend of intensified demand for AI and the inevitable growth of high-powered computing, in order to mitigate the detrimental ramifications caused by HPC data centers, we sincerely suggest that the United Nations Advisory Board include a more detailed section on the environmental impacts of HPC data centers in your scheduled developmental goals for 2030 to raise people's awareness and urge governments all around the world to take more measures. It will make a huge difference to us, to you, to mankind, to animals, and to our heartful home: Earth.

Yours,
Team 15248

# 10 References


[1] Gartner predicts power shortages will restrict 40% of AI data centers by 2027. (n.d.). https://www.gartner.com/en/newsroom/press-releases/2024-11-12-gartner-predicts-power-shortages-will-restrict-40-percent-of-ai-data-centers-by-20270

[2] David Mytton . (2022, August 17). Sources of Data Center Energy Estimates: A comprehensive review. Joule. https://www.sciencedirect.com/science/article/pii/S2542435122003580#

[3] Benhari, A. (n.d.). Green HPC: An analysis of the domain based on TOP500. https://arxiv.org/pdf/2403.17466

[4] BYJU'S. (2022, January 3). Linear interpolation formula with solved examples. BYJUS. https://byjus.com/linear-interpolation-formula/

[5] Forecasting: Principles and practice (2nd ed). 7.2 Trend methods. (n.d.). https://otexts.com/fpp2/holt.html

[6] IEA. (n.d.). Countries & Regions. IEA. https://www.iea.org/countries

[7] U.S. Energy Information Administration - EIA - independent statistics and analysis. Renewable energy explained - renewable portfolio and clean energy standards - U.S. Energy Information Administration (EIA). (n.d.). https://www.eia.gov/energyexplained/renewable-sources/portfolio-standards.php

[8] Priorities 2019-2024. European Commission. (n.d.). https://commission.europa.eu/strategy-and-policy/priorities-2019-2024_en

[9] Hansen , J. P. (2016, November 30). Limits to growth in the Renewable Energy Sector. Renewable and Sustainable Energy Reviews. https://www.sciencedirect.com/science/article/abs/pii/S1364032116310371

[10] Iea. (n.d.-b). Massive global growth of renewables to 2030 is set to match entire power capacity of major economies today, moving world closer to tripling goal - news. IEA. https://www.iea.org/news/massive-global-growth-of-renewables-to-2030-is-set-to-match-entire-power-capacity-of-major-economies-today-moving-world-closer-to-tripling-goal

[11] Frequently asked questions (faqs) - U.S. energy information administration (EIA). Frequently Asked Questions (FAQs) - U.S. Energy Information Administration (EIA). (n.d.-a). https://www.eia.gov/tools/faqs/faq.php?id=74&t=11

[12] Carbon dioxide emissions from electricity. World Nuclear Association. (n.d.). https://world-nuclear.org/information-library/energy-and-the-environment/carbon-dioxide-emissions-from-electricity

[13] Hansen, J., Sato, M., Kharecha, P., & von Schuckmann, K. (2011, December 22). Earth's energy imbalance and implications. Atmospheric Chemistry and Physics. https://acp.copernicus.org/articles/11/13421/2011/acp-11-13421-2011.html

[14] Radiative forcing. Radiative Forcing - an overview | ScienceDirect Topics. (n.d.). https://www.sciencedirect.com/topics/earth-and-planetary-sciences/radiative-forcing#:~:text=Radiative%20forcing%20is%20a%20measure,radiation%20within%20the%20Earth's%20atmosphere.

[15] New estimates of radiative forcing due to well mixed greenhouse gases - myhre - 1998 - geophysical research letters - wiley online library. (n.d.-b). https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/98GL01908

[16] Hansen, J. E. (2024, January 12). Annual global temperature 2023. Columbia University. Retrieved from https://www.columbia.edu/~jeh1/mailings/2024/AnnualT2023.2024.01.12.pdf

[17] Lindsey, R. (2024, April 9). Climate change: Atmospheric carbon dioxide. NOAA Climate.gov. https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide#:~:text=Before%20the%20Industrial%20Revolution%20started,was%20280%20ppm%20or%20less.

[18] 2024. (n.d.). http://theory.people.com.cn/n/2012/1218/c40531-19935445.html

[19] COMMISSION OF THE EUROPEAN COMMUNITIES (CEC) (1992), CORINAIR Inventory, Default Emissions Factor Handbook, 2nd edition, CEC-DG XI.

[20] IEA (1992a), Electricity supply in the OECD, OECD.

[21] IEA (2024), Global coal consumption, 2000-2026, IEA, Paris https://www.iea.org/data-and-statistics/charts/global-coal-consumption-2000-2026, Licence: CC BY 4.0

[22] IEA (1991), Greenhouse Gas Emissions: The Energy Dimension, OECD.

[23] IEA (1988), Emission Controls in Electricity Generation and Industry, OECD.

[24] Wu Guanghe (ed.). Physical geography. Beijing: Higher Education. 2008: 88-89. ISBN 978-7-04-022876-2.

[25] Zhu Xiang; Liu Xinmin (eds.). General high school textbooks: Geography (compulsory course 1). Changsha: Hunan Education Press. 2019. ISBN 978-7-5539-4784-6.

[26] Roser, M. (2024, March 18). Why did renewables become so cheap so fast?. Our World in Data. https://ourworldindata.org/cheap-renewables-growth

[27] Statista Research Department, & 11, N. (2024, November 11). Hydropower installation cost worldwide 2022. Statista. https://www.statista.com/statistics/799341/global-hydropower-installation-cost/

[28] Change Oracle. (2024, January 14). Renewable energy is growing rapidly but is it enough to stop climate change? https://changeoracle.com/2023/01/31/renewables-are-growing-but-is-it-enough-to-stop-climate-change/

[29] US data center locations and Power Demand. Aterio. (n.d.). https://www.aterio.io/datasets/lst_us_data_centers

[30] How many data centers are there and where are they being built?. ABI Research: The Tech Intelligence Experts. (n.d.). https://www.abiresearch.com/blogs/2024/07/16/data-centers-by-region-size-company/#:~:text=By%20the%20end%20of%202024,be%20in%20operation%20by%202030.