ENVIRONMENTAL RESEARCH LETTERS

LETTER • **OPEN ACCESS**

Assessment of the Climate Trace global powerplant CO_2 emissions

ENVIRONMENTAL RESEARCH
LETTERS

CrossMark

OPEN ACCESS

RECEIVED
13 July 2024

REVISED
16 September 2024

ACCEPTED FOR PUBLICATION
4 October 2024

PUBLISHED
17 October 2024

LETTER

Assessment of the Climate Trace global powerplant CO_2 emissions

Kevin R Gurney[1,*], **Bilal Aslam**[1], **Pawlok Dass**[1], **Lech Gawuc**[1], **Toby Hocking**[1,2], **Jarrett J Barber**[1] **and Anna Kato**[1]

[1] School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, United States of America
[2] Département d'informatique, Université de Sherbrooke, 2500 Boulevard de l'Université, Sherbrooke QC J1K-2R1, Canada
[*] Author to whom any correspondence should be addressed.

E-mail: kevin.gurney@nau.edu

Keywords: CO2 emissions, machine learning, powerplants

Supplementary material for this article is available online

Abstract

Accurate estimation of planetary greenhouse gas (GHG) emissions at the scale of individual emitting activities is a critical need for both scientific and policy applications. Powerplants represent the single largest and most concentrated form of global GHG emissions. Climate Trace, co-founded and promoted by former U.S. Vice President Al Gore, is a new effort using, in part, artificial intelligence (AI) approaches to estimate asset-scale GHG emissions. Climate Trace recently released a database of global powerplant CO_2 emissions at the facility-scale that uses both AI and non-AI estimation approaches. However, no independent peer-reviewed assessment has been made of this important global emissions database. Here, we compare the Climate Trace powerplant CO_2 emissions to an atmospherically calibrated, multi-constraint estimate of powerplant CO_2 emissions in the United States. The 3.7% (65) of compared facilities that used an AI-based approach show a mean relative difference (MRD) of −1.1% (SD: 46.4%) in the year 2019. The 96.3% (1726) of the facilities that used a non-AI-based approach show a MRD of −50.0% (SD: 117.7%). Of the non-AI estimated facilities, 151 (8.7%) facilities agree to within ±20%. The large differences between Climate Trace and Vulcan-power emission estimates for these facilities is primarily caused by Climate Trace' use of a national-mean power plant capacity factor (CF) which is a poor representation of the reported power plant CFs of individual US facilities and leads to very large errors at those same 1726 facilities.

1. Introduction

Accurate quantification of anthropogenic greenhouse gas (GHG) emissions is an essential ingredient of both climate change science and climate change policy. Numerous research-based anthropogenic GHG emissions estimates have been produced to meet the needs of carbon cycle research, climate change research, and GHG emissions mitigation policy [1]. The various efforts reflect different methods, data sources, space/-time resolution and geographic domains. There is increasing interest and need for emissions quantification at the scale of the emitting asset [2]. This is important for identifying specific emitters and prioritizing mitigation options. The business community, cities, and citizen groups all need asset-scale emissions characterization. While there is recent research and practitioner work at the city, province and national scales, asset-scale quantification at the global scale is particularly challenging, especially in portions of the planet where data transparency and/or data collection is limited.

Climate Trace is an effort to better quantify and track GHG emissions globally at the asset scale. Climate Trace represents a coalition of organizations and investigators begun in 2019 and is now generating and openly distributing datasets on GHG emissions along with methodology documents from the Climate Trace website [3]. The project has gained considerable media attention owing, in part, to its co-founder and chief spokesperson, former vice-president Al Gore. For example, the project has been presented prominently at both the conference of the parties 27 and 28 to the United Nations Framework

Convention on Climate Change in 2022 and 2023, respectively [4, 5]. The Project claims 'unprecedented granularity that pinpoints nearly every major source of GHG emissions around the world and provides independently produced estimates of how much each emits.' [6]. Recent reporting indicates that large companies such as Boeing, Tesla, and General Motors are signing up to use the Climate Trace data [7]. Finally, the effort utilizes new techniques and approaches to solving the global GHG emissions problem, for example, 'by training artificial intelligence (AI) algorithms to fuse data across multiple wavelengths and from more than 300 satellites; 11 100 air-, land-, and sea-based sensors; and other data streams to identify all of the largest point sources and track them over time.' [8]. The granularity and pioneering approaches make it an important contribution to the collection of emissions data products being pursued in the research and practitioner communities [1].

Given its stature, importance and the recent data availability, it is useful to compare it to other efforts that have quantified emissions at commensurate spatial resolution using independent approaches and data sources. A good example of this is the Vulcan Project which, while limited to the U.S. geographic domain, quantifies emissions at the spatial resolution of individual points, lines, and polygons using U.S.-specific data [9].

Powerplants are a particularly helpful place to start with a comparison. For one, they are the largest single sector-based fossil fuel CO_2 emitter in the United States. They are monitored at the individual unit-scale by multiple federal agencies using independent measurements and have been carefully and comprehensively quantified by the Vulcan Project under the Vulcan-power data product release [10]. Comparisons also offer useful insights on methodological improvements, data gaps, and uncertainty.

This paper presents a facility-by-facility comparison in the United States between the Vulcan-power version 1.1 powerplant CO_2 emissions and the Climate Trace powerplant CO_2 emissions for the years 2019, 2020, and 2021. Differences are quantified at aggregate and individual facility scales and we attempt to identify the underlying explanations for the discrepancies found.

2. Methods

2.1. Climate Trace

The Climate Trace powerplant emissions data for the years 2019, 2020, and 2021 was downloaded from the Climate Trace website [11]. The Climate Trace power plant CO_2 estimation uses two distinct approaches. The first uses a combination of machine learning (ML) and remotely sensed water vapor plumes to generate time-dependent capacity factors (CFs) at the facility scale. This approach is used to estimate 12.5% of the total 8333 global facilities estimated across the 2019–2022 time period [12].

The second, and majority of individual facility estimates (87.5%), relies on a simple calculation using national mean powerplant properties. The fuel type-specific annual/national mean CFs (MCFs) (percentage or share of maximum generation at a given time) are multiplied by individual facility nameplate capacity (the maximum generation state) to estimate facility-specific generation. This, in turn, is multiplied by a CO_2 emission rate (e.g. tonnes carbon dioxide per megawatt-hour) to estimate an annual emission amount (tonnes of carbon dioxide per year).

2.2. Vulcan-power

The Vulcan-power v1.1 dataset in 2019 is comprised of 14 258 grid-connected power generation facilities in the United States for which annual fossil fuel CO_2 and total CO_2 emissions are estimated [10]. It represents one sector of the larger Vulcan Project which estimates fossil fuel CO_2 emissions for all sectors in gridded form [9]. The pointwise Vulcan-power is primarily built from two data sources. The first is the Environmental Protection Agency's Clean Air Markets Division (CAMD) data [13]. The CAMD data are collected under the Acid Rain Program, which was instituted in 1990 under Title IV of the Clean Air Act [14–16]. Though the CAMD data set does not include all power plants in the United States, it accounts for a very large proportion of the CO_2 emissions. The CAMD data used in Vulcan are reported as hourly CO_2 emissions monitored from an emitting stack or through a calculation, based on records of fuel consumption [17]. The annual reporting is also used for additional information related to the facility.

The second primary data source used is the Department of Energy's Energy Information Administration (EIA) reporting data [18]. The EIA data set is derived from the EIA reporting form 923, which reports monthly data on receipts and cost of fossil fuel, fuel stocks, generation, consumption of fuel for generation, and environmental data at each power plant [19]. Fuel consumption is reported as a heat input value (e.g. British thermal units). CO_2 emission factors specific to fuel (tonnes of carbon per British thermal unit) are then utilized to calculate the quantity of CO_2 emitted. In order to maintain consistency with the data source, the CO_2 emission factors used by the EIA are adopted to estimate the $FFCO_2$ emissions from these facilities [20].

Some manual corrections are performed to the geocoordinates of both the CAMD and EIA electricity production data, as a result of searching in Google Earth or via alternative online information resources (e.g. utility websites).

A hierarchy was employed given that there was overlap between the two data sets. This was performed at the unit level given that a single facility might have individual power units reporting to CAMD and another only reporting to the EIA. Where overlap did exist at this scale, preference was made to retain the CAMD data. Further details and rationale can be found in Gurney *et al* [20, 21]. The Vulcan-power dataset can be accessed online [10].

2.3. Facility matching

Since there is no common unique ID linking the power plants in the Vulcan and Climate Trace datasets, a mixture of geographic proximity and facility name string matching was used. Facilities were included if the Euclidian distance between facility pairs was within 50 meters. This produced 2109 pairs of facilities in 2019. Facility name matching was further applied to this set of paired facilities using the 'stringdist' R library. The match threshold was set to 0.25, meaning that 25% of the name strings were sequentially identical. This reduced the number of matched pairs to 1791 in 2019. The match threshold was arrived at by visual inspection of the location matched facility set and determining where potential mismatches began. Sensitivity to both the Euclidian distance and name matching thresholds were performed with results available in the supplementary information, tables S1 and S2.

3. Results

A series of statistics are used to compare the two datasets for all three years (table 1). We separate the facilities for which Climate Trace used AI methods versus those that do not, instead relying on a traditional calculation approach (see methods). We refer to these as 'AI' and MCF-based or 'MCF' methods, respectively. The number of matched power plant facility pairs in 2019 was 1726 for the MCF method and 65 for the AI method (in 2020: 1698 and 62; in 2021: 1683 and 60). The matched set of facilities analyzed here comprise 80% of the 2019 U.S. power plants reported in the Climate Trace dataset and 80.7% of their total reported U.S. power plant CO_2 emissions.

In 2019, the total relative CO_2 emissions difference (TRD) between Climate Trace and Vulcan-power is −5.8% and −14.0% for the AI and MCF facilities, respectively, with the Climate Trace total less than the Vulcan-power equivalent across all years. The mean emissions of the two datasets are statistically different for the MCF method ($D = 0.25$; p-value $= 2.2 \times 10^{-16}$) but not statistically different for the AI method ($D = 0.09$; p-value $= 0.95$) based on a Kolmogorov–Smirnov two-sample test.

However, in spite of the closeness of the total and mean emissions, the differences at the individual facility scale are larger. The mean paired difference (MD) between the Climate Trace and Vulcan-power datasets in 2019 is 97 650 tC/yr and 21 508 tC/yr for the AI and MCF methods, respectively. The same general dynamic is seen when the sum of the individual facility differences are considered without sign (absolute). The mean absolute difference (MAD) in 2019 is 421 190 tC/yr and 101 212 tC/yr for the AI and MCF methods, respectively. A MAD that is many times larger than the MD indicates a large amount of sign cancelation in the paired differences.

When the MDs are expressed in relative terms, differences also emerge between Climate Trace and Vulcan-power. The mean relative difference (MRD) across the paired facilities in 2019 is −1.1% (SE: 5.8%) and −50.0% (SE: 2.8%) for the AI and MCF methods, respectively. As with the total summed emissions, the Climate Trace facility estimates are, on average, smaller than the Vulcan-power estimates. The AI method, however, shows a far smaller MRD value which could be attributable to greater symmetry about the 0 difference value or overall smaller differences, regardless of sign. The MRD of the AI facilities is well within the Vulcan power plant $FFCO_2$ estimation uncertainty for all three years (see supplementary information). However, the MCF differences are not.

Figures 1 and 2 provide further insight into the relative paired differences. Figure 1 shows scatterplot diagrams of all the 2019 paired power plant facilities analyzed here in both unscaled (figure 1(a)) and logged (figure 1(b)) form (see supplementary material for years 2020 and 2021). This demonstrates why the MD of the individual facilities using the AI method is larger than those using the MCF method but for which the MRD is, contrastingly, much smaller. The AI method was predominantly used among the largest overall emitting facilities leading to average absolute differences that are larger in relative terms, but smaller when placed in the context of the total facility emission values.

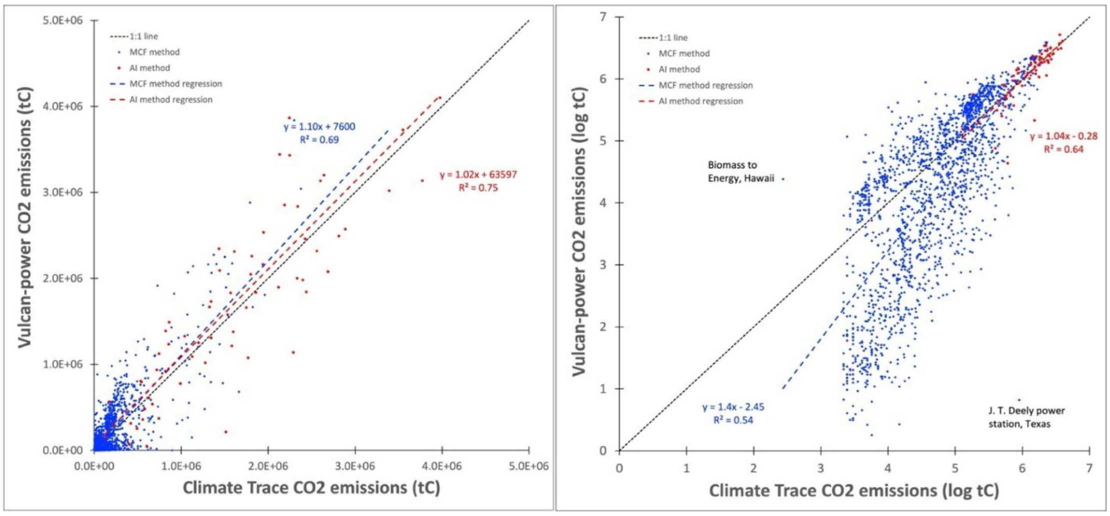
A linear regression of the unscaled paired facilities shows a slope coefficient of 1.10 (SE: 0.018) and an R^2 value of 0.69 for the facilities estimated with the MCF method while the facilities using the AI method show a slope coefficient of 1.02 (SE: 0.074) and an R^2 of 0.75. This further confirms that the Climate Trace estimates are on average smaller than the Vulcan-power estimates and also demonstrates the greater symmetry (equal amount of positive versus negative differences to Vulcan-power facilities—figure 2(b)) in the distribution of the paired differences compared to the MCF-based facilities.

Additionally, figure 1(b) shows a 'low-end' emissions threshold (2183.14 tC) to the Climate Trace power plant CO_2 emissions estimates (with a single exception; the 'biomass to energy facility' in Kauai, Hawaii). Disagreement is greater between the two

Table 1. Comparison statistics for the paired power plant facilities in the Climate Trace (CT) and Vulcan-power (Vulcan) datasets. Threshold values are: proximity—50 meters, string match: 25%[a].

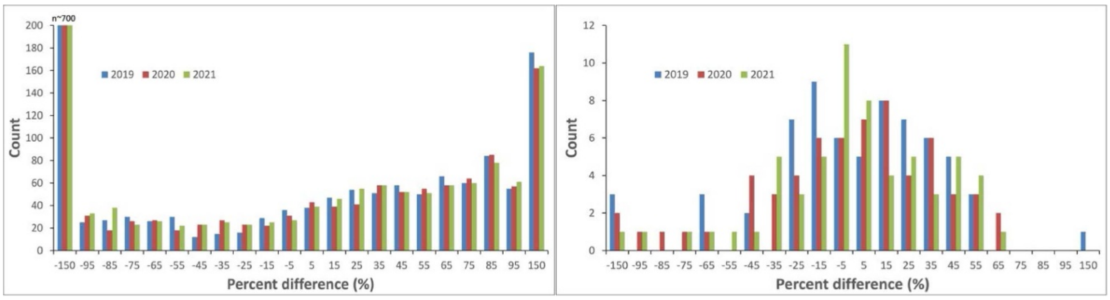
Method	Year	N pairs	Total emit CT (tC)	Total emit Vulcan (tC)	TD (MtC)	TRD (%)	MD (tC)	MAD (tC)	MRD (%)	MRD StDev (%)	SD (MtC)	SAD (MtC)	Pearson correlation
AI	2019	65	106.33	112.68	−6.35	−5.8%	97 650	421 190	−1.1	46.4	6.3	27.4	0.87
	2020	62	89.15	93.30	−4.14	−4.5%	66 821	380 493	−3.1	42.1	4.1	23.6	0.85
	2021	60	98.06	104.36	−6.30	−6.2%	104 999	409 442	−0.2	38.6	6.3	24.6	0.84
MCF	2019	1726	245.66	282.79	−37.12	−14.0%	21 508	101 212	−50.0	117.7	37.1	174.7	0.83
	2020	1698	229.03	263.97	−34.94	−14.2%	20 577	101 776	−51.3	117.0	34.9	172.8	0.79
	2021	1683	235.89	271.41	−35.51	−14.0%	21 101	96 707	−47.0	114.0	35.5	162.8	0.85

[a] differences are calculated as Climate Trace minus Vulcan-power.
TD = total emission difference.
TRD = total emission relative difference.
MD = mean facility difference.
MAD = mean absolute facility difference.
MRD = mean facility relative difference: Calculated as (CT-Vulcan)/avg(CT,Vulcan).
SD = summed facility difference.
SAD = summed absolute facility difference.



a) b)

Figure 1. Comparison of facility-scale Vulcan-power to Climate Trace powerplant carbon dioxide emissions in the year 2019. The Climate Trace facilities using the MCF versus AI method are noted separately (blue and red symbols, respectively). (a) Unscaled paired emissions, 1:1 line (solid black) and regression line/equation (dashed blue) included; (b) same as (a) but for log-scaled paired emission values.



a) b)

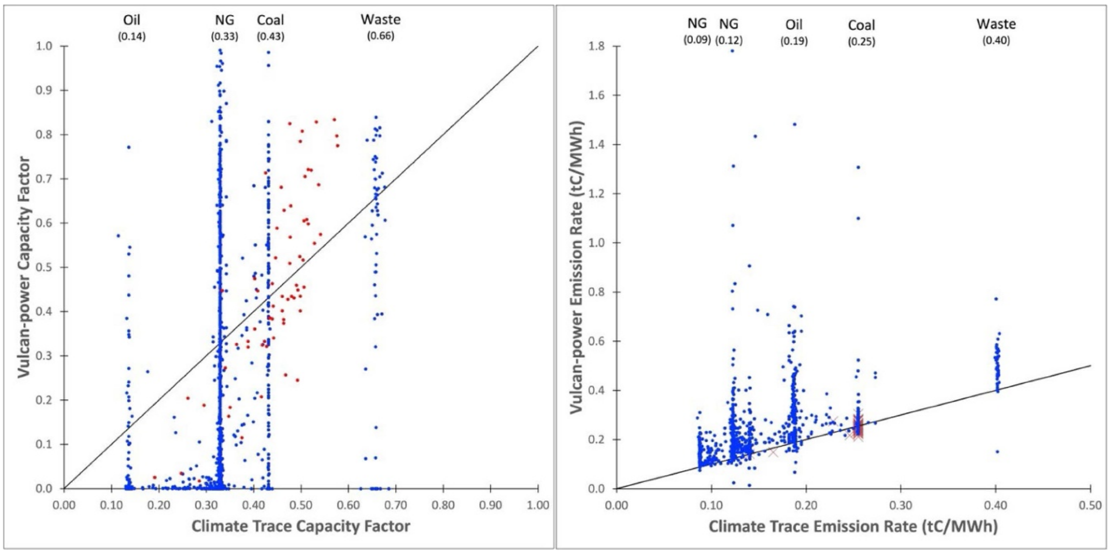
Figure 2. Frequency of the individual facility relative difference between Climate Trace (CT) and Vulcan-power for the years 2019, 2020, and 2021. (a) MCF method; (b) AI method. Calculated as (CT-Vulcan)/avg(CT,Vulcan).

datasets at the smaller emission values, converging somewhat as the emission magnitude increases (see also supplementary information figure S2), consistent with the MRD value for the facilities estimated using the MCF (larger MRD value) versus AI methods (smaller MRD value).

Figure 2 presents the frequency distribution of the individual paired relative difference values across all three years providing further insight into the paired differences. For those facilities using the MCF method (figure 2(a)), there are a large number of paired differences (>700) that are greater than the −95% relative difference frequency bin. Similarly, there are a large number of paired differences that are greater than the +95% relative difference frequency bin, facilities for which Climate Trace emissions exceed those in the Vulcan-power dataset. The much larger negative individual relative difference values account for the large negative MRD (−50.0% for the year 2019). Aside from these two extremes, the majority of differences between the paired facilities in all three years show a positive relative difference with the peak count between +80% and +90% relative difference.

With the Vulcan-power emissions 95% confidence interval of −19.5%/+10.4% in 2019, nearly all of the differences noted using the MCF method are statistically significant at the 2-sigma level [21]). Out of the 1726 facility differences in 2019 using the MCF method, 151 (8.7%) agree to within ±20% of the Vulcan-power emissions estimates, leaving 1575 with relative differences outside these bounds. For the AI method, 28 (43.1%) agree to within ±20% (leaving 37 outside these agreement boundaries).

As previously noted, The MRD across the paired facilities in 2019 is −1.1% and −50.0% for the AI and MCF methods, respectively. The median values are +3.0% and −54.4%, respectively. Though the mean/median of the MRD values for the AI method is much smaller (−1.1%/+3.0%) than the MCF method (−50.0%/−54.4%), the AI method standard deviation remains large at 46.4%.



a) b)

Figure 3. Comparison of key calculation parameters from Vulcan-power and Climate Trace across all paired facilities in this study. (a) Capacity factor; (b) emission rate. The mode value of the Climate Trace parameters for the different primary fuels analyzed here is shown. The Climate Trace facilities using the MCF versus AI method are noted separately (blue and red symbols, respectively).

4. Discussion

The Climate Trace emissions estimation for 96.3% of the total facilities analyzed here relies upon a combination of country/fuel-specific MCF, individual unit-scale nameplate capacity, and country/fuel-specific emissions rates (CO_2 per MWh). All three of these key variables for each facility are provided in the Climate Trace dataset. This allows for some diagnostic exploration of the reasons for the large facility differences found here.

Figure 3(a) shows the Climate Trace CFs plotted against the CFs from the Vulcan-power database (see Methods) separately noted as those using the MCF method (blue symbols) versus the AI method (red symbols). The country-MCFs used by Climate Trace for each fuel type category (i.e. coal, oil, natural gas, waste) is evident from the large number of individual facilities with an identical CF (predominantly the blue symbols in figure 3(a)). There is a second group of paired facilities (predominantly the red symbols in figure 3(a)) that show greater correspondence with the Vulcan CFs. These are facilities for which a fuel mix was present (and hence, averaging of multiple fuel type-specific CFs occurred) or the AI-based estimation procedure was used for a given facility.

The reason the individual Climate Trace power plant CO_2 estimates using the MCF method are subject to large differences to the Vulcan-power estimates is partly due to the fact that the CF of individual facilities deviates substantially from the Climate Trace country/fuel-average. However, of greater importance is that actual fuel-specific CFs across the 1726 facilities examined here have a non-normal distribution, obviating the validity of an arithmetic mean. A frequency distribution of the Vulcan CFs for the paired power plants, grouped by fuel, suggests that both the oil and natural gas facilities have distributions inappropriate for an arithmetic mean (see supplementary information figure S3). In the case of natural gas, almost 32% of the facilities have a CF of 0.02 (the mode center of the distribution) compared to the ostensible CF used by Climate Trace of 0.33. Similarly, for facilities using oil as fuel, 71.8% of facilities have a CF of 0.02, compared to the ostensible CF used by Climate Trace of 0.14. By contrast, the dominant CF for coal and waste-fired facilities is much closer to the mean value adopted by Climate Trace. Though in the case of coal-fired facilities, there is large spread of common CFs ranging from 0.02 to 0.74, highlighting the non-gaussian nature of power production infrastructure in the U.S.

The other component of the Climate Trace MCF-based calculation worth exploring is the emission rate. Figure 3(b) shows the Climate Trace emission rates plotted against those in the Vulcan-power dataset. Climate Trace derives the individual facility emission rates as the product of a 'base' emission rate, specific to fuel type and technology, and a 'country calibration factor' to account for regional differences globally. As with the CF, it is clear that there are clusters of nearly identical emission rates that tend to group according to fuel type. Unlike the CF analysis, the AI-based facilities (dominated by coal-burning facilities) cluster similarly to the MCF method facilities. The emission rates used by Climate Trace are consistent with the mode values of the distribution of fuel-specific emission rates from the Vulcan-power dataset (see supplementary information figure S4) with the exception of waste-burning facilities. Nevertheless, as was found with the CF analysis, the narrow set of values used by Climate Trace is likely not capturing the true variance of the emission rate at individual facilities and therefore contributes to the potential error at the facility-scale.

There is an urgent need for independent asset-scale GHG emissions information. As businesses, cities, and other sub-national entities engage in emissions mitigation planning, accurate GHG emissions with information on geolocation, fuel, sector, and combustion type are useful and allow for better mitigation prioritization and strategic planning ([2]; etc). This was part of the original motivation, for example, when constructing the U.S. Vulcan emissions data product. The 'bottom-up' or engineering approach followed by the Vulcan project is not feasible in most countries, save a few with large amounts of data collection and transparent archive systems.

Efforts such as Climate Trace are an attempt to overcome that geographic limitation, adopting techniques that have global application and rely on consistent across-country data. Power plants offer perhaps the best opportunity to estimate emissions from individual assets globally given their more manageable asset scale (in comparison to, for example, road segments or residential buildings) and relatively concentrated ('point-source') emissions characteristics. However, as we have demonstrated here, asset-scale estimation using the techniques adopted by Climate Trace is accompanied by large individual estimation bias, particularly for those facilities that relied upon the MCF approach.

While the approach combining ML and high-resolution satellite imagery remains a promising technique, it was applied to a small minority of the individual powerplant facilities. This was due to understandable limitations in the remote sensing information needed to actuate the AI method. As a result, the majority of the facilities in the U.S. example rely on an approach that, by definition, cannot accurately estimate powerplant emissions at the asset scale. Because of the nature of the underlying power plant operating characteristics, this method also cannot properly capture the mean behavior. Either the approach combining ML with high-resolution remote sensing is expanded to cover all asset-scale estimation (challenging given obstacles such as cloud cover and scene availability), or alternative techniques are developed that do not rely on national-mean characteristics of power plant activity.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgment

This work was made possible through support from the National Oceanic and Atmospheric Administration Grant NA19OAR4310167. The first author owns equity in CrossWalk Lab, a company distributing the Vulcan data product.

Ethical statement

This research did not involve humans or animals

ORCID iDs

Kevin R Gurney https://orcid.org/0000-0001-9218-7164
Bilal Aslam https://orcid.org/0000-0001-7308-5285
Jarrett J Barber https://orcid.org/0000-0002-9225-351X
Anna Kato https://orcid.org/0000-0003-2758-6817

References

[1] National Academies of Sciences, Engineering, and Medicine 2022 *Greenhouse Gas Emissions Information for Decision Making: a Framework Going Forward* (The National Academies Press) (See table 2.1) (https://doi.org/10.17226/26641)
[2] Gurney K R *et al* 2015 Climate change: track urban emissions on a human scale *Nature* **525** 179–81
[3] Climate Trace Independent greenhouse gas emissions tracking (available at: https://climatetrace.org)
[4] United Nations Climate Change, Side Events 2022 Al gore and the climate trace coalition release detailed inventory of the sources of GHG emissions worldwide (available at: https://unfccc.int/event/al-gore-and-the-climate-trace-coalition-release-detailed-inventory-of-the-sources-of-ghg-emissions)
[5] United Nations Climate Change, Special Event 2023 Al Gore and Climate Trace Unveil Game-Changing Greenhouse Gas Emissions Inventory (available at: https://unfccc.int/event/al-gore-and-climate-trace-unveil-game-changing-greenhouse-gas-emissions-inventory)
[6] Climate Trace, News & Insights 2023 Climate TRACE unveils open emissions database of more than 352 million assets (available at: https://climatetrace.org/news/climate-trace-unveils-open-emissions-database-of-more-than)
[7] Bloomberg 2023 Al gore-backed group has a tool to decarbonize supply chains (available at: www.bloomberg.com/news/articles/2023-12-03/cop28-al-gore-backed-group-releases-a-new-tool-to-track-co2-emissions)
[8] Gore A 2021 Measure emissions to manage emissions *Science* **378** 455
[9] Gurney K R, Liang J, Patarasuk R, Song Y, Huang J and Roest G 2020 The Vulcan version 3.0 high-resolution fossil fuel CO_2 emissions for the United States *J. Geophys. Res.: Atmos.* **125** e2020JD032974
[10] Gurney K R, Kato A and Dass P 2024 Vulcan-power version 1.1 *Zenodo* (https://doi.org/10.5281/zenodo.13755992)
[11] Climate Trace Data downloads (available at: https://climatetrace.org/data) (Accessed 20 December 2023)
[12] Couture H D *et al* 2024 Estimating carbon dioxide emissions from power plant water vapor plumes using satellite imagery and machine learning *Remote Sens.* **16** 1290

[13] United States Environmental Protection Agency 2015 40 DFR part 60, EPA-HQ-OAR-2013-0602; FRL-XXXX-XX-OAR, RIN 2060- AR33, carbon pollution emission guidelines for existing stationary sources: electric utility generating units (Accessed 3 August 2015)

[14] Code of Federal Regulations 2008 Protection of environment, environmental protection agency, 40 CFR part 75, continuous emission monitoring, code of federal regulations (available at: www.ecfr.gov/cgi-bin/text-idx?SID=4719db7a48cd26050b0732d0f9adc3ad&mc=true&node=pt40.2.51&rgn=div5) AERR summary www.epa.gov/air-emissions-inventories/air-emissions-reporting-requirements-aerr#rule-summary (Accessed 24 January 2008)

[15] United States Environmental Protection Agency 2005 *Plain English Guide to the Part 75 Rule* (U.S. Environmental Protection Agency, Clean Air Markets Division)

[16] United States Environmental Protection Agency 2010 *Draft Part 75 Emissions Monitoring Policy Manual* (U.S. Environmental Protection Agency, Clean Air Markets Division)

[17] United States Environmental Protection Agency, Clean Air Markets Program Data (available at: https://campd.epa.gov/data)

[18] Department of Energy/Energy Information Administration 2003 *Electric Power Monthly March 2003* Energy Information Administration, Office of Coal, Nuclear, and Alternate Fuels, U.S. Department of Energy

[19] United States Energy Information Administration 2023 Electricity, form EIA-923 detailed data with previous form data (EIA-906/920) (available at: www.eia.gov/electricity/data/eia923)

[20] Department of Energy/Energy Information Administration 2011 *Electric Power Annual 2010* (available at: www.eia.gov/cneaf/electricity/epa/epa_sum.html)

[21] Gurney K R, Huang J and Coltin K 2016 Bias present in US federal agency power plant CO_2 emissions data and implications for the US clean power plan *Environ. Res. Lett.* **11** 064005