Estimating the Energy Consumption of Data Centres in the European Union

Policy Implications for AI Deployment and Sustainability

Monserrat López Pérez

Hertie School

A dissertation submitted to the Hertie School for the degree of Master of Data Science for Public Policy.

Supervisor: Lynn Kaack

Hertie School

April 2025

Word count: 7,082

Contents

Executive Summary **1**

1. Introduction **2**

2. Literature Review **4**

2.1. Energy Requirements and Emissions of LLMs 4

2.2. Energy Consumption and Environmental Impacts of Data Centres 5

2.2.1. Estimations in the European Union 7

3. Data and Methods **9**

3.1. Data collection, preprocessing, and cleaning 9

3.1.1. Data Scrapping 9

3.1.2. Data Cleaning and Feature Extraction 10

3.1.3. Geocoding and Spatial Validation 11

3.1.4. Data Centre Classification 12

3.1.5. Imputation of Missing Values 13

3.1.6. Carbon Intensity 13

3.2. Data Centre Energy Consumption and Emissions Estimation . . 14

3.2.1. Energy Consumption Estimation 14

3.2.2. CO_2 Emissions Estimation 15

4. Results **17**

4.0.1. Spatial Distribution and Clustering 17

4.1. Energy Consumption 18

4.2. Carbon Emissions 20

5. Discussion **23**
5.1. Limitations 26
5.2. Policy Implications for AI Deployment and Sustainability in the Region 28

6. Conclusion **30**

A. Appendix **31**

Bibliography **37**

Executive Summary

This thesis estimates data centre electricity consumption and related carbon emissions for the 27 European Union Member States. It employs a bottom-up approach and an area-based methodology, using benchmarks from the literature on facility size, power usage effectiveness (PUE), and IT power densities. Three scenarios (low, mid, and high) are developed, assuming different levels of energy efficiency, leading to different energy consumption levels.

For doing so, data was scraped from a public facility listing, retrieving information about each data centre located within the European Union. The results reveal that data centre energy consumption across the Member States is estimated from 54.5 to 60 TWh annually, representing approximately 2.2% to 2.6% of the EU's total electricity consumption. At the same time, the carbon emissions related to this consumption are estimated between 14.16 and 17.16 Mt CO_2e/year.

The analysis reveals a high concentration of data centres clusters in four main cities: Frankfurt, Paris, Dublin, and Amsterdam. Although the electricity consumption of data centres in these countries is high, there is a disparity in emissions related to the differences in the energy mix among European countries.

These findings can help to inform policy discussions on the implications of the projected increase in energy demand related to the deployment of AI technologies. They also highlight the need to create policies for optimizing the electricity mix of countries with high carbon intensities and ensuring robust digital infrastructure to meet the ambitious suitability goals of the region.

1. Introduction

In recent years, demand for digital services has increased at an unprecedented rate, particularly driven by the deployment and development of Artificial Intelligence (AI) technologies. AI and machine learning (ML) models require substantial processing power and, consequently, significant amounts of energy, varying from the type of algorithm used, size of the models, and stage in the ML life cycle (inference, training, and deployment) (Kaack et al., 2022). Since 2010, the computing power needed to train AI models has grown at a rate of 4.6× per year (Epoch AI, 2024). This exponential growth in AI adoption directly affects the energy consumption of data centres.

Data centres are the physical infrastructure that supports the Information and Communication Technology (ICT) sector and the digital economy. These facilities house computing infrastructure for data processing, storage, and digital communications (Shehabi et al., 2016), and contain power, backup systems, and cooling infrastructure to maintain optimal conditions for the equipment stored in them.

Globally, data centres account for around 1.5% of final electricity demand, with the United States (US) being the largest consumer in 2024 (45%), followed by China (25%) and Europe (15%). This figure is expected to double by 2030, driven by AI adoption and Deployment (International Energy Agency, 2025).

The rising energy consumption of data centres raises environmental concerns, as their energy demands are supplied by the local electricity grid, which in many regions still relies heavily on fossil fuels. Their carbon footprint varies according to their geographic location and the carbon intensity of the regional energy mix.

The environmental impacts of data centres are particularly relevant in Europe, where regulatory frameworks strongly emphasise sustainability goals in the region. The European Green Deal aims to reduce greenhouse gas emissions by 55% by 2030, relative to 1990 levels, and make Europe the first climate-neutral continent by 2050 (European Commission, 2019). At the same time, the European Union has announced plans to invest €200 billion in artificial intelligence, intending to establish Europe as an "AI continent" (European Commission, 2025).

This thesis uses a bottom-up model to estimate the energy consumption and carbon emissions of data centres across the European Union using geospatial and electricity grid information. The central research question is: How can geospatial data and electricity grid information be used to estimate the energy consumption and CO_2 emissions of data centres in the EU, and what are the implications for AI deployment and compliance with EU sustainability goals?

The scope of this research is limited to commercial data centres in the 27 EU Member States. A data science pipeline was developed to extract information from a public data centre listing, enriching these records with operational benchmarks and location carbon intensity values linked to electricity data. The findings aim to provide a geographical understanding of the environmental footprint of data centres in the EU and inform policy discussions on the sustainable development of digital infrastructure in the region.

2. Literature Review

There is growing research interest in estimating the energy demand associated with the development of AI systems and their environmental impacts. These assessments have been done using two primary approaches: one focusing on estimating the emissions generated during the training and inference phases of Deep Learning (DL) models and the other on estimating data centres' energy consumption and related emissions, considering infrastructure characteristics and other relevant indicators.

2.1. Energy Requirements and Emissions of LLMs

One significant research stream focuses on quantifying the carbon emissions generated by deep learning algorithms, particularly Large Language Models (LLMs). Pioneer work by Strubell et al. (2019) estimated the energy required to train and develop four different Natural Language Processing (NLP) tasks. According to their estimations, training a large Transformer model could emit 626,155 lbs of CO_2 (284 tonnes), equivalent to about five times the average lifetime emissions of a single car. While this study highlighted the issue's magnitude, it predates the release of publicly available LLMs such as ChatGPT and the subsequent growth of the AI industry.

Luccioni et al. (2022) estimated the carbon footprint of the BLOOM model (176 billion parameters) across its life cycle. The final training emitted approximately 24.7 tonnes of CO_2-equivalent (CO_2eq) considering only dynamic power consumption and 50.5 tonnes when including all processes from equipment manufacturing to operational energy use. They also estimated the carbon emissions related to inference deployment. Using the CodeCarbon tool on a Google Cloud Platform (GCP) instance, they tracked energy use over 18 days, estimating approximately 19 kg CO_2eq per day. While results vary by hardware and location, the example provides a valuable reference for the environmental impact of ML deployment.

Additionally, several tools have emerged to track energy use and carbon emissions from AI workloads (Heguerte et al., 2023), such as CodeCarbon (Courty et al., 2024), Carbon Tracker (Anthony et al., 2020), Green Algorithms (Lannelongue et al., 2021), or eco2AI (Budennyy et al., 2022). These tools contribute to the broader field of Green AI, which emphasizes awareness of AI's computational and environmental costs and promotes the measurement, monitoring, and transparency of computational efficiency (Schwartz et al., 2020).

While this first approach provides task-level benchmarks of AI energy use, quantifying emissions per model or training run, it does not account for the broader infrastructure required to support AI systems, particularly the data centres where training and inference occur.

2.2. Energy Consumption and Environmental Impacts of Data Centres

A second stream of research focuses on estimating data centres' energy consumption and their environmental footprint. Different methodological approaches have been used to estimate these impacts, including top-down, bottom-up, extrapolation, and hybrid methods. Bottom-up approaches estimate data centre energy use based on detailed information on data centre types locations, IT equipment, and energy efficiency trends (Masanet et al., 2020). Top-down approaches rely on aggregated data from national governments or industry reports. Extrapolation methods forecast future energy consumption by projecting total energy use from bottom-up or top-down estimates under different scenarios (Shehabi et al., 2024). Finally, hybrid approaches combine multiple sources and estimation methods (Kamiya & Bertoldi, 2024).

A widely cited bottom-up study from Shehabi et al. (2016) used historical data of equipment stock, data centre classifications, and trends in efficiency to estimate data centre electricity consumption in the United States (US). The authors found that US data centres consumed approximately 70 billion kWh (70 TWh) in 2014, accounting for around 1.8% of national electricity use. More recently, an updated study version incorporated more granular data, including shipment records for AI-specific hardware (GPUs) and cooling systems (Shehabi et al., 2024). Based on these updates, total energy consumption by US data centres was estimated at 176 TWh in 2023. The significant difference between the 2016 and 2024 estimations is related mainly to the AI-driven computing demands and the industry's shift toward power-intensive accelerated computing. While those reports are detailed bottom-up estimations, the data sources usually come from private industry reports, limiting their reproducibility.

In an attempt to use public data sources, Siddik et al. (2021) used data from prior studies for small and mid-size data centres, while colocation and hyperscale facilities were mapped using commercial listings such as Data Center Map and Baxtel. With that information, they analysed the environmental impacts of data centres in the US, considering both carbon emissions and water consumption using an area-based approach, applying IT load benchmarks from Shehabi et al. (2011) and type-specific PUE assumptions. Based on this methodology, the authors estimated that in 2018, total GHG emissions attributable to US data centres reached approximately 31.5 million tons of greenhouse gases in 2018, representing nearly 0.5% of the country's total greenhouse gas emissions.

More recently, Guidi et al. (2024) collected and analysed data from 2,132 US data centres, gathering information by combining web scraped data with proprietary information from Baxtel.com. The authors estimated electricity consumption using a formula based on maximum power capacity, annual hours (8,760), and uptime operational at full capacity. Based on this methodology, the authors estimated that US data centres consumed 192.64 TWh in 2022, which represented around 4.59% of the total energy consumption in the country, with associated CO_2 emissions of 105.59 million metric tons, or 1.66% of total US greenhouse gas emissions in 2022.

While the bottom-up studies are more accurate in estimating data centres energy demand and environmental impacts, this approach has its limitations (Shehabi et al., 2016). Accurately assessing energy consumption and related emissions is challenging due to the limited availability of public data on energy usage and key indicators such as Power Usage Effectiveness (PUE) and Water Usage Effectiveness (WUE), installed equipment base, capacity utilization rates, and other critical metrics. This absence of information leads to the necessity of scrapping information, buying private data, developing models for estimating some metrics or using industry benchmarks.

2.2.1. Estimations in the European Union

In the European context, few studies provide bottom-up insights into the energy consumption, water usage, and emissions of data centres. At the national level, Jerléus et al. (2024) collected data from diverse sources, including company websites, permit applications, questionnaire responses, and web-based repositories, such as Data Center Map and Baxtel, of 148 Swedish data centres. The authors used reported values for installed power or IT load and Power Usage Effectiveness (PUE) where available. In cases where these were missing, they applied an area-based estimation approach using average IT power density and PUE values stratified by facility size, based on benchmarks from Shehabi et al. (2011) and Shehabi et al. (2016). This area-based method also forms the foundation of the present study. According to their findings, Swedish data centres emitted approximately 130,000 tonnes of CO_2eq in 2020, equivalent to around 0.13 Mt or 0.3% of Sweden's total GHG emissions.

At the EU level, a recent study by the Joint Research Centre used hybrid methods and official sources for estimating data centres' energy consumption in the EU county members (Kamiya & Bertoldi, 2024). According to their estimations, data centres in the EU-27 consumed between 45 and 65 TWh of electricity in 2022, equivalent to 1.8–2.6% of the total electricity demand in the region. Most of this consumption was concentrated in Germany, France, the Netherlands, and Ireland.

While the discussed studies provide valuable references to data centre energy use, many rely on proprietary data sources or national-level aggregations that limit transparency and spatial analysis. This thesis builds on the area-based methodology proposed by Jerléus et al. (2024), using scraped data as a primary source and adapting it to the EU-27 countries. The following section outlines the data collection pipeline, preprocessing steps, feature extraction techniques, and energy modelling framework developed for this analysis.

3. Data and Methods

This research uses a bottom-up approach to estimate data centres' energy consumption and carbon emissions across the European Union (EU). The primary data used for energy and emissions modelling was scraped from Data Center Map, an industry listing that provides location, infrastructure, and textual metadata for commercial facilities (Data Center Map, 2025). All data were collected exclusively for academic and non-commercial purposes as part of this master's thesis project. Before the publication of the GitHub repository[1], all data were anonymised to protect any sensitive information. Only aggregated or non-identifiable fields are included, and no direct links to operators or individual facilities are provided.

The overall pipeline of this research consists of five main components: (1) data collection, preprocessing, and cleaning; (2) geocoding and spatial validation; (3) data centre classification; (4) imputation of missing data; and (5) energy and emissions estimation.

3.1. Data collection, preprocessing, and cleaning

3.1.1. Data Scrapping

First, data was extracted from `datacentermap.com`. XML files were downloaded from the sitemap and parsed to extract all data centre page URLs.

[1]https://github.com/monlo/thesis-dc-emissions

These links were then filtered to retain only those corresponding to data centres located in the EU-27 countries. Each facility page was visited using a Selenium browser configured with rotating user agents, retrieving metadata fields including address, service description, specifications (build-up power capacity and whitespace), and website links.

The final dataset resulting from this collection stage consisted of a structured CSV file containing metadata for 1,795 EU-based commercial data centres. Subsequent cleaning steps removed duplicate entries, excluded facilities marked as archived, planned, or under construction, and standardized country and city fields. After geocoding and spatial validation, the cleaned dataset included 1,600 operational facilities with validated location data across the EU-27, used for further classification, imputation, and energy modelling.

3.1.2. Data Cleaning and Feature Extraction

The scraped data contained some variables that required preprocessing and feature extraction. This step was essential to prepare the dataset for geocoding, classification, and energy modelling.

Power Usage Effectiveness (PUE). PUE is a key indicator of infrastructure efficiency, calculated as the total energy used by a data centre divided by the energy consumed by its IT equipment (Masanet et al., 2020). An extraction procedure was implemented to identify PUE values from unstructured provider descriptions, using patterns such as exact values (e.g., "PUE is 1.2"), inequality expressions (e.g., "PUE less than 1.3"), and intervals (e.g., "PUE between 1.2 and 1.4"). However, due to the commercial nature of the source data and the risk of containing promotional biases or lack of standardisation, these extracted values were not used in the final estimations. Instead, PUE values were assigned based on benchmark values stratified by facility size, as detailed in Section 3.2.1.

Whitespace and Building Size. The whitespace area (representing the usable IT floor space) and the total building size were extracted using regex patterns and standardised to square meters. This is one of the main features used for the energy estimation model.

Power Capacity. Values for power capacity were extracted and normalized to megawatts (MW). Multiple phrasings such as "available power," "built-out power," or "live capacity" were detected using pattern matching. Where power values were reported in kilowatts, they were converted to MW for consistency.

Filtering and Deduplication. From the initial dataset of 1,795 entries, 180 were excluded as they corresponded to listings marked as "under construction," "planned," or "archived," based on pattern detection in the description field. Country names were mapped to ISO-2 codes, city names were standardized, and each facility was assigned to one of four EU macro-regions (North, South, East, West). The cleaned dataset included 1,615 entries with fields such as operator, data centre name, address, whitespace, power capacity, PUE estimate, and standardized country and city identifiers.

3.1.3. Geocoding and Spatial Validation

After feature extraction and data cleaning, the dataset was geocoded. A geocoding pipeline was implemented using two open-source services, Photon and Nominatim. Facilities with ambiguous or missing addresses (49 entries) were geocoded using city centroids, using city and country information for the geocoding. The resulting coordinates, representing the geographic centroid of the respective city or postal region, were flagged as approximate facility locations.

In cases where both API services failed to return results and no valid city or country match was available due to vague address content, the facilities were removed, reducing the total number of facilities from 1,615 to 1,600. The final dataset contains only facilities with validated latitude and longitude coordinates within EU Member States.

3.1.4. Data Centre Classification

A critical step in the imputation process was to categorize each data centre into one of three operational categories: *hyperscale*, *colocation*, or *enterprise*. Hyperscale facilities are large-scale infrastructures typically owned and operated by major cloud providers (e.g., AWS, Google, Microsoft) and optimized for public cloud services and AI workloads. Colocation data centres are shared facilities where multiple clients house their servers within infrastructure managed by a third-party provider (European Commission. Joint Research Centre., 2020). Enterprise data centres are facilities owned by a single organization and operated for internal use (Pilz & Heim, 2023).

The classification was conducted in two stages. In the first stage, a rule-based approach assigned provisional categories based on refined keywords and operator names. Facilities were labeled as hyperscale if their descriptions included terms such as "hyperscale" or "cloud-scale" or if the operator matched a curated list of major cloud providers. Colocation facilities were identified using terms like "carrier neutral" or "multi-tenant" or matching known colocation providers. Enterprise facilities were labeled based on keywords such as "private," "corporate," or "on-premise," and served as the default when no other match was found (see Appendix A.1).

A supervised Natural Language Processing (NLP) model was trained to refine these rule-based labels in the second stage. Facility descriptions, operator names, and data centre names were cleaned, tokenized, lemmatized, and converted into TF-IDF feature vectors. A Random Forest classifier was trained on these features using an 80/20 random split. The model achieved a 97% overall accuracy on the validation set, with particularly strong performance for colocation and enterprise classes. The final classification output identified 801 data centres as enterprise, 669 as colocation, and 130 as hyperscale across the EU-27 dataset (see Appendix A.2).

3.1.5. Imputation of Missing Values

The final classification allowed the imputation of missing whitespace values, critical for estimating energy consumption, as described in Section 3.2.1. Several key attributes necessary for estimating energy consumption were missing in the final geocoded and cleaned dataset. In particular, whitespace was missing in 61% of the entries. For that reason, an imputation strategy was used to enable further modelling.

The imputation procedure followed two main steps. First, for data centres that lacked reported whitespace but provided total building size, usable whitespace was estimated as 50% of the gross building area, which was applied to 47 entries. For all remaining entries with missing whitespace and no reported building size, imputation was based on the median observed whitespace area within each data centre category. Medians were calculated from the non-missing values across enterprise, colocation, and hyperscale facilities, yielding 700 m^2 for enterprise, 1,500 m^2 for colocation, and 7,250 m^2 for hyperscale data centres. These type-specific medians were used to impute 929 additional entries.

3.1.6. Carbon Intensity

Carbon intensity data for 2024 were collected from Electricity Maps for each country bidding zone (Electricity Maps, 2025). Each zone's dataset included hourly values based on a life-cycle assessment (LCA) approach, measured in grams of CO_2e per kilowatt-hour (gCO_2e/kWh). Each zone's annual average carbon intensity was calculated by taking the mean of its hourly values.

Carbon intensity reflects the amount of carbon dioxide equivalent emitted per unit of electricity generated. Countries with a higher share of renewable energy in their grid mix tend to exhibit lower carbon intensities. The LCA-based metric used here captures emissions during generation and upstream emissions related to fuel extraction, transportation, and infrastructure.

Each data centre was geolocated and spatially matched to a bidding zone. The geometries of the bidding zones were sourced from the official Electricity Maps Zone Finder repository in GeoJSON format (Electricity Maps, 2024). After merging the zone geometries by name, a spatial join was performed to associate each data centre with its corresponding bidding zone. This allowed the assignment of average grid carbon intensity values specific to each zone to estimate the emissions.

3.2. Data Centre Energy Consumption and Emissions Estimation

3.2.1. Energy Consumption Estimation

An area-based formula was used to estimate each data centre's annual energy consumption, adapted from Jerléus et al. (2024). Each facility was assigned an area class based on thresholds for whitespace floor area (m^2). These classes were then used to assign IT Power Density (ITD) and Power Usage Effectiveness (PUE) values based on benchmarks taken from Jerléus et al. (2024), originally developed by Shehabi et al. (2011, 2016).

Annual electricity consumption was calculated using the following formula:

$$E = \left(\frac{\text{ITD}}{1000}\right) \times \text{PUE} \times \text{Area} \times 8760 \tag{3.1}$$

where E is the annual energy use in megawatt-hours (MWh), ITD is the average power density of IT equipment in kilowatts per square meter (kW/m^2), Area is the whitespace area in square meters (m^2), and 8,760 is the number of hours in a year. The division by 1,000 converts the total IT load (in kilowatts) to megawatts.

In Jerléus et al. (2024), ITD refers to the average electrical power consumed by IT equipment per square meter of data centre floor space (kW/m^2). It reflects the intensity of computational infrastructure hosted within a data centre. Higher ITD values indicate a greater concentration of servers or equipment per unit area, resulting in increased energy demand.

Three scenarios were defined: *low-energy*, *mid-energy*, and *high-energy* to model for different levels of efficiency. These reflect various operational conditions, from highly optimized to less efficient configurations. Specifically, the *low-energy* scenario assumes optimal infrastructure performance. It assigns the lowest PUE value for each area class, while the *high-energy* scenario assumes less efficient operation and assigns the highest PUE value. The *mid-energy* scenario adopts the baseline PUE values reported by Jerléus et al. (2024), and ± 0.20 margins are used to define the low- and high-energy scenarios. In all cases, ITD remained fixed across scenarios.

Hyperscale data centres were treated as a distinct category in the energy estimation model. Following Jerléus et al. (2024), these facilities are distinguished not by their physical size but by their operational characteristics, such as cloud-scale workload optimization and high infrastructure efficiency. For hyperscale facilities, a fixed IT power density (ITD) of 1.10kW/m^2 and a PUE value of 1.13 were applied. The complete specification of ITD benchmarks and scenario-based PUE values by area class is provided in Appendix A.3 and Appendix A.4. The final energy consumption of each facility was calculated in megawatt-hours per year (MWh/year) using the area-based formula. These values were then converted to terawatt-hours per year (TWh/year) to enable aggregation at the national and EU levels.

3.2.2. CO_2 Emissions Estimation

To estimate the annual CO_2 emissions associated with the electricity consumption of each data centre, they were computed by multiplying the data centres energy consumption with the average carbon intensity of electricity in its corresponding bidding zone.

The emissions per facility were calculated using the following formula:

$$\text{Emissions (tons CO}_2\text{e)} = E \times CI \tag{3.2}$$

where E is the annual electricity consumption expressed in terawatt-hours (TWh), and CI is the average grid carbon intensity in grams of CO_2-equivalent per kilowatt-hour (gCO_2e/kWh). To ensure unit consistency, the result was scaled by a factor of 1,000 to account for the conversion of energy from TWh to kWh and mass from grams to metric tons.

The emissions estimation was performed for each of the three energy consumption scenarios (*low*, *mid*, and *high*), and results were aggregated at both the national and EU levels. Final values were reported in metric tons (tCO_2e/year) and megatons ($MtCO_2e$/year) for comparison and reporting.

4. Results

The final cleaned and geocoded dataset contains 1,600 data centres across the EU-27, distributed among 801 enterprise (50.1%), 669 colocation (41.8%), and 130 hyperscale facilities (8.1%). Most data centres are concentrated in Western European countries, such as Germany (341), France (221), the Netherlands (148), Italy (129), and Spain (111). However, data centres tend to be distributed in clusters, specially in major urban areas and interconnection hubs.

4.0.1. Spatial Distribution and Clustering

To identify top clusters, a Density-Based Spatial Clustering of Applications with Noise (DBSCAN). The DBSCAN algorithm was selected for its ability to detect clusters of arbitrary shapes without requiring a predetermined number of clusters (Ester et al., 1996). The algorithm was applied to all geolocated facilities, using a 15 km radius and considering a minimum of 5 facilities to form a dense region. This yielded 73 unique clusters across the Union.

Figure 4.1 displays the resulting clusters, where the size of each point is proportional to the number of data centres in the cluster. The largest cluster is located in Frankfurt, Germany, containing 95 data centres, followed by Paris (74), Dublin (72), and Amsterdam (48). These cities, including London, are known as the "FLAP-D market hubs", Europe's main data centre hubs which concentrate the majority of capacity and activity in Western Europe (German Datacenter Association, 2024).

Figure 4.1.: Spatial distribution of clustered data centres

4.1. Energy Consumption

The estimated annual electricity consumption of EU-based data centres ranges from 54.5 TWh in the low scenario to 65.8 TWh in the high scenario, with a mid-scenario of 60.1 TWh. Figure 4.2 shows the estimated annual data centre electricity consumption by EU grouped by country and region under the mid-scenario, with intervals spanning the low and high scenarios. At the national level, the highest estimated consumption in the mid-scenario is observed in Germany (18 TWh), followed by the Netherlands (8.3 TWh), France (7 TWh), and Ireland (5.3 TWh). These countries also host some of the largest infrastructure clusters identified in Section 4.0.1.

The data centres in Western Europe have the largest energy consumption in the region with approximately 35.6 TWh/year (63% of EU-27 consumption), followed by Northern Europe (11.3 TWh/year, 20%), Southern Europe (8.6

TWh/year, 15%), and Eastern Europe (4.7 TWh/year, 8%). Table A.5 provides a detailed comparison of the estimations presented here and figures collected in Kamiya and Bertoldi (2024).

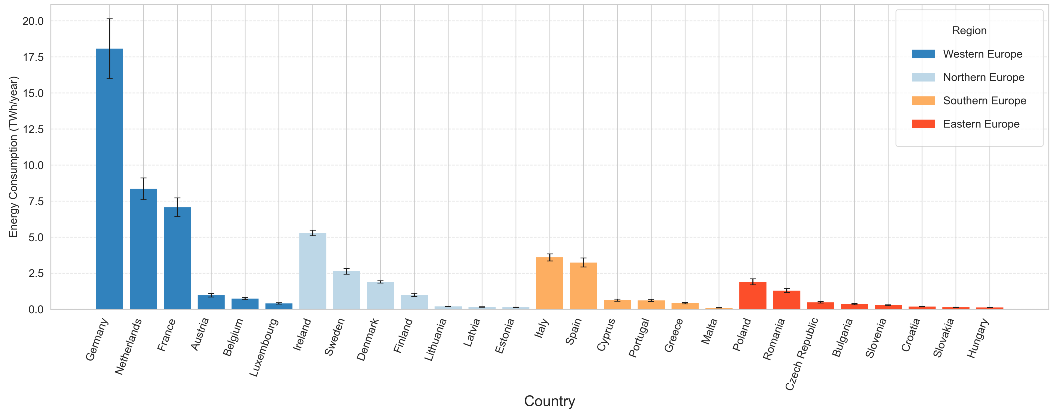

Figure 4.2.: Estimated Data Centre Electricity Consumption by Country and Region (Mid Scenario)

To contextualise data centre electricity consumption, the share of national electricity consumption attributable to data centres was calculated using aggregated final electricity consumption figures for 2021 from Eurostat (2022). Across all EU Member States, data centres account from 2.2% of the total energy consumed in 2022 in the low scenario to 2.6% in the high scenario, with 2.4% in the middle scenario.

At a national level, country shares vary considerably. For example, data centres account for about 18% of national electricity consumption in Ireland, 13.37% in Cyprus, and 7.73% in the Netherlands. In contrast, in larger countries such as Germany and France, the shares are 3.62% and 1.64%, respectively. In the case of Cyprus, the small size of the national electricity system likely amplifies the relative weight of data centre electricity consumption. These results may also reflect limitations in the underlying dataset and the imputation strategy used. Further discussion of these uncertainties is provided in Section 5.1.

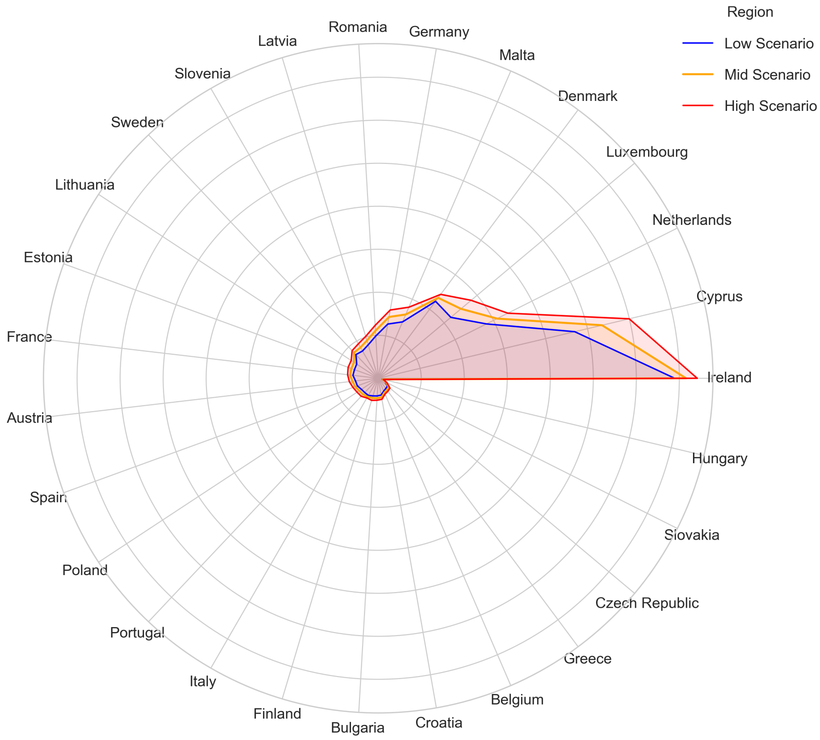

Figure 4.3.: Data Centre Share of National Electricity Consumption by Scenario

4.2. Carbon Emissions

Emissions associated with data centres in the EU were estimated using the energy consumption figures described in the previous section and carbon intensity averages (see Section 3.2.2). According to this estimation, the total data centre emissions in the member states range from 14.16 Mt CO_2e/year in the low scenario to 17.16 Mt CO_2e/year in the high scenario, with a mid-scenario of 15.66 Mt CO_2e/year.

At the national level, the distribution of emissions mirrors the patterns observed for energy consumption, with Germany leading by a substantial margin at approximately 6 Mt CO_2e/year, followed by the Netherlands (2.3 Mt CO_2e/year) and Ireland (2.1 Mt CO_2e/year). Countries like Poland and Italy also generate substantial emissions, with 1.3 Mt CO_2e/year and 0.9 CO_2e/year, respectively. In contrast, the Nordic countries, such as Sweden (0.05 Mt) and Finland (0.07 Mt), have relatively low emissions despite having data centre clusters. Figure 4.4 presents a map with the estimated data centre emissions across the EU27 under the mid scenario. The colour scale shows total emissions in megatonnes (Mt CO_2/year), with darker colours indicating higher emission levels. On top of the national emissions map, clustered markers are shown, representing concentrations of data centres identified through spatial analysis.

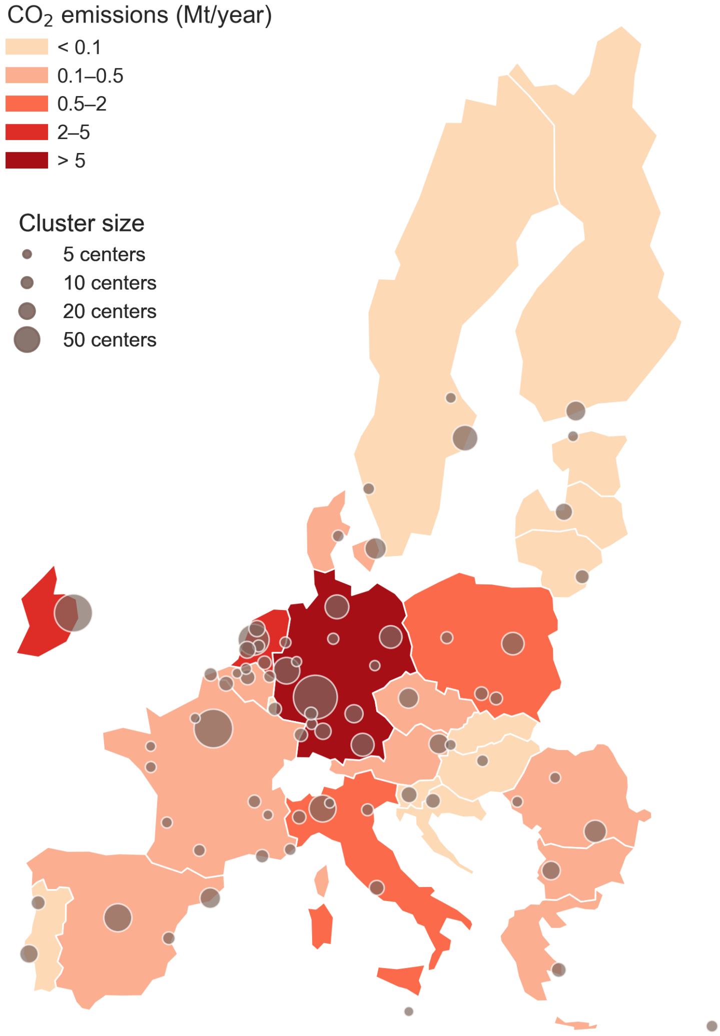

Figure 4.4.: Estimated CO_2 Emissions and Data Center Clusters in the EU (2024, Mid Scenario)

Figure 4.5 compares each country's estimated mid-scenario data centre energy consumption and CO_2 emissions on a log-log scale, allowing comparisons across several orders of magnitude. Since emissions are calculated as the product of energy consumption and carbon intensity, a strong positive correlation is expected. If carbon intensities were identical across countries, emissions would scale linearly with energy use. Deviations from the fitted trend line reflect differences in electricity carbon intensity. Countries positioned above the line, such as Poland, Ireland, and Cyprus, have more carbon-intensive grids. In contrast, countries below the line, such as Sweden, France, and Spain, have cleaner electricity mixes and, despite concentrating a large number of data centres and high energy consumption, exhibit lower emissions.

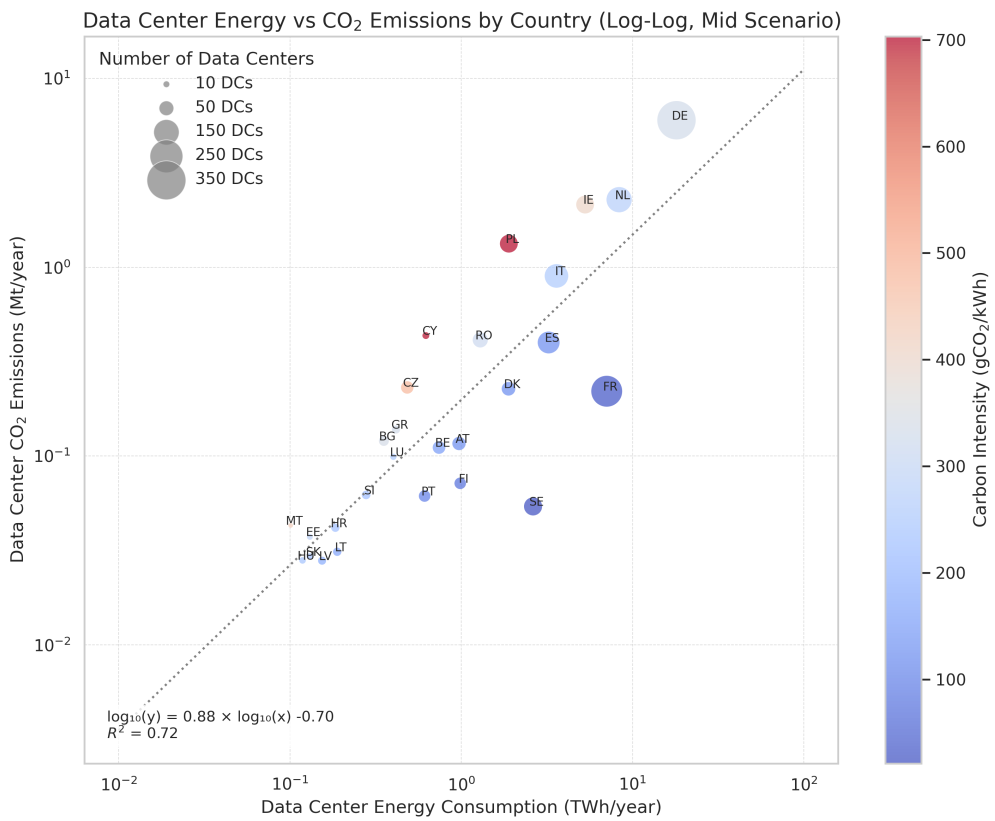

Figure 4.5.: Data Center Energy Consumption vs CO_2 Emissions by Country (Log-Log, Mid Scenario)

5. Discussion

The results show that data centres across the Union are principally located in four Western European countries: Germany, the Netherlands, France, and Ireland. However, they also tend to be located in clusters, the cities within the "FLAP-D" hub that concentrate most data centres. This reflects some of the characteristics of the market, where data centres tend to be located closely to benefit from existing infrastructure, access to Internet Exchange Points, and network effects. In addition, proximity to major metropolitan areas helps minimise latency and improve user connectivity.

The estimations also helped identify three possible scenarios for data centre energy consumption. The low scenario reflects an efficient operation with low PUE values, while the high scenario assumes a less efficient operation (high PUE values). The middle one assumes industry benchmarks.

According to this modelling, data centres in the EU-27 countries consume annually from 54.5 to 65.8 TWh, with 60 TWh in the middle scenario, accounting for 2.19% to 2.64% share of the total electricity demand in the region. These results are broadly consistent with previous assessments. For instance, Kamiya and Bertoldi (2024) reported that data centres in the European Union consumed between 45 and 65 TWh of electricity in 2022. More recently, International Energy Agency (2025) estimated that data centre energy consumption in Europe reached approximately 70 TWh in 2024, highlighting an increasing trend linked to the expansion of AI deployment.

Appendix A.5 contains a table summarizing previous estimates of data centre energy consumption at the national level for selected European countries, based on Kamiya and Bertoldi (2024), alongside the mid-scenario estimations from this study. The mid-scenario estimates produced here closely match previous assessments for countries such as Germany and Ireland. Germany's estimated 18.07 TWh/year aligns with the 17.9 TWh reported for 2022, while Ireland's 5.29 TWh/year compares very closely with the 5.3 TWh figure reported by the Central Statistics Office of Ireland for the same year.

However, significant differences exist for other countries where previous estimates are available. For instance, this study estimates France's data centre electricity consumption at 7.07 TWh/year, approximately 39% lower than the 11.6 TWh reported in previous estimations. Similarly, this study estimates 8.35 TWh/year for the Netherlands compared to earlier figures ranging between 2.7 and 6.3 TWh/year. Differences may reflect variations in the methodology, scope, or data used.

However, no prior national estimates were available for many EU-27 countries, including Austria, Bulgaria, Croatia, and Romania. In these cases, this study provides the first harmonized figures derived from a bottom-up and geospatial perspective, filling an important gap in the literature and offering a consistent baseline for future monitoring and policy evaluation.

Regarding the share of data centre energy consumption relative to national electricity use, the regional estimate of 2.2%–2.6% in this study broadly aligns with the figures reported by Kamiya and Bertoldi (2024), which range between 2.6% and 2.8%. At the national level, the estimates also are closely aligned, particularly for Ireland (17.9% in this study compared to 18% in Kamiya and Bertoldi (2024)), the Netherlands (7.7% vs. 5.2%), Luxembourg (6.2% vs. 4.8%), Denmark (5.8% vs. 4.5%), Germany (3.6% vs. 3%), Sweden (2% vs. 2.3%), and France (1.64% vs. 2.2%).

One important outlier to discuss is the Cyprus case, where, according to the estimations presented here, the energy consumption of data centres in this country is about 13.4% of its total energy consumption. In this case, only one facility of the eleven contemplated reported information necessary for the estimations, relying principally on the imputation strategy. The only facility that reported information is a big facility with around 30,000 square meters, representing around 65% of the estimated data centre energy consumption of this country. Additionally, Cyprus's total electricity consumption is relatively low compared to other EU Member States. This amplifies the weight of a few large data centres in the national figures.

Respecting the estimation of the carbon emissions associated with the data centre's energy consumption in the EU, they range from 14.16 Mt CO_2e/year in the low scenario to 17.16 Mt CO_2e/year in the high scenario, with a mid-scenario of 15.66 Mt CO_2e/year.

Appendix A.6 compares emissions estimates from the reviewed literature with those calculated in this report. The review reveals a substantial difference between US and EU emissions estimates. The most recent US figure from Guidi et al. (2024) reports 105.6 Mt CO_2-eq annually, approximately seven times higher than the EU-27 estimate of 15.66 Mt CO_2e/year. Although the EU hosts one of the largest concentrations of data centres globally, this disparity cannot be explained solely by the number of facilities. Instead, it reflects significant differences in carbon intensity between the two regions. The United States has a national grid average of 400 gCO_2e/kWh, while the EU-27 average is 246 gCO_2e/kWh, both measured on a LCA basis (Electricity Maps, 2025). This highlights the critical role of electricity grid composition in shaping the environmental impact of digital infrastructure. The relatively lower emissions in the EU-27 demonstrate the potential benefits of more decarbonized electricity grids in mitigating data centre climate impacts.

These effects are also present within EU countries. For example, in Poland, relatively high emissions result from a carbon-intensive electricity mix rather than from the size of the country's data centres. On the contrary, countries like France and Sweden show lower emissions even when they concentrate more facilities, reflecting low-carbon power systems.

5.1. Limitations

This research presents several limitations regarding data quality, methodological assumptions, and scope. The main limitation is the lack of reliable and publicly available data on data centres' performance and energy use. This limitation leads to relying on data from an industry listing webpage, which discloses information in marketing terms for selling proposes, and where operators could have limited incentives to transparently disclose detailed and accurate information about their energy metrics and infrastructure.

More importantly, probably not all data centres are publicly listed. This limitation impacts the completeness of the dataset and, in turn, the accuracy of the estimations. This is evident when comparing the Swedish data used in this study to that of Jerléus et al. (2024). While this analysis identifies 76 operational data centres in Sweden, Jerléus et al. include 130 facilities. Although both studies apply the same area-based methodology benchmarks, this study estimates national electricity consumption in Sweden at approximately 2.6 TWh/year, compared to 6.7 TWh/year reported by Jerléus et al.

Along with the missing facilities, this study presents other methodological limitations, such as the absence of crucial information and imputation strategy. About 60% of the facilities lacked whitespace information, a crucial measure for the area-based estimations. The classification of data centres, which was a key strategy for imputing those missing values, was performed using a natural language processing (NLP) approach based on facility descriptions, which as discussed before, were filled with marketing content, not necessarily reliable operational information. While the NLP modelling and random forest classification achieved a strong overall accuracy of 97%, it exhibited a relatively low performance in classifying hyperscale facilities. This likely reflects the class imbalance in the dataset and the conservative nature of the supervised classifier, which can bias imputation.

Additionally, the energy estimations in this study were conducted using an area-based approach, drawing directly from Jerléus et al. (2024). While appropriate given the available data, this method does not fully capture actual operational energy use or other characteristics important for bottom-up estimations, such as power capacity, IT load profiles, infrastructure configurations, and utilization rates. A fixed IT power density of 1.10 kW/m^2 was assigned to hyperscale facilities following Jerléus et al. (2024). However, this may underestimate energy use, particularly in facilities deploying accelerated servers for AI workloads, which could have much higher power densities. In addition, the model does not incorporate advancements in cooling technologies, such as liquid cooling systems, which can significantly improve efficiency and reduce total facility energy consumption. Further research is needed to continuously update these benchmarks and better capture the evolving impacts of AI on electricity demand and associated emissions.

Regarding the broader environmental impacts, in the developed model, emissions estimates are linearly derived from modelled energy consumption. Therefore, any over- or underestimation in energy consumption directly propagates into the emissions figures. Additionally, the model accounts only for electricity related emissions and does not include emissions related to construction, embodied carbon from equipment and materials, or emissions from backup generator usage.

Finally, renewable energy procurement strategies, such as Power Purchase Agreements (PPAs), are also not incorporated, as a location-based estimation approach was used, relying on the average grid carbon intensity of each facility's region. While this allows for a spatially explicit analysis, it does not capture corporate decarbonization efforts that may partially mitigate facility-level emissions. A final and important limitation is the omission of water usage, which is a critical environmental indicator for data centre cooling infrastructure. The exclusion of water intensity considerations limits the comprehensiveness of the environmental impact assessment. However, these aspects fall outside the scope of this study and could be addressed in future research.

5.2. Policy Implications for AI Deployment and Sustainability in the Region

The estimations developed here can help lead a discussion on the policy implications of AI development in the region and identify where, why, and how data centres create challenges or opportunities for sustainable AI deployment in the Member States.

A crucial step in understanding and improving estimation models of electricity consumption and emissions is access to reliable data. The revised Energy Efficiency Directive (EED) of the European Union, implemented through Delegated Regulation (EU) 2024/1364, requires data centres with an installed IT power demand of at least 500 kW to report relevant energy-related performance indicators annually to a dedicated European database, starting in September 2024 (European Commission, 2024). These reporting obligations aim to improve transparency and support improvements in energy efficiency across the sector. According to the regulations, the data will be collected through the European database. They will be publicly available in aggregated form, including information such as the number of reporting data centres, total installed IT power, data centre size categories, total energy consumption, and average PUE. From a policy perspective, this new reporting framework will enhance the sector's transparency and enable more precise monitoring of compliance with EU climate and energy targets.

While access to reliable data is an important step for assessing energy consumption and emissions, broader policy concerns remain regarding the rising energy consumption of data centres. Although the impacts of AI deployment on energy consumption are difficult to quantify, recent projections from the International Energy Agency (IEA) estimate that global electricity consumption by data centres could reach around 945 TWh by 2030 under the Base Case scenario. In this projection, the electricity consumption in accelerated servers will grow by 30% annually, compared to 9% for conventional servers, meaning that accelerated servers could account for about half of the net increase in global data centre electricity consumption (International Energy Agency, 2025).

The impacts of AI-driven workloads in energy consumption estimations present relevant challenges for energy supply. AI is developing rapidly, but the speed at which energy infrastructure can grow differs. In this sense, there is a path dependency problem, where electricity grids are under strain in many places, and there is a lack of infrastructure capable of supporting new data centre developments and energy requirements. For example, connection queues can take up to ten years in the Netherlands and up to seven years in Germany (International Energy Agency, 2025).

Several initiatives can be implemented to mitigate the risks, such as strategically incentivise data centre locations in areas with higher renewable energy disposition and grid availability. Data centres in Europe already emit less CO_2e/year compared to other regions such as the US. This gives the EU a strategic advantage in aligning the sector with its climate goals. However, infrastructure bottlenecks could undermine this advantage, particularly in regions with limited grid capacity or delayed renewable deployment.

As sustainability becomes a growing concern in the European Union, locating data centres close to renewable energy sources can help reduce the environmental impact of the projected AI-related increase in energy demand. Countries like Sweden, Finland, and Spain, with abundant renewable resources and available land, are well-positioned to attract new developments.

Optimistically, AI could also be used to reduce emissions by accelerating innovation in the industry. However, it remains unclear whether the gains from AI adoption will be sufficient to outweigh the increase in emissions from data centres, or whether rebound effects could neutralise these benefits. Meanwhile, robust data collection, investment in resilient digital infrastructure, and the strategic siting of data centres near renewable energy sources will be essential to ensure that the expansion of AI deployment in the region remains compatible with long-term climate goals.

6. Conclusion

This study provides one of the first systematic, bottom-up estimates of data centre energy consumption and associated carbon emissions across EU Member States. Using scraped data from a public industry listing, the results show that data centre electricity consumption across the EU-27 is estimated to range between 54.5 and 60 TWh per year, representing approximately 2.2% to 2.6% of total electricity demand. Carbon emissions related to this consumption are estimated between 14.16 and 17.16 Mt CO_2e per year.

While the highest energy-consuming countries are predominantly located in Western Europe, emissions vary significantly depending on the carbon intensity of national electricity grids. Countries such as Poland, Cyprus, and Ireland exhibit higher carbon intensities than Sweden and France, which benefit from cleaner electricity mixes.

These findings highlight the importance of considering not only the scale of energy consumption but also the underlying energy mix in assessing the environmental impacts of digital infrastructure. As AI-driven workloads are expected to grow and the region intends to position itself as an AI-competitive market, policy measures that promote strategic planning of data centres near renewable energy sources, improve infrastructure capacity, and enforce transparent energy reporting will be crucial to ensure that digital sector expansion remains compatible with EU climate goals.

A. Appendix

Table A.1.: Rule-based classification logic for data center type classification

Type	**Keyword Indicators**	**Known Operators**
Hyperscale	hyperscale, hyperscaler, global cloud, cloud-scale, hyperscale facility, hyperscale infrastructure, hyperscale data center, hyperscale data centre	Amazon, AWS, Google, Microsoft, Azure, Meta, Facebook, Apple, Alibaba, Tencent, Oracle, IBM, Baidu, Huawei, OVH, Switch, Cloudflare, Salesforce
Colocation	colocation, co-location, colo, carrier-neutral, multi-tenant, wholesale, neutral facility	Equinix, Digital Realty, NTT, CyrusOne, QTS, CoreSite, Global Switch, Interxion, Telehouse, Colt, EdgeConneX, Vantage, Iron Mountain, Atman, Data4, Dataplex, Nexcenter
Enterprise	enterprise, private, in-house, on-premise, corporate, dedicated data center, dedicated facility	—

Table A.2.: Performance of supervised classification model (Random Forest)

Class	**Precision**	**Recall**	**F1-score**
Colocation	0.98	0.97	0.97
Enterprise	0.96	1.00	0.98
Hyperscale	0.96	0.79	0.86
Overall Accuracy	0.97		

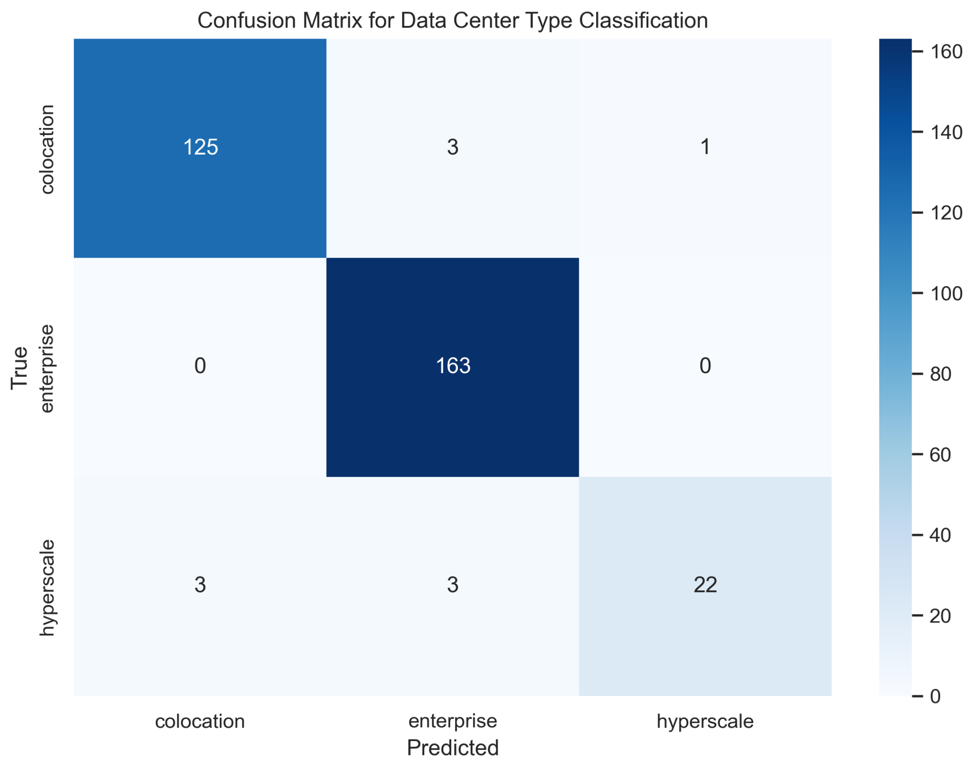

Figure A.1.: Confusion matrix of supervised classification model for data centre types (Random Forest, 80/20 train-test split).

Table A.3.: Area class definitions and energy benchmark parameters. Adapted from Jerléus et al. (2024), based on **Shehabi2016**; Shehabi et al. (2011).

Class	Whitespace Range (m^2)	IT Power Density (kW/m^2)	PUE
A	< 9.3	—	2.00
B	9.3 – 92.8	0.43	2.35
C	92.8 – 185.7	0.65	1.88
D	185.7 – 1858	0.86	1.79
E	> 1858	1.10	1.60
Hyperscale	Not applicable	1.10	1.13

Table A.4.: Scenario-based PUE values by area class.

Class	PUE (Low-Energy)	PUE (Mid-Energy)	PUE (High-Energy)
A	2.00	2.00	2.00
B	2.15	2.35	2.55
C	1.68	1.88	2.08
D	1.59	1.79	1.99
E	1.40	1.60	1.80
Hyperscale	1.13		

Note: The mid-energy scenario uses the original benchmarks from Jerléus et al. (2024). Low and high scenarios apply a ± 0.20 margin to represent more and less efficient operations.

Table A.5.: Overview of studies estimating the energy use of data centres in Europe and estimations of this study, adapted from Kamiya and Bertoldi (2024)

Country	Publication	Estimation	This Study
Austria	–	–	0.97 TWh

(Continued on next page)

(Continued from previous page)

Country	Publication	Estimation	This Study (Mid scenario 2025)
Belgium	Belgian Digital Infrastructure Association (2022)	0.38 TWh in 2021	0.74 TWh
Bulgaria	–	–	0.35 TWh
Croatia	–	–	0.18 TWh
Cyprus	–	–	0.62 TWh
Czech Republic	–	–	0.48 TWh
Denmark	COWI (2021)	0.88 TWh in 2020	1.89 TWh
	Danish Energy Agency (2021–2023)	1.1 TWh in 2021	
Estonia	–	–	0.13 TWh
Finland	Hiekkanen et al. (2021)	0.25 TWh in 2018	0.99 TWh
France	Ademe & Arcep (2022)	11.6 TWh in 2020	7.07 TWh
	Bordage et al. (2021)	5.2 TWh in 2020	
	CITIZING (2020)	9 TWh in 2019	
Germany	Hintemann et al. (2020–2023)	17.9 TWh in 2022	18.07 TWh
	BloombergNEF et al. (2021)	7.2 TWh in 2021	–
Greece	–	–	0.42 TWh

(Continued on next page)

(Continued from previous page)

Country	**Publication**	**Estimation**	**This Study (Mid scenario 2025)**
Hungary	–	–	0.12 TWh
Ireland	CSO Ireland (2021–2023)	5.3 TWh in 2022	5.29 TWh
	BloombergNEF et al. (2021)	4.7 TWh in 2021	
Italy	–	–	3.60 TWh
Latvia	–	–	0.15 TWh
Lithuania	–	–	0.19 TWh
Luxembourg	–	–	0.40 TWh
Malta	–	–	0.10 TWh
Netherlands	Statistics Netherlands (2021)	2.7 TWh in 2019	8.35 TWh
	BloombergNEF et al. (2021)	6.3 TWh in 2021	
Poland	–	–	1.90 TWh
Portugal	–	–	0.61 TWh
Romania	–	–	1.29 TWh
Slovakia	–	–	0.13 TWh
Slovenia	–	–	0.28 TWh
Spain	–	–	3.24 TWh
Sweden	Swedish Energy Agency (2023); Radar (2020)	2.8–3.2 TWh in 2022	2.63 TWh
	Radar (2020)	2.4 TWh in 2020	

Table A.6.: Comparison of data centre CO_2-equivalent emissions across regions and studies.

Study	**Region**	**Year**	**Energy Use (TWh)**	**Emissions (Mt CO_2-eq)**	**% of Global/Regional GHG**
International Energy Agency (2025)	Global	2024	415	180	0.5%*
Siddik et al. (2021)	USA	2018	72.2	31.5	0.5%
Guidi et al. (2024)	USA	2022	192.6	105.6	1.66%
Jerléus et al. (2024)	Sweden	2020	6.7	0.13	0.3%
This study (Mid-scenario 2024)	EU-27	2024	59.0	15.6	0.45%**

* Percentage of global fuel combustion CO_2 emissions.

** Based on total EU GHG emissions of approximately 3,400 MtCO_2-eq in 2022.

Bibliography

Anthony, L. F. W., Kanding, B., & Selvan, R. (2020). Carbontracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models. https://doi.org/10.48550/ARXIV.2007.03051

Budennyy, S. A., Lazarev, V. D., Zakharenko, N. N., Korovin, A. N., Plosskaya, O. A., Dimitrov, D. V., Akhripkin, V. S., Pavlov, I. V., Oseledets, I. V., Barsola, I. S., Egorov, I. V., Kosterina, A. A., & Zhukov, L. E. (2022). eco2AI: Carbon Emissions Tracking of Machine Learning Models as the First Step Towards Sustainable AI. *Doklady Mathematics*, *106*(S1), S118–S128. https://doi.org/10.1134/S1064562422060230

Courty, B., Schmidt, V., Goyal-Kamal, MarionCoutarel, Feld, B., Lecourt, J., LiamConnell, SabAmine, inimaz, supatomic, Léval, M., Blanche, L., Cruveiller, A., ouminasara, Zhao, F., Joshi, A., Bogroff, A., Saboni, A., de Lavoreille, H., ... MinervaBooks. (2024, May). Mlco2/codecarbon: V2.4.1. Retrieved April 24, 2025, from https://zenodo.org/doi/10.5281/zenodo.11171501

Data Center Map. (2025). Data Center Map. Retrieved March 19, 2025, from https://www.datacentermap.com

Electricity Maps. (2024). Electricity Maps Zone Finder (GeoJSON Repository). Retrieved April 16, 2025, from https://github.com/electricitymaps/zone-finder/tree/main

Electricity Maps. (2025). Electricity Maps. Retrieved April 15, 2025, from https://portal.electricitymaps.com/datasets

Epoch AI. (2024, June). Data on Notable AI Models. Retrieved March 26, 2025, from https://epoch.ai/data/notable-ai-models

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. https://dl.acm.org/doi/10.5555/3001460.3001507

European Commission. (2019). The European Green Deal. Retrieved March 26, 2025, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A640%3AFIN

European Commission. (2024). Delegated Regulation (EU) 2024/1364 of the Commission. Retrieved April 25, 2025, from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401364

European Commission. (2025). EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence. Retrieved March 27, 2025, from https://ec.europa.eu/commission/presscorner/detail/en/https:%5C/%5C/ec.europa.eu%5C/commission%5C/presscorner%5C/detail%5C/en%5C/ip_25_467

European Commission. Joint Research Centre. (2020). *Development of the EU Green Public Procurement (GPP) criteria for data centres, server rooms and cloud services: Final technical report*. Publications Office. Retrieved April 8, 2025, from https://data.europa.eu/doi/10.2760/964841

Eurostat. (2022). Final energy consumption by sector. https://doi.org/10.2908/TEN00124

German Datacenter Association. (2024). European Data Center Overview. Retrieved April 25, 2025, from https://www.germandatacenters.com/fileadmin/documents/publications/European_Data_Center_Overview.pdf

Guidi, G., Dominici, F., Gilmour, J., Butler, K., Bell, E., Delaney, S., & Bargagli-Stoffi, F. J. (2024, November). Environmental Burden of United States Data Centers in the Artificial Intelligence Era. https://doi.org/10.48550/arXiv.2411.09786

Heguerte, L. B., Bugeau, A., & Lannelongue, L. (2023). How to estimate carbon footprint when training deep learning models? A guide and review. *Environmental Research Communications*. https://hal.science/hal-04120582/file/guide_arxiv.pdf

International Energy Agency. (2025). *Energy and AI* (tech. rep.). Retrieved April 25, 2025, from https://iea.blob.core.windows.net/assets/dd7c2387-2f60-4b60-8c5f-6563b6aa1e4c/EnergyandAI.pdf

Jerléus, K., Ibrahim, M. A., & Augustsson, A. (2024). Environmental footprints of the data center service sector in Sweden. *Heliyon*, *10*(11), e31290. https://doi.org/10.1016/j.heliyon.2024.e31290

Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, *12*(6), 518–527. https://doi.org/10.1038/s41558-022-01377-7

Kamiya, G., & Bertoldi, P. (2024). Energy Consumption in Data Centres and Broadband Communication Networks in the EU. (KJ-NA-31-841-EN-N (online)). https://data.europa.eu/doi/10.2760/706491

Lannelongue, L., Grealey, J., & Inouye, M. (2021). Green Algorithms: Quantifying the Carbon Footprint of Computation. *Advanced Science*, *8*(12), 2100707. https://doi.org/10.1002/advs.202100707

Luccioni, A., Viguier, S., & Ligozat, A.-L. (2022). Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model. https://arxiv.org/pdf/2211.02001

Masanet, E., Shehabi, A., Lei, N., Smith, S., & Koomey, J. (2020). Recalibrating global data center energy-use estimates. *Science*, *367*(6481), 984–986. https://doi.org/10.1126/science.aba3758

Pilz, K., & Heim, L. (2023). Compute at Scale: A Broad Investigation into the Data Center Industry. https://doi.org/10.48550/ARXIV.2311.02651

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, *63*(12), 54–63. https://doi.org/10.1145/3381831

Shehabi, A., Masanet, E., Price, H., Horvath, A., & Nazaroff, W. W. (2011). Data center design and location: Consequences for electricity use and greenhouse-gas emissions. *Building and Environment*, *46*(5), 990–998. https://doi.org/10.1016/j.buildenv.2010.10.023

Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., Masanet, E., Horner, N., Azevedo, I., & Lintner, W. (2016, June). *United States Data Center Energy Usage Report* (tech. rep. No. LBNL–1005775, 1372902). https://doi.org/10.2172/1372902

Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., Masanet, E., Horner, N., Azevedo, I., & Lintner, W. (2024). *United States Data Center Energy Usage Report* (tech. rep. No. LBNL–1005775, 1372902). Retrieved April 10, 2025, from http://www.osti.gov/servlets/purl/1372902/

Siddik, M. A. B., Shehabi, A., & Marston, L. (2021). The environmental footprint of data centers in the United States. *Environmental Research Letters*, *16*(6), 064017. https://doi.org/10.1088/1748-9326/abfba1

Strubell, E., Ganesh, A., & McCallum, A. (2019, June). Energy and Policy Considerations for Deep Learning in NLP. https://arxiv.org/pdf/1906.02243