Power sector: Emissions from Electricity Generation

Jeremy Freeman[1,4], Ali Rouzbeh Kargar[1,4], Heather D. Couture[2,4], Jeyavinoth Jeyaratnam[1,4], Jordan Lewis[1,4], Madison Alvara[1,4], Hannes Koenig[1,4], Tiffany Nakano[3,4], Aaron Davitt[1,4], Christy Lewis[1,4], and Gavin McCormick[1,4]
1) WattTime, 2) Pixel Scientia Labs, 3) Global Energy Monitor, 4) Climate TRACE

The methodology provides changes and updates since our winter 2024 journal article by Couture et al. (2024): "Estimating Carbon Dioxide Emissions from Power Plant Water Vapor Plumes Using Satellite Imagery and Machine Learning" – please see this document for more detailed description of the method employed for this sector.

1. Introduction

The mission of Climate TRACE is to make meaningful climate action faster and easier. We do this by leveraging all tools at our disposal to track greenhouse gas (GHG) and, for the first time this year, non-GHG air pollutant emissions with high detail and speed, delivering information that is relevant to all parties working to reduce global emissions.

The sector which contributes the most to human-caused climate change is the energy sector. Depending on the data source, from 2015 to 2020, the energy sector contributed, on average, ~76% of global anthropogenic CO_2 emissions, or between 33 and 37 $GtCO_2$ per year [1,2,3]. Within the energy sector, electricity generation by combustion power plants (plants which burn fossil fuels, biomass, and/or waste to generate electricity) is responsible for the majority of GHG emissions. This sub-sector alone accounts for an average of ~15 $GtCO_2$, representing ~31% of total global GHG emissions from 2015 to 2020 [1].

The environmental impact of combustion power plants goes beyond GHGs; however, they also release other harmful pollutants into the air and water which pose a major health hazard to people, animals, and ecosystems [4]. Along with GHG emissions, Climate TRACE is providing and tracking the following air pollutants, non-GHGs, associated with combustion power plants: sulfur dioxide (SO_2), nitrogen oxides (NO_X), and PM2.5. Although this method is in its early stages of development, ongoing improvements and collaboration with the scientific community will improve these non-GHG emissions estimates over time.

Climate TRACE has developed a novel approach to independently track an individual power plants' emissions: we use proxy signals in visible satellite imagery (i.e., water vapor plumes co-released with emissions) and a harmonized power plant database pulling from several inventories (see Table 1 below) to estimate power plant electricity generation, which can be translated to emissions. Where available and deemed good quality by a series of data validation tests, we publish electricity generation data from reporting entities (currently ONS, ENTSO-E, and CAMPD; see Table 1). We obtained satellite imagery (Landsat 8 C2, Sentinel-2H, and PlanetScope PSScene) for each power plant from 2015 to most recently available imagery, then trained machine learning (ML) models on reported electricity generation data, rather than emissions, as generation is far more widely-available at hourly resolution to align with satellite imagery. Our models are thus designed to estimate electricity generation first and foremost; we then apply power plant specific emissions factors to infer CO_2, SO_2, NO_X, and PM2.5 emissions.

Table 1 Datasets employed to create a harmonized Global Combustion Power Plant Inventory

Dataset Name	Plant Metadata Used	Published (use varying by source/country)
Brazil ONS (Operador Nacional do Sistema Elétrico)	electricity generation	Source Level Data, Aggregated into Country Level Data
ENTSO-E (European Network of Transmission System Operators for Electricity)	electricity generation, operation start/end dates, capacity	Source Level Data, Aggregated into Country Level Data
USA CAMPD (Clean Air Markets Program Data)	electricity generation, operation start/end dates, capacity	Source Level Data, Aggregated into Country Level Data
US Energy Information Administration EIA-860, EIA-860m	Plant Name, Unit Fuel Type, Location, Unit Capacity, Unit Operating Dates, Unit Cooling type, Unit Pollution Control Tech SO_2	Source Level Data, Aggregated into Country Level Data
World Resources Institute (WRI) Global Power Plant Database (GPPD)	Plant Name, Plant Fuel Type, Location, Plant Capacity, Plant Operating Dates	Source Level Data, Aggregated into Country Level Data
S&P Global/Platts World Electric Power Plant (WEPP) database	Unit Fuel Type, Unit Capacity, Unit Operating Dates, Unit Cooling Type, Unit Pollution Control Tech SO_2	The source level dataset is proprietary and is used internally only.
Global Energy Monitor (GEM) Global Coal Plant Tracker (GCPT), Global Oil and Gas Plant Tracker (GGPT), and Global Bioenergy Power Tracker (GBPT)	Plant Name, Unit Fuel Type, Location, Unit Capacity, Unit Operating Dates	Source Level Data, Aggregated into Country Level Data
Other Sources (e.g., press releases, newspaper articles, company websites)	All	Source Level Data, Aggregated into Country Level Data

Overall, Climate TRACE tracks over 35,500 unique power plants and provides GHG and non-GHG emissions estimates at the annual and monthly level within 2-3 months of recorded activity. Additionally, we publish capacity, capacity factor, electricity generation, fuel type, geolocation, and plant names for the over 9,000 largest plants worldwide where known.

2. Materials and Methods

2.1 Overview

Climate TRACE has published monthly electricity generation, carbon emissions, and non-GHG air pollutant (SO_2, NO_X, and PM2.5) emissions at both the country level and at the asset level for global coverage of combustible power plants. We focus on combustible, also called thermal, power because this is the set of power plants that burn fuel to generate electricity and therefore directly generate emissions. Combustible power can be broken into three general categories: fossil (e.g. coal, gas, oil), waste (e.g. municipal garbage, tires) and biomass (e.g. trees, biogas). Burning biomass for power represents ~4% of carbon emissions for the power sector 2015-present, but since biomass carbon emissions are encompassed by Climate TRACE forestry and land use sectors, we subtract biomass-related carbon emissions out from our overall CO_2 emissions estimates to avoid double-counting. However, we continue to provide monthly biomass-related carbon emissions via the "other_4" column of our published asset-level data, which can be summed with the CO_2 column to get total combustible CO_2 emissions. Note that our capacity, capacity factor, electricity generation and non-GHG air pollutant emissions estimates (SO_2, NO_X, and PM2.5) do include biomass, however, because these are not accounted for by the other Climate TRACE sectors.

Climate TRACE publishes electricity generation data from reporting entities (currently ONS, ENTSO-E, and CAMPD; see Table 1) for power plants where the data is available and passes validation tests. The tests evaluate the reported electricity data for completeness (e.g. missed hours or days of reporting, units failing or not required to report, unrealistic values). They are run at the plant-month level, but we currently only publish the reported generation data for plants that pass the tests for all months from January 2019 to present, allowing for up to 20% missing reported data within any given month (which we gap-fill using the average for that month).

For reporting entities with a delay in reporting the most recent months, we impute the missing data. For reporting entities that publish net generation, we convert it to gross generation. We assign higher confidence and lower uncertainty values to emissions estimates derived from this reported data, with lower confidence and higher uncertainty for plants with imputed missing reported data (see breakdown in Table S3). Improvements to the tests and efforts to expand usage of reported electricity data from additional plants and regions are ongoing.

For those power plants where reported electricity generation data is not used Climate TRACE's baseline approach for estimating electricity generation at individual power plants is to synthesize the unit-level capacity, fuel, and prime mover information contained in our harmonized power plant inventory for each plant (see Couture et al 2024, Appendix A.1) with EIA and EMBER Yearly Electricity Data country-level annual estimates of capacity and generation by fuel type to calculate the annual average fuel-specific capacity factor in each year reported for each country in the world. We then assume the same capacity factors within each country for each part of the plant with the associated fuel-type. This baseline method lacks detail and is therefore reported with higher uncertainties; however, it is helpful as it is applicable to every power plant in the world.

To augment this baseline, Climate TRACE leverages remotely-sensed (RS) proxy signals to estimate power generation at the monthly level for individual power plants. By training machine learning (ML) models to understand the association between electricity generation and water vapor plumes from stacks (in layman's terms, the exhaust pipes of a power plant) with flue gas desulfurization (a common pollution control for coal), wet natural draft cooling towers (the large hyperbolic structures commonly seen primarily at coal plants), and – for the first time this year – mechanical draft cooling towers (fan-like structures, common across coal and gas). Figure 1 provides an overview of the remote sensing and machine learning (RS+ML) proxy signal approach employed to estimate individual power plant emissions. Figure 2 provides an example of the RS+ML approach applied to two observations of a particular power plant. For more detailed explanations of this proxy signal approach, referred to as "RS+ML" in this document, please see Couture et al. (2024). Note however that the structures which produce water vapor plume proxy signals do not exist at every power plant in the world. Yet, the 4% of plants worldwide that do have these structures amount to nearly half (43%) of carbon dioxide emissions from all non-biomass combustion (fossil and waste) power generation 2015-2024.

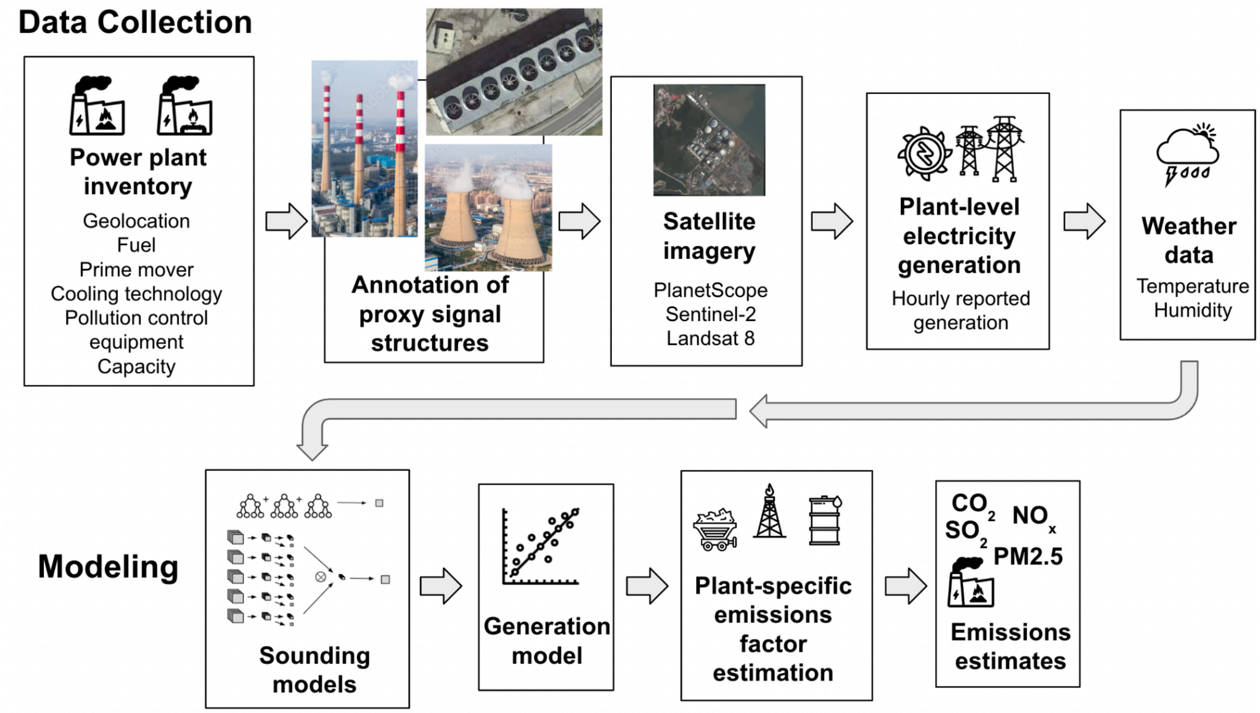

Figure 1: Diagram of RS+ML-enabled proxy signal modeling for monitoring power plant activity.

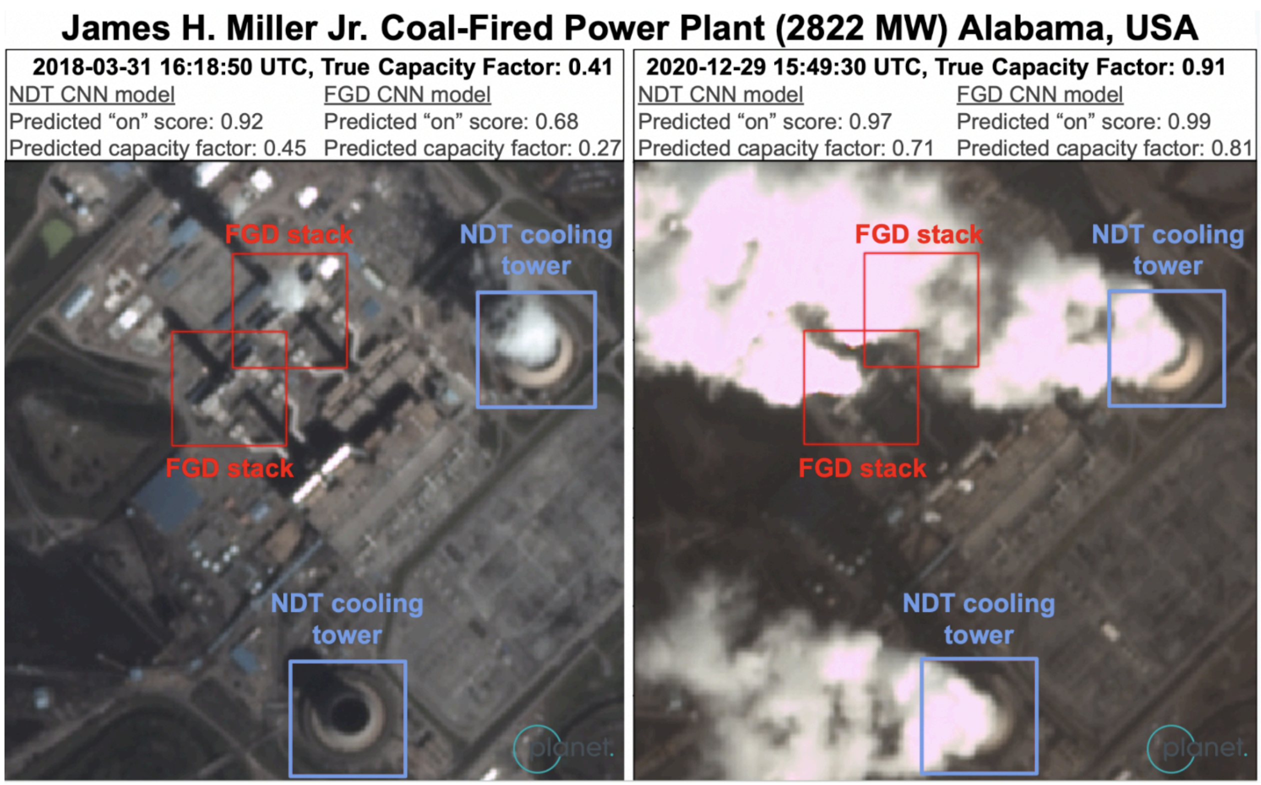

Figure 2: ML predictions on the James H. Miller Jr. power plant at low vs. high generation on two observation dates, using PlanetScope PSScene imagery. Separate natural draft cooling tower (NDT) and flue gas desulfurization

(FGD) models predicted on NDT (blue) and FGD stacks (red) patches. These predictions were ingested by subsequent models to estimate generation, then CO_2, SO_2, NO_X, and PM2.5, for the plant. © 2023 IEEE [8].

Our final published electricity generation estimates are the ensemble of the RS/ML-enabled proxy signal estimates (where available) and the country-, fuel-, and prime-mover-specific averages (Figure 2). Climate TRACE asset-level data provides a flag to indicate whether the estimate leveraged the remote sensing and machine learning approach (the column “other1” contains either the string “i” if the estimate leveraged ML, or the string “a” if it did not). Additionally, non-GHG emissions were generated for 2015 to 2024. See section 2.2.1 “Non-GHG Air Pollutant Emissions” for more information.

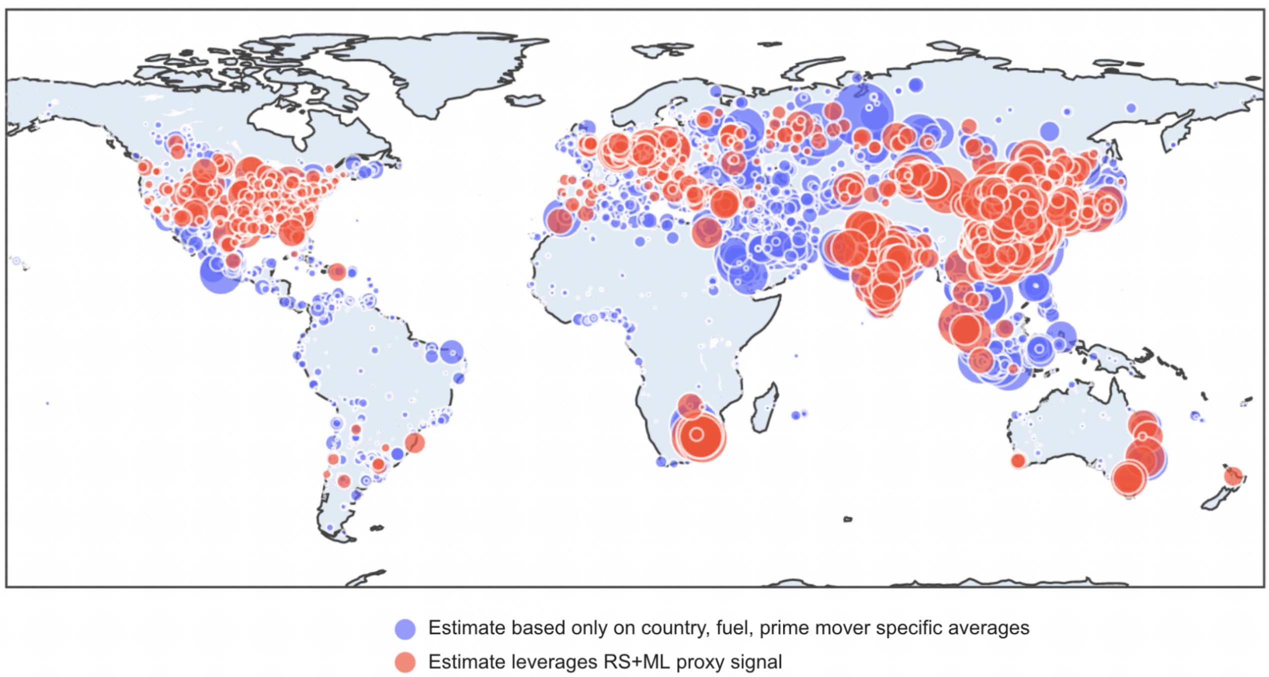

Figure 3: Asset-level Climate TRACE carbon dioxide emissions estimates for January 1, 2023, to December 31, 2024, color-coded by whether the estimate incorporated the RS+ML proxy signal modeling (red dots) method or country, fuel, and prime mover specific averages alone (blue dots). Each dot represents one powerplant, sized according to the amount of carbon emissions, ranging from 80 to nearly 35,000,000 metric tons of CO_2.

Climate TRACE provides both facility-level and country-level CO_2, SO_2, NO_X, and PM2.5 combustion power plant emissions estimates. Facility-level estimates were published at the monthly level, January 2019 through the end of July 2024, for all plants with capacity 100 MW or greater for which we know the latitude and longitude coordinates, totaling over 6000 plants. We also have an additional ~3000 plants under 100 MW capacity but for which we have capacity, latitude and longitude, so we also published these at the facility-level as long as their capacity was at least 10 MW. This constituted just over 9000 unique plants reported at facility-scale (up 8% from last year’s, 2023, Climate TRACE release) across 176 countries, totaling 12.9 $GtCO_2$ in 2023 which amounts to ~96% of total non-biomass global combustible power CO_2 emissions. Country-level estimates are published at monthly resolution (January

2015 through July 2024, inclusive) and represent our best estimate of all power plant emissions in each country, with no lower limit on capacity nor need to have precise location, capacity, name, etc. information. These country-level estimates are provided for all 250 countries in the world, totaling 35 573 plants which collectively amount to 13.4 GtCO_2 (non-biomass) and 16.7 MtSO_2, 23.5 MtNO_x, and 1.1 Mt$PM_{2.5}$ (fossil, waste, biomass) in 2023.

2.2 2024 Methodology Updates

Our 2024 methodology updates, new since our data published to the Climate TRACE website in 2023 (upon which our early 2024 journal article was based) are expanded upon in the following sections:

2.2.1 Non-GHG Air Pollutant Emissions

The combustion of fossil fuels, biomass, and waste to generate electricity emits air pollutants including sulfur dioxide (SO_2), nitrogen oxides (NO_X) and particulate matter (PM 2.5) which are linked to adverse health outcomes. We estimate emission factors for SO_2, NO_X, and PM 2.5 at the power plant unit-level based on fuel type, boiler type, and pollution control technology. Emissions factors are sourced from the 2023 EMEP/EEA air pollutant emission inventory guidebook and 2023 EIA Electric Power Annual. We then multiply these emission factors with the Climate TRACE estimated power generation for each unit to obtain SO_2, NO_X, and PM 2.5 emissions estimates. Fuel, boiler type and pollution control technology are obtained from our harmonized power plant information inventory, while the pollution control technology efficiency is based on the US EPA's AP-42 as a primary source.

Pollution control technology is relevant because emissions are reduced when a plant operates its control equipment. In the absence of pollution control information, we generally assume that plants lack pollution controls, resulting in higher estimates. The exception is that for coal units with missing SO_2 pollution control information, we used the efficiency of the most prevalent type of pollution control in the specific region if the percentage of units with missing pollution control information was less than 50%. This restriction was applied in order to avoid introducing potential bias into the emissions estimates. We assess uncertainty using the standard deviations of uncontrolled emissions factors across boiler types and assign confidence scores based on data availability. We validated against Europe's Beyond Fossil Fuels data, US CAMPD-reported SO_2 and NO_X and EGRID for PM2.5, and Australia's National Pollutant Inventory, showing a moderate to strong correlation between reported and estimated emissions. Challenges include missing data, uncertainties in reported emissions factors and day-to-day plant operation, while future work could incorporate more specific fuel and boiler type information, regional variations in emissions factors, and extending validation to additional regions. For more information on non-GHG estimates, please contact us.

2.2.2 Grid Marginal Operating Emissions Intensity

New this year, we have included average annual Grid Marginal Operating Emissions Intensity (tCO_2 per MWh) for 2022 to 2024 year to date, for each available region of the world (stored in the "other_7" column of asset-level data) derived from WattTime's historical data API. This data represents the change in emissions caused by a change in electricity use (or generation) at a particular time and place. For more details, please see WattTime's website.

2.2.3 Monthly Data

Monthly power plant electricity generation and emissions (CO_2, SO_2, NO_X, and PM2.5) for January 1, 2015, through July 31, 2024, was produced for the first time this year. As opposed to simply assuming equal generation in each month (dividing annual values by 12), we scaled our monthly electricity generation (and thus emissions) estimates based on estimated monthly electricity demand. Monthly demand was estimated for each country, year, and month as the average between the following two approaches:

1. Synthetic Electricity Demand – adapted by WattTime based on Mattsson et al [5].
2. Estimated Load Profile – created by TransitionZero based on [6] and [7].

While this monthly data offers increased recency and understanding of seasonal patterns, it does come with the cost of higher uncertainty, driven both by fewer remote sensing observations (for the RS+ML method) and uncertainty in the estimate of monthly generation.

2.2.4 Mechanical Draft Cooling

New this year, we added an additional proxy signal, mechanical draft cooling towers, to monitor power plants from satellite imagery. This is the first proxy signal that targets gas plants in addition to coal, given that our existing two proxy signals are largely coal-focused. Mechanical draft cooling towers, like natural draft cooling towers, are evaporative cooling structures used to dissipate heat resulting from the electricity generation process and release steam plumes into the atmosphere (Figure 4). Unlike natural draft cooling towers, mechanical draft is more discreet, simply appearing as large fans on the rooftops of industrial buildings, and the water vapor plume they release is often far fainter and harder to see. Because cooler and wetter ambient conditions produce larger, more visible water vapor plumes, we apply a restricted temperature and humidity filter (stricter than that applied to natural draft). This new signal poses a greater modeling challenge, with our mechanical draft models performing slightly worse than our existing models, yet still well above the country-/fuel-/prime-mover-specific averages model.

Figure 4: Examples of mechanical draft cooling structures (source: Google Maps).

3. Results

3.1 Validation

Climate TRACE publishes asset-level (facility-scale) data from January 1, 2019, through July 31, 2024, so electricity generation and emissions were validated over the same period of time, data availability permitting (with more recent years sometimes lacking). For asset-level validation, Climate TRACE estimates were compared to facility-level hourly or daily reported electricity generation data (summed to monthly/annual totals) from Europe (ENTSO-E), Türkiye (EPİAŞ) the USA (CAMPD), India (NPP), Taiwan (Taipower, 台灣電力公司) and Australia (NEM), displayed in Table 2. In Table 3, we compared against reported emissions CO_2, SO_2, NO_x, and PM 2.5. For CO_2, we compared annual Climate TRACE estimates against reported CO_2 emissions from Europe (EU ETS), USA (CAMPD; also used for monthly comparisons), India (CEA), and Australia (NGER). For non-GHG air pollutants, we compared Climate TRACE estimates against annual reported SO_2 and NO_x data from Australia (National Pollutant Inventory), the USA (CAMPD), and Europe (Beyond Fossil Fuels) as well as $PM_{2.5}$ reported data from Australia (National Pollutant Inventory) and the USA (EGRID). A Pearson correlation and root mean squared error was used to compare across all plants successfully matched against the reported data sources mentioned above. Thus, the majority of estimates in Tables 1 and 2 employ the baseline country, fuel, prime mover specific averages model alone, while the more sophisticated (yet limited-applicability) RS+ML model is used for roughly 30%. Nonetheless, all comparisons show a positive correlation with statistically significant p-values (alpha=0.01). Monthly estimates, while more challenging, also do not show any major loss in performance.

Table 2: Overall asset-level validation against reported electricity data, overall across both model types

	ELECTRICITY GENERATION		**CAPACITY FACTOR (generation / capacity)**	
	annual n=8632	monthly n=100996	annual n=8632	monthly n=100996
Pearson Correlation (R)	0.89 (p<0.01)	0.87 (p<0.01)	0.58 (p<0.01)	0.53 (p<0.01)
Root Mean Squared Error (RMSE)	1.7 TWh	166 GWh	0.22	0.25
Value Range	0 - 37.2 TWh	0 - 3385 GWh	0 - 1	0 - 1

Table 3: Overall asset-level validation against reported emissions data, overall across both model types

	CO_2		**SO_2**	**NO_x**	**$PM_{2.5}$**
	annual n=7163	monthly n=50756	annual n=638	annual n=638	annual n=373
Pearson Correlation (R)	0.88 ($p<0.01$)	0.87 ($p<0.01$)	0.8 ($p<0.01$)	0.67 ($p<0.01$)	0.47 ($p<0.01$)
Root Mean Squared Error (RMSE)	1.25 $MtCO_2$	0.1 $MtCO_2$	5,533 tSO_2	7,072 tNO_x	448 $tPM_{2.5}$
Value Range	0 - 35 $MtCO_2$	0 - 2.1 $MtCO_2$	0 - 48742 tSO_2	0 - 35522 tNO_x	0 - 1961 $tPM_{2.5}$

Table 4 demonstrates the improvement offered by remote sensing with machine learning (RS+ML) over the country, fuel, prime mover specific averages approach, evaluated on the same set of plants (i.e. the set of plants which exhibit the water vapor proxy signals and for which we have reported data). Monthly estimates, while more challenging, also do not show any substantial loss in performance. Though fewer in raw plant count, the plants to which we apply this more accurate RS+ML method amount to nearly half (43%) of non-biomass electricity CO_2 emissions. Thus, for these largest plants, the ability to apply remote sensing and machine learning adds an important lift.

Table 4: Asset-level validation against reported data, comparing averages-based vs. RS+ML over the same set of plants

	CAPACITY FACTOR		**ELECTRICITY GENERATION**		**CO_2**	
MODEL	annual n=2410	monthly n=29938	annual n=2410	monthly n=29938	annual n=2017	Monthly n=14109
Method A: Baseline country, fuel, prime mover specific averages	R: 0.49 ($p<0.01$) RMSE: 0.24	R: 0.44 ($p<0.01$) RMSE: 0.26	R: 0.80 ($p<0.01$) RMSE: 1.7 TWh	R: 0.75 ($p<0.01$) RMSE: 229 GWh	R:0.81 ($p<0.01$) RMSE: 2.13 $MtCO_2$	R: 0.80 ($p<0.01$) RMSE: 0.18 $MtCO_2$
Method I: Ensembled with remote sensing & machine learning (RS+ML)	R: 0.71 ($p<0.01$) RMSE: 0.17	R: 0.61 ($p<0.01$) RMSE: 0.22	R: 0.89 ($p<0.01$) RMSE: 1.8 TWh	R: 0.85 ($p<0.01$) RMSE: 184 GWh	R: 0.91 ($p<0.01$) RMSE: 1.15 $MtCO_2$	R: 0.87 ($p<0.01$) RMSE: 0.15 $MtCO_2$

Since asset-level reporting mechanisms do not include power plants below a certain capacity threshold and the majority of countries do not report on individual power plants, in order to validate Climate TRACE country-total estimates, we compare against EDGAR 1.A.1.a Main Activity Electricity and Heat Production emissions. This is available for 196 countries, January 1, 2015, through December 31, 2022; see table 5 below for the comparison.

Table 5: Country-level (aggregated national total, per time period) RMSE

	Country-level CO_2 Climate TRACE vs. EDGAR 1.A.1.a Main Activity Electricity and Heat Production
Annual n=1560	R: 0.998 ($p<0.01$) RMSE: 62.67 $MtCO_2$ (range: 0 - 5896.78 $MtCO_2$)
Monthly n=18685	R: 0.992 ($p<0.01$) RMSE: 6.38 $MtCO_2$ (range: 0 - 623.86 $MtCO_2$)

4. Conclusion

Climate TRACE provides the data that makes it possible to answer questions that drive impactful, precise climate action. Whether the priority is to reduce greenhouse gas emissions or air pollution, or both, this question can be directly answered by plots above and the data continually updated on climatetrace.org. Coupled with Climate TRACE data on other sectors - manufacturing, shipping and aviation, etc. - a life cycle analysis is possible along the entire supply chain, of which power generation is just one piece. This enables businesses, investors, governments, and even consumers to quantify the greenhouse gas and/or air pollution footprint(s) of products and compare the trade-offs between one supply chain vs. another. With this information and the lines of thinking enabled by it, we can make climate action faster and easier by uncovering how to slash the maximum total emissions on a global scale with the fastest and easiest actions possible.

Future work includes improving electricity generation, and thus emissions, estimates by incorporating additional data sources and alternate modeling approaches to improve the baseline model as well as expand the set of plants to which we can apply remote sensing and machine learning via additional proxy signals (e.g. heat signatures via thermal satellite imagery). Non-GHG air pollutant emissions estimates, while new and carrying high uncertainties given technology differences across the world, offer an additional lens to explore the intersections between industry, climate change, and environmental justice at high spatial and temporal resolution. We look forward to collaborating to drive increased accuracy and actionable insights.

5. Supplementary materials metadata

Country-level emissions estimates for electricity generation are available for download at ClimateTRACE.org, and the following table summarizes this data.

Table S1 General dataset information for "country-climate-trace_electricity-generation_091924.csv"**.**

General Description	Definition
Sector definition	*Electricity Generation*
UNFCCC sector equivalent	*1.A.1.a.i Electricity Generation, 1.A.1.a.ii Combined Heat and Power Generation*
Temporal Coverage	*January 1, 2015 – July 31, 2024 (inclusive)*
Temporal Resolution	*Monthly*
Data format(s)	*CSV*
Coordinate Reference System	*EPSG:4326, decimal degrees*
Number of sources available for download and percent of global emissions	*9,004 unique sources across January 1, 2019 - July 31, 2024; totaling 12.9 GtCO_2 in 2023, representing 96% of all fossil/waste fuel power plant emissions in 2023.*
Total global emissions for 2023	*13.4 Gt CO_2*
Ownership	*Ownership data was obtained from Global Energy Monitor (GEM) and the U.S. Energy Information Administration (EIA) for the U.S.*
What emission factors were used?	*Carbon intensity values for combinations of energy source and prime mover technology were modeled from USA EPA, JRC data and IEA data. Non-GHG air pollutant emissions factors were sourced from the EMEP/EEA air pollutant emission inventory guidebook 2023.*
What is the difference between a "NULL / none / nan" versus "0" data field?	*"0" values are for true non-existent emissions. If we know that the sector has emissions for that specific gas, but the gas was not modeled, this is represented by "NULL/none/nan"*
total_CO2e_100yrGWP and total_CO2e_20yrGWP conversions	*Climate TRACE uses IPCC AR6 CO_2e global warming potentials (GWPs). CO_2e conversion guidelines are here: [https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_FullReport_small.pdf.](https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_FullReport_small.pdf)*
Do the estimates include emissions from biomass fuels?	*No, due to concerns of overcounting since biomass is accounted for by the Climate TRACE forestry and land use sector. Electricity CO_2 emissions reported at the country level are from fossil/waste fuel electricity generation only. Emissions resulting from biomass are not included in the country totals but are available at the source-level for 2019-2024 (via the "other_4" column) and upon request at the country level for 2015-2024. For plants that burn both fossil/waste and biomass, emissions are appropriately broken out based on the power plant's capacity of each fuel type.*

Source-level emissions estimates for electricity generation are available for download at ClimateTRACE.org, and the following two tables summarize this data.

Table S2: Asset level metadata description for asset-climate-trace_electricity-generation_091924.csv. Note, fields marked with an asterisk (*) indicate fossil/waste fuel CO_2 emissions, and GWP are broken out separately from biomass fuel to avoid double-counting emissions from the Climate TRACE forestry sector. Capacity, capacity factor, generation, SO_2 emissions, NO_x emissions, PM 2.5 emissions do include biomass, however (they include all combustible fuels – fossil, waste, biomass).

Data attribute	Definition
sector	*Electricity Generation*
asset definition	*Each asset is an individual power plant.*
start_date	*[string] Gregorian time period start in YYYY-MM-DD format. For example, for the month of January 2024,* start_date *is* "2024-01-01".
end_date	*[string] Gregorian time period end as a string in YYYY-MM-DD format. For example, for the month of January 2024,* end_date *is* "2024-01-31".
asset_identifier	*[string] Unique identifier for assets (power plants). For example, use the* asset_identifier *column in* asset *to match to* ownership*'s* asset_id *column.*
model_number	*[string] Date results were produced*
asset_name	*[string] Power plant name*
iso3_country	*[string] ISO 3166 Alpha-3 country code*
location	*[shapely.Point] Coordinates x,y (longitude, latitude) for the power plant's location*
type	*[string] power plant fuel type. note: multiple fuel types are listed in order of prevalence for power plants that burn more than one type of fuel.*
capacity description	*[integer] power plant fossil/waste/biomass fuel capacity in MW sourced from harmonized power plant database described in Table 1, representing the instantaneous maximum electricity the power plant is capable of generating under normal conditions.*
capacity_units	*megawatts (MW)*
capacity_factor description	*[float] Climate TRACE estimated power plant average fossil/waste/biomass fuel utilization rate for a given plant and time period calculated as generation / capacity * hours in time period*
capacity_factor_units	*Proportion of power plant utilization multiplied by number of hours in time period, (MWh per MW capacity) on a scale from 0 to 1*
activity_description	*[integer] Climate TRACE estimated power plant total fossil/waste/biomass fuel generation in MWh for a given power plant and time period.*
activity_units	*megawatt-hour (MWh)*
CO2_emissions_factor	*[float] Carbon intensity values for combinations of energy source and prime mover technology were modeled from USA EPA, JRC data and IEA data.*
CH4_emissions_factor	*NaN (CH_4 not tracked for electricity generation)*
N2O_emissions_factor	*NaN (N_2O not tracked for electricity generation)*
SO2_emissions_factor	*[float] We estimate emission factors for SO_2, NO_x, and PM 2.5 at the power plant unit-level based on fuel type, boiler type, and pollution control technology, according to*

Data attribute	Definition
NOX_emissions_factor	*guidance provided in the 2023 EMEP/EEA air pollutant emission inventory guidebook, 2023 EIA Electric Power Annual, and the US EPA's AP-42.*
PM2_5_emissions_factor	
CO2_emissions	*[integer] Climate TRACE estimated power plant total CO_2 emissions from fossil/waste fuels* for a given time period.*
CH4_emissions	*NaN (CH_4 not tracked for electricity generation)*
N2O_emissions	*NaN (N_2O not tracked for electricity generation)*
SO2_emissions	*[integer] Climate TRACE estimated power plant total SO_2 emissions from fossil/waste/biomass fuels for a given time period.*
NOX_emissions	*[integer] Climate TRACE estimated power plant total NO_x emissions from fossil/waste/biomass fuels for a given time period.*
PM2_5_emissions	*[integer] Climate TRACE estimated power plant total PM 2.5 emissions from fossil/waste/biomass fuels for given time period.*
total_CO2e_100yrGWP	*[integer] Climate TRACE estimated power plant total fossil/waste fuel* CO_2 100-year global warming potential (GWP), rounded to the nearest integer for a given time period. Since CO_2* is *the only gas currently tracked, this is equal to* CO2_emissions.
total_CO2e_20yrGWP	*[integer] Climate TRACE estimated power plant total fossil/waste fuel* CO_2 20-year global warming potential (GWP), rounded to the nearest integer for a given time period. Since CO_2 is the only gas currently tracked, this is equal to* CO2_emissions.
other1_description	*[string] Describes where the remote sensing and machine learning proxy signal (RS+ML) method was applied by denoting whether the Climate TRACE estimate comes from the country-/fuel-/prime-mover-specific averaging alone (*a*) by averaging this method with the RS+ML method (*i*).*
other1_units	a = *estimated by country-/fuel-/prime-mover-specific averages alone;* i = *estimated by averaging the RS+ML method and country-/fuel-/prime-mover-specific averages*
other2_description	*[integer] power plant biomass-only capacity in MW sourced from harmonized power plant database described in Table 1, representing the instantaneous maximum electricity from biomass the power plant is capable of under normal conditions.*
other2_units	*megawatts (MW)*
other3_description	*[integer] Climate TRACE estimated power plant total biomass-only generation in MWh for a given power plant and time period.*
other3_units	*megawatt-hour (MWh)*
other4_description	*[integer] Climate TRACE estimated power plant total CO_2 emissions from biomass-only* in metric tons (tonnes) for a given time period.*
other4_units	*metric tons (tonnes) abbreviated tCO_2*
other5_description	*[float] capacity factor*
other5_units	*Unitless proportion of power plant utilization, MW per MW, on a scale from 0 to 1*
other6	*N/A (not used)*

Data attribute	Definition
other7_description	*[float] grid marginal operating emissions intensity*
other7_units	*tCO_2 per MWh*

Table S3 describes source level metadata description confidence and uncertainty for "confidence-climate-trace_electricity-generation_091924.csv" and "uncertainty-climate-trace_electricity-generation_091924.csv"**.**

Confidence is defined on a 5-point scale from very low to very high:

- Very low (1): Purely assumption-driven or engineering estimates with details not verified by anything;
- Low (2): Purely assumption-driven or engineering estimates with few details calibrated;
- Medium (3): Estimated from a machine learning model applied outside the training data set on data with fairly similar physical characteristics;
- High (4): Estimated from a machine learning model applied to the training data set or estimated;
- Very High (5): Estimated by multiple independent sources with agreement

Table S3: Source level metadata description confidence and uncertainty for "confidence-climate-trace_electricity-generation_091924.csv" and "uncertainty-climate-trace_electricity-generation_091924.csv"**.**

Data attribute	Confidence Definition	Uncertainty Definition
type	Confidence in fuel type reported. Set to 4 = High for all plants because data is sourced from the harmonized power plant inventory as described in Table 1	N/A
capacity_description	Confidence in capacity reported. Set to 4 = High for all plants because data is sourced from the harmonized power plant inventory as described in Table 1	Uncertainty in capacity reported. Estimated at 3% for all plants
capacity_units	megawatts (MW)	megawatts (MW)
capacity_factor_description	Capacity factor is electricity generation divided by power plant capacity. Of the ML proxy signals (NDT, FGD, and MDT), NDT performs the best. Therefore, capacity factor confidence was set to the following: 5 = Very High for reported data with up to 20% missing reporting events; 4 = High for plant-months with imputed reported data OR plants which used the NDT model and country-/fuel-specific averages; 3 = Medium for plants that did not use the NDT model but did use the FGD model and country-/fuel-specific averages;	For reported data: uncertainty is the RMSE between CAMPD reported data vs. EIA-923, then applied to all regions; For ML models: reported as the RMSE between predicted and reported capacity factor calculated over regions with reported generation data (India, Türkiye, US, Europe and Australia). This was further broken down by whether the plant was ML modellable with the NDT, FGD, or MDT models. For the country-/fuel-specific averages approach, this was broken down by plant capacity.

Data attribute	Confidence Definition	Uncertainty Definition
	2 = Low for plants that had access to only the country-/fuel-specific averages method and/or MDT.	
capacity_factor_units	on a scale from 0 to 1 (proportion)	on a scale from 0 to 1 (proportion)
activity_description	uncertainty was calculated as the square root of the sum of the squared fractional uncertainties for capacity and capacity factor, assuming independent random errors. Thus, for electricity generation $g = cf$ for capacity c and capacity factor f with independent random errors ϵ_c and ϵ_f respectively, the electricity generation uncertainty $\epsilon_g = g\sqrt{(\epsilon_c/c)^2 + (\epsilon_f/f)^2}$.	average between capacity and capacity factor confidences, resulting in some set to 4="high" (those modeled with ML+satellites and with some amount of NDT) and the remaining majority set to 3="medium."
activity_units	megawatts (MW)	megawatts (MW)
CO2_emissions_factor	2 = "low" for all plants	Calculated as 25% of the CO_2 emissions factor, across all plants.
CH4_emissions_factor	N/A	N/A
N2O_emissions_factor	N/A	N/A
SO2_emissions_factor	We assign the confidence score based on the data availability and our fallback approach. If we are able to match the data attributes required for both, uncontrolled EF and pollution control efficiency, without applying our fallback strategy described above, we assign a confidence score of 3 ("medium"). If we have to rely on other fuel types to obtain control efficiencies or average across all boiler types of the unit's fuel to obtain the uncontrolled EF, we assign a confidence score of 2 ("low"). In all other cases, we assign a confidence score of 1 ("very low").	To obtain the uncertainty value, we calculate the standard deviation across a boiler type's uncontrolled EFs associated with a fuel that belongs to the unit's fuel type class. For example, we calculate the standard deviation across the EFs of all coals, e.g., lignite, subbituminous, etc. per boiler type. If we are able to match the boiler type from our database to a boiler type available in the source data, we report this standard deviation as the uncertainty. If that match is not available, we calculate the standard deviation across all EFs associated with the fuel type class and report this standard deviation as the uncertainty.
NOX_emissions_factor		
PM2_5_emissions_factor		
CO2_emissions	taken as the average between the confidence scores for capacity, capacity factor, and CO_2 emissions factor, which	For capacity c, capacity factor f, and CO_2 emissions factor e, with

Data attribute	Confidence Definition	Uncertainty Definition
	ended up being 2 = “medium” for all plants	uncertainties ϵ_c, ϵ_f, ϵ_e respectively, CO_2 emissions uncertainty ϵ_m for CO_2 emissions m are propagated as the square root of the sum of the squared fractional uncertainties of each of these three factors: $\epsilon_m = m\sqrt{(\epsilon_c/c)^2 + (\epsilon_f/f)^2 + (\epsilon_e/e)^2}$
CH4_emissions	N/A	N/A
N2O_emissions	N/A	N/A
SO2_emissions	taken as the average between the confidence scores for capacity, capacity factor, and CO_2 emissions factor	For capacity c, capacity factor f, and emissions factor e, with uncertainties ϵ_c, ϵ_f, ϵ_e respectively, emissions uncertainty ϵ_m for emissions m are propagated as the square root of the sum of the squared fractional uncertainties of each of these three factors: $\epsilon_m = m\sqrt{(\epsilon_c/c)^2 + (\epsilon_f/f)^2 + (\epsilon_e/e)^2}$
NOX_emissions		
PM2_5_emissions		
total_CO2e_100yrGWP	same as CO_2 emissions, since the GWP factor for CO_2 is 1	same as CO_2 emissions, since the GWP factor for CO_2 is 1
total_CO2e_20yrGWP	same as CO_2 emissions, since the GWP factor for CO_2 is 1	same as CO_2 emissions, since the GWP factor for CO_2 is 1

Data citation format:
Freeman, J., Rouzbeh Kargar, A., Couture, H., Jeyaratnam, J., Lewis, J., Alvara, M., Koenig, H., Nakano, T., Davitt, A., Lewis, C., and McCormick, G. (2025). *Power sector- Emissions from Electricity Generation.* WattTime, USA Transition Zero, UK, Pixel Scientia Labs, USA and Global Energy Monitor, USA, Climate TRACE Emissions Inventory. https://climatetrace.org [Accessed date]

Geographic boundaries and names (iso3_country data attribute): The depiction and use of boundaries, geographic names and related data shown on maps and included in lists, tables, documents, and databases on Climate TRACE are generated from the Global Administrative

Areas (GADM) project (Version 4.1 released on 16 July 2022) along with their corresponding ISO3 codes, and with the following adaptations:

- HKG (China, Hong Kong Special Administrative Region) and MAC (China, Macao Special Administrative Region) are reported at GADM level 0 (country/national);
- Kosovo has been assigned the ISO3 code 'XKX';
- XCA (Caspian Sea) has been removed from GADM level 0 and the area assigned to countries based on the extent of their territorial waters;
- XAD (Akrotiri and Dhekelia), XCL (Clipperton Island), XPI (Paracel Islands) and XSP (Spratly Islands) are not included in the Climate TRACE dataset;
- ZNC name changed to 'Turkish Republic of Northern Cyprus' at GADM level 0;
- The borders between India, Pakistan and China have been assigned to these countries based on GADM codes Z01 to Z09.

The above usage is not warranted to be error free and does not imply the expression of any opinion whatsoever on the part of Climate TRACE Coalition and its partners concerning the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its borders.

Disclaimer: The emissions provided for this sector are our current best estimates of emissions, and we are committed to continually increasing the accuracy of the models on all levels. Please review our terms of use and the sector-specific methodology documentation before using the data. If you identify an error or would like to participate in our data validation process, please contact us.

6. Errata

1. We have reason to suspect biomass PM 2.5 emissions for some units were underestimated in our current release. This will be corrected in future monthly updates.
2. We currently report an emissions factor uncertainty of 0 in cases where the underlying data has only a single emissions factor. This should not be interpreted as no uncertainty in our estimate and will be replaced with a more informed value in the future.

7. References

1. Ge, M.; Friedrich, J. 4 Charts Explain Greenhouse Gas Emissions by Countries and Sectors; World Resources Institute: Washington, DC, USA, 2020; Available online: https://www.wri.org/blog/2020/02/greenhouse-gas-emissions-by-country-sector (accessed on 6 June 2023).

2. Climate Watch. Climate Watch Historical GHG Emissions; World Resources Institute: Washington, DC, USA, 2022; Available online: https://www.climatewatchdata.org/ghg-emissions (accessed on 1 July 2023).

3. IEA (International Energy Agency). Greenhouse Gas Emissions from Energy Data Explorer; IEA: Paris, France, 2023; Available online: https://www.iea.org/data-and-statistics/data-tools/greenhouse-gas-emissions-from-energy-data-explore (accessed on 1 July 2023).

4. Karn Vohra, Alina Vodonos, Joel Schwartz, Eloise A. Marais, Melissa P. Sulprizio, Loretta J. Mickley, "Global mortality from outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem," Environmental Research, Volume 195, 2021, 110754, ISSN 0013-9351, https://doi.org/10.1016/j.envres.2021.110754. (https://www.sciencedirect.com/science/article/pii/S0013935121000487)

5. Niclas Mattsson, Vilhelm Verendel, Fredrik Hedenus, Lina Reichenberg, "An autopilot for energy models – Automatic generation of renewable supply curves, hourly capacity factors and hourly synthetic electricity demand for arbitrary world regions," Energy Strategy Reviews, Volume 33, 2021, 100606, ISSN 2211-467X, https://doi.org/10.1016/j.esr.2020.100606. (https://www.sciencedirect.com/science/article/pii/S2211467X20301590)

6. Hampp, Johannes. "Euronion/Synde." *GitHub*, 23 Aug. 2021, github.com/euronion/synde.

7. Maarten Brinkerink, Brian Ó. Gallachóir, Paul Deane, "Building and Calibrating a Country-Level Detailed Global Electricity Model Based on Public Data," Energy Strategy Reviews, Volume 33, 2021, 100592, ISSN 2211-467X, https://doi.org/10.1016/j.esr.2020.100592. (https://www.sciencedirect.com/science/article/pii/S2211467X20301450)

8. Hobbs, M.; Kargar, A.R.; Couture, H.; Freeman, J.; Söldner-Rembold, I.; Ferreira, A.; Jeyaratnam, J.; O'Connor, J.; Lewis, J.; Koenig, H.; et al. Inferring carbon dioxide emissions from power plants using satellite imagery and machine learning. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 16–21 July 2023.