Spatial Disaggregation of Remainder Emissions

Gary Collins[1,3], April Nellis[1,3], Krsna Raniga[2,3], Nicole Brown[1,3], Michael Pekala[1,3], Zoheyr Doctor[2,3], Dan Moore[2,3], Elizabeth Reilly[1,3], Marisa Hughes[1,3], and Gavin McCormick[2,3]

1) The Johns Hopkins University Applied Physics Laboratory, 2) WattTime, 3) Climate TRACE

1.   Introduction

For many bottom-up inventories, emissions are estimated at a regional level then distributed spatially in proportion with some proxy. For example, the Open-source Data Inventory for Anthropogenic CO2 (ODIAC) dataset estimates and distributes fossil fuel emissions in proportion to nightlight intensity as measured by the satellite remote sensing (Oda et al. 2018) and the Emissions Database for Global Atmospheric Research (EDGAR) supplements its bottom-up inventory with sector-specific proxies such as built-up nonresidential surface area (Crippa et al. 2024). A distinguishing characteristic of Climate TRACE data is that it is typically generated in a different order: asset-level locations are identified first then emissions are estimated upon them. This approach produces higher spatial accuracy of emissions estimates and allows for better attribution of responsibility of sources of emissions.

However, despite the efforts of the Climate TRACE coalition, global coverage of emissions at an asset-level is not yet achieved. For some sectors, this is due to the lack of asset-level information (e.g. lime manufacturing locations, a large emitting sector). In other sectors, the specific spatial distribution of emissions is hard to identify (e.g. manure management of free grazing cattle). To fill this gap, Climate TRACE estimates the quantity of "remainder emissions" by taking the difference between asset-level emissions and the highest quality country-level emissions estimate available (be it Climate TRACE or other emission inventories). These remainder emissions require distribution spatially in order to be useful for policy-makers, scientists, and activists alike; and hence specialized spatial proxies are used and the methodology for developing these proxies are described herein.

2.   Data and Methods

2.1.   Overview
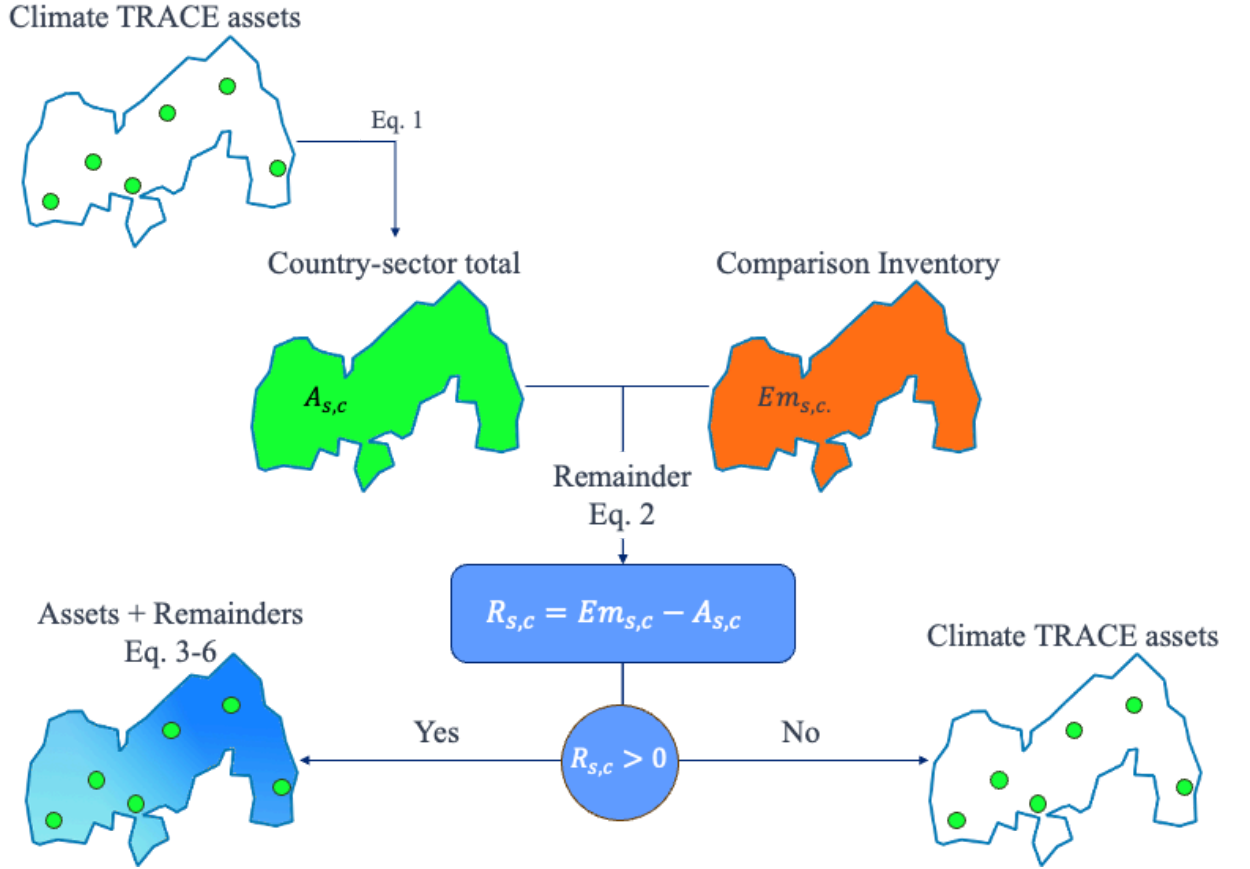
Figure 1. Climate TRACE remainder disaggregation process. This primarily consists of comparing Climate TRACE's asset total for a given country c and sector s to some comparison inventory (EDGAR, UNFCCC, etc.) and disaggregating that remainder is greater than 0.

Climate TRACE spatial products are a combination of both location-specific asset-level data and proxy-distributed country/sector-level remainder emissions. To generate these products, asset-level data was produced on some spatial support z (be it gridded data or administrative boundaries like Global Administrative Areas (GADM) (Global Administrative Areas. 2022). Then an asset-level country-sector sum was produced by,

$$A_{s,c} = \Sigma_z \, a_{s,c}(z), \text{ Eq. 1,}$$

where "$a_{s,c}(z)$" is the sum of the asset emissions for sector s at location z. This may differ from the "best" country-level estimate of the emissions, $Em_{s,c}$, since some sectors do not have all their emissions accounted for. The "best" country-level emissions estimate may be an output of a separate model, which was the case for wastewater treatment and disposal, or

it could be the emissions estimate from another bottom-up inventory like EDGAR. The remainder emissions are then computed by taking the difference between the two, expressed by

$$R_{s,c} = Em_{s,c} - A_{s,c}, \text{Eq. 2.}$$

In the following, we cover the details of the method implementation. We discuss the proxies and data used to spatially distribute remainders in 2.2, and provide more insight into our methods in 2.3. We first cover the single-proxy approach which is applied to a majority of the sectors in the database in 2.3.1. Then, we showcase a novel learned ensemble proxy which leverages high-confidence bottom-up Climate TRACE data in 2.3.2. The benefits of this data-driven method are discussed for rice cultivation, our selected pilot sector.

2.2. Data

2.2.1. Proxy Data

We use several proxies from the Global Human Settlement Layer (GHSL), as well as nightlights and population. Specifically, we used the following proxies to facilitate spatial remainder disaggregation:

Table 1. Emissions proxy data for remainder disaggregation.

Proxy	Source
Agricultural regions (Levels 11 and 12 of the GHS Settlement Model Degrees of Urbanisation)	*Schiavina M., Melchiorri M., Pesaresi M. (2023)*: **GHS-SMOD R2023A - GHS settlement layers, application of the Degree of Urbanisation methodology (stage I) to GHS-POP R2023A and GHS-BUILT-S R2023A, multitemporal (1975-2030).** European Commission, Joint Research Centre (JRC) PID: http://data.europa.eu/89h/a0df7a6f-49de-46ea-9bde-563437a6e2ba
GHS Distribution of Built-up Surfaces (Non-residential)	*Pesaresi M., Politis P. (2023)*: **GHS-BUILT-S R2023A - GHS built-up surface grid, derived from Sentinel2 composite and Landsat, multitemporal (1975-2030)** European Commission, Joint Research Centre (JRC) PID: http://data.europa.eu/89h/9f06f36f-4b11-47ec-abb0-4f8b7b1d72ea
GHS Distribution of Built-up Volumes	*Pesaresi M., Politis P. (2023)*: **GHS-BUILT-V R2023A - GHS built-up volume grids derived from joint assessment of Sentinel2, Landsat, and global DEM data, multitemporal (1975-2030).** European Commission, Joint Research Centre (JRC) PID: http://data.europa.eu/89h/ab2f107a-03cd-47a3-85e5-139d8ec63283
Night Light Intensity	Elvidge, C.D, Zhizhin, M., Ghosh T., Hsu FC, Taneja J. **Annual time series of global VIIRS nighttime lights derived from monthly averages:2012 to 2019**. Remote Sensing 2021, 13(5), p.922, doi:10.3390/rs13050922
Population	*Schiavina M., Freire S., Carioli A., MacManus K. (2023)*: **GHS-POP R2023A - GHS population grid multitemporal**

Proxy	Source
	(1975-2030). European Commission, Joint Research Centre (JRC) PID: http://data.europa.eu/89h/2ff68a52-5b5b-4a22-8f40-c41da8332cfe.

Table 3 in the supplemental section shows the mapping between Climate TRACE sectors and their associated proxies. These proxies were selected by referencing proxies used by other greenhouse gas databases. Namely, ODIAC uses night light intensity to approximate CO_2 emissions (Oda et al. 2016), while EDGAR uses GHSL settlement layers as well as population data and built-up nonresidential surface area and volume for various sectors (Crippa et al. 2024).

2.2.2. Climate TRACE Database Data

Section 2.3 discusses the two different approaches to spatial disaggregation using proxies, both of which rely on Climate TRACE asset-level data aggregated to the GADM 2 level as discussed in Eq. 1 and illustrated in Figure 1.

2.3. Methods

2.3.1. Remainders

As shown in Table 2 in the supplemental section, in 2022 approximately 15 GT of global CO_2 equivalent emissions (computed from CO_2, CH_4, and N_2O) were not accounted for by Climate TRACE assets, which is approximately 20% of all CO_2 equivalent emissions. Figure 2 and Figure 3 display the global distribution of remainders. While China has the largest nominal remainder emissions, many other regions have significant relative remainder emissions, such as Djibouti and Ireland. Few regions have small relative gaps, like in the United States where reporting standards are very high.
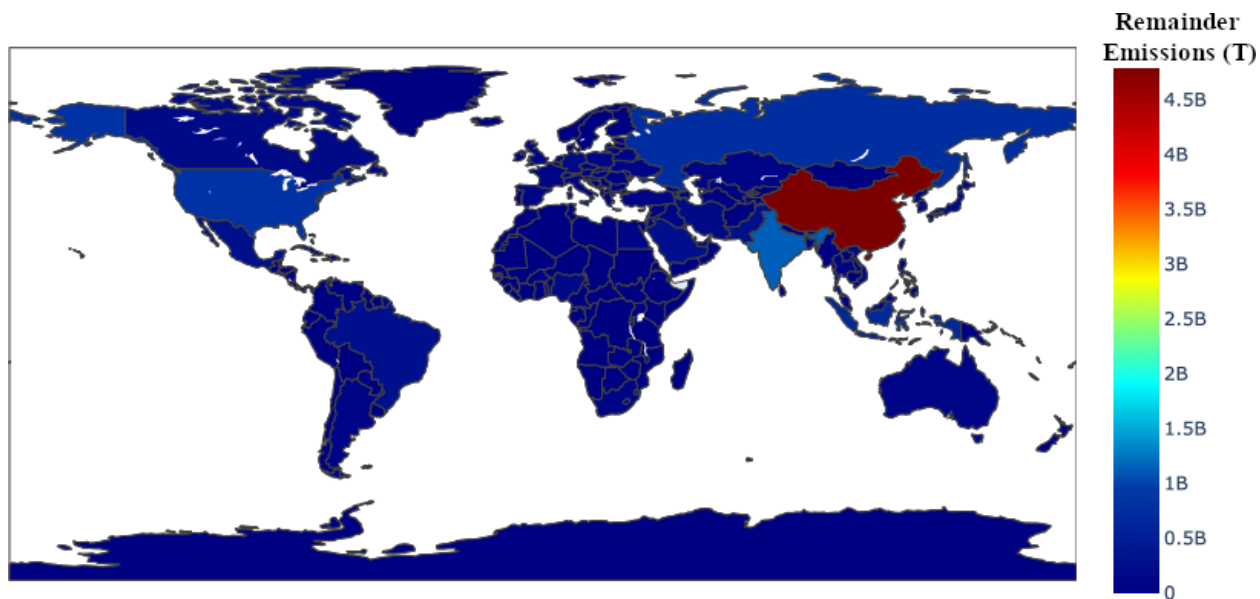
Figure 2. Distribution of nominal CO_2 equivalent remainder values in 2022.
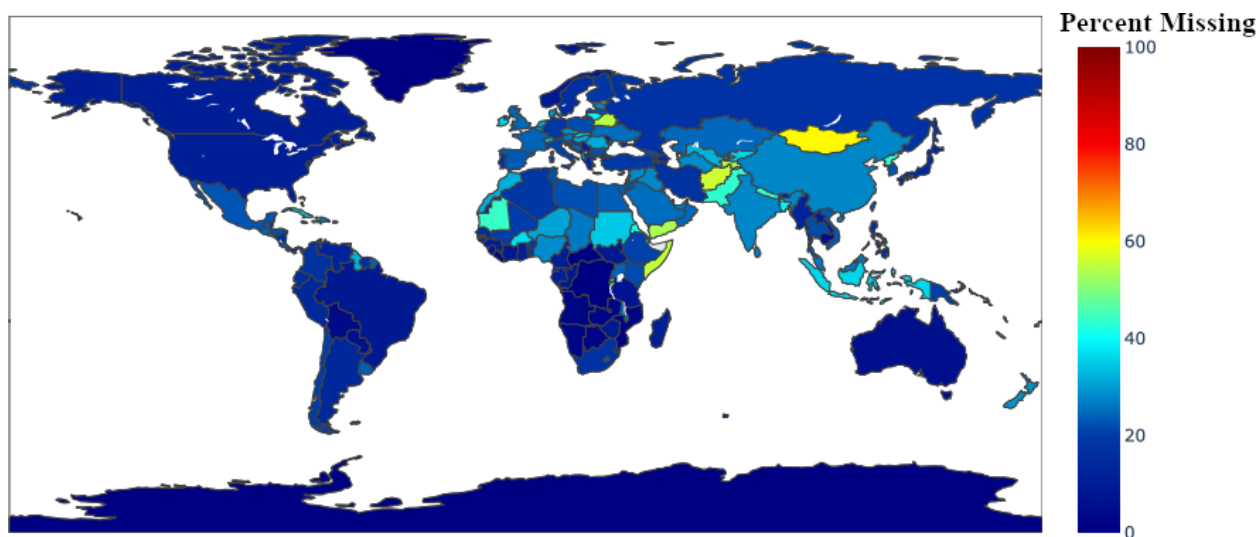
Figure 3. Distribution of CO_2 equivalent remainder values in 2022 as a percentage of total emissions in each country.

Many of the largest sectors in which these emissions were missing are the "other" sectors, which are composed by taking the difference between larger EDGAR sectors and Climate TRACE subsectors (e.g. subtracting all Climate TRACE industrial sectors from the wider EDGAR IND sector). These unaccounted-for emissions have no corresponding asset information and thus redistributing their emissions requires proxying. Other gaps exist due to there being a difference between country-level bottom-up estimates vs. total asset-level emissions, like cement.

In most inventories, the distribution of country-level emissions can be expressed by,

$$em_{s,c}(z) = \frac{p_{s,c}(z)}{\Sigma_z p_{s,c}(z)} \cdot Em_{s,c} \text{ Eq. 3,}$$

where $p_{s,c}(z)$ is the country-sector proxy value at location z. Unlike other inventories, Climate TRACE modifies the proxy in Equation 3 to take into account asset-level information for a given sector. As Climate TRACE may have more complete emissions estimates for a given sector in some subregions than others, the modified proxy is developed by removing distributed assets emissions from Equation 1, then renormalizing the output before distributing remainders, defined by

$$\tilde{p}_{s,c}(z) = max(em_{s,c}(z) - a_{s,c}(z),\ 0) \text{ Eq. 4,}$$

where the maximum is taken to prevent negative proxy values. The remainder distribution for Climate TRACE uses these modified proxies with,

$$r_{s,c}(z) = \frac{p_{s,c}(z)}{\Sigma_z p_{s,c}(z)} \cdot R_{s,c}, \text{ Eq. 5,}$$

with the total emissions of a location z for country c arising from sector s defined with,

$$em_{s,c}(z) = a_{s,c}(z) + r_{s,c}(z) \text{ Eq. 6.}$$

A typical use for the spatial distribution of emissions is producing gridded, latitude-longitude data that is readily usable for climate modeling. However, Climate TRACE produces data also at administrative boundaries. Hence, we use a generic spatial variable z.

A visual example of modified proxy generation and remainder distribution is shown in Figure 4.

Equations 3 and 4

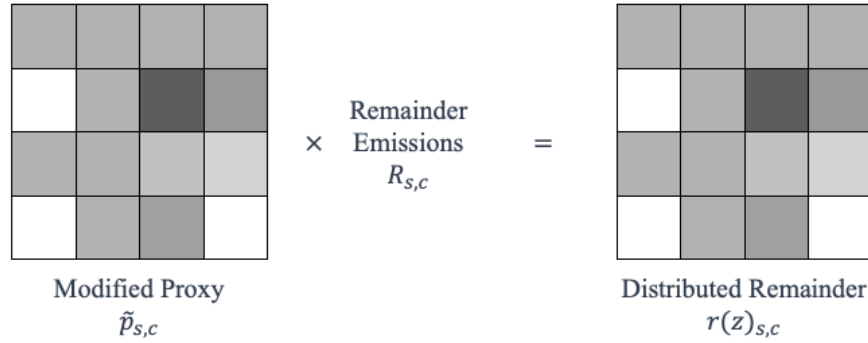Proxy of Principality of Square $\dfrac{p_{s,c}(z)}{\Sigma p_{s,c}}$ $\times$ Comparison Inventory $Em_{s,c.}$ $-$ Climate TRACE Assets Distributed $a_{s,c}(z)$ $=$ Modified Proxy $\tilde{p}_{s,c}$

Equations 5

Modified Proxy $\tilde{p}_{s,c}$ $\times$ Remainder Emissions $R_{s,c}$ $=$ Distributed Remainder $r(z)_{s,c}$
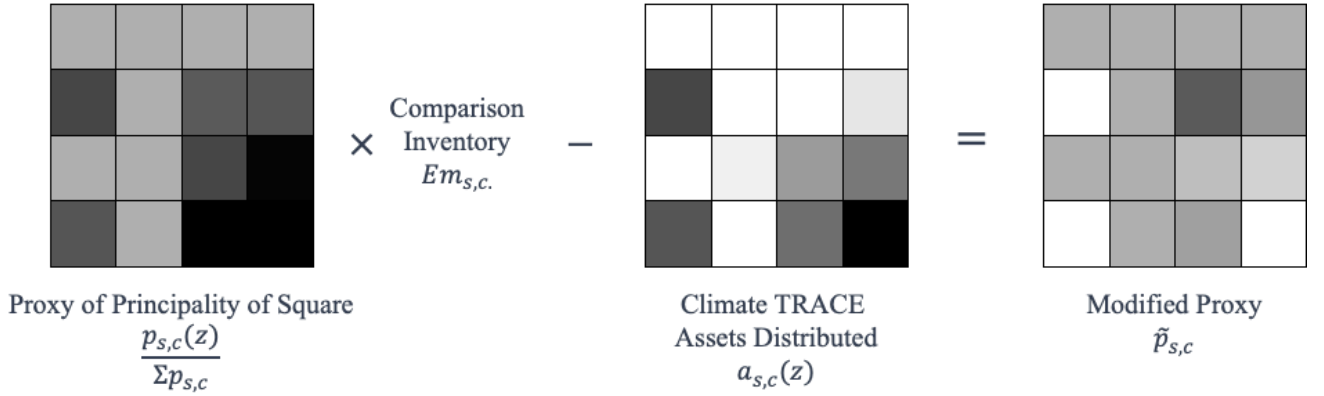
Figure 4. A visualization of Equations 3, 4, and 5. Equation 3 and 4 takes a proxy of a region and modifies it by removing mass from Climate TRACE asset locations. This modified proxy is then used to disaggregate the remainder with Equation 5.

2.3.2. Ensemble Proxies

A host of factors affect the spatial distribution of GHG emissions, from proximity of transportation to natural terrain. While it is an established approach to consider sector-dependent proxies, as in Crippa et al. (2024) and Climate TRACE methodology documents (versions Fall 2024), much of the underlying complexity is still lost when relying on a single proxy to estimate the spatial distribution of remainder emissions. Therefore, we have developed a method for determining the spatial distribution of emissions remainders using an *ensemble* of proxies. This method was prototyped for the rice-cultivation sector, whereas all other sectors used proxies listed in Table 3 of the supplemental for remainder disaggregation.

Since Climate TRACE has amassed an extensive database of point-source GHG emitters (assets) with both location and emissions data, this information can be used as the basis for data-driven ensemble proxy development. In countries and sectors where a vast majority of

GHG emissions data is accounted for by assets in the Climate TRACE database, proxies can be evaluated based on their consistency with this granular data using a variety of metrics. Furthermore, the ability to evaluate proxies against a high-certainty dataset opens the door to training and evaluating ensemble proxies. For this process, we use only data from countries where Climate TRACE assets account for at least 99.99% of total country emissions. These countries include the top five global rice producers (China, India, Bangladesh, Indonesia, and Vietnam), which produce a combined 72% of the world's rice (Foreign Agriculture Service. 2024).

As an initial case study, we apply a random forest (RF) regressor to the problem of distributing remainder methane emissions associated with rice cultivation across GADM 2 regions. The input data for the RF for a given GADM region is a vector of the various proxy values, while the output data is the predicted emissions level. We evaluate the performance of the trained output using the GADM-level methane emissions calculated from Climate TRACE rice cultivation assets, normalized by country. The accuracy of the proxy is assessed using the mean squared error between the ensemble proxy and the Climate TRACE distribution. In order to maximize the transferability of the model to countries with sparse asset-level data, we perform hyperparameter optimization and cross-validation on the random forest regressor. The out-of-sample performance of the regressor is evaluated by training the random forest on a subset of the countries with high completion and validating the trained random forest on the remaining countries with high completion. In addition, the random forest's performance is evaluated on each country individually, which helps identify and avoid overfitting.

Because this approach performed well on methane emissions due to rice cultivation, we hope to expand the methodology to more sectors in 2025.

3.   Results

3.1. Single Proxy

Using the proxy methodology discussed in 2.3.1, remainder emissions from each sector except rice cultivation (the pilot sector for the ensemble proxy) were spatially distributed based on the single proxy in Table 3 corresponding to the given sector.

Roughly one third of the emissions from cement is unaccounted for among Climate TRACE data as shown in Table 2. These emissions are disaggregated using average nightlight intensity. Figure 5 shows the distribution of cement emissions in Mexico, aggregated to GADM Administration Level 1 boundaries.
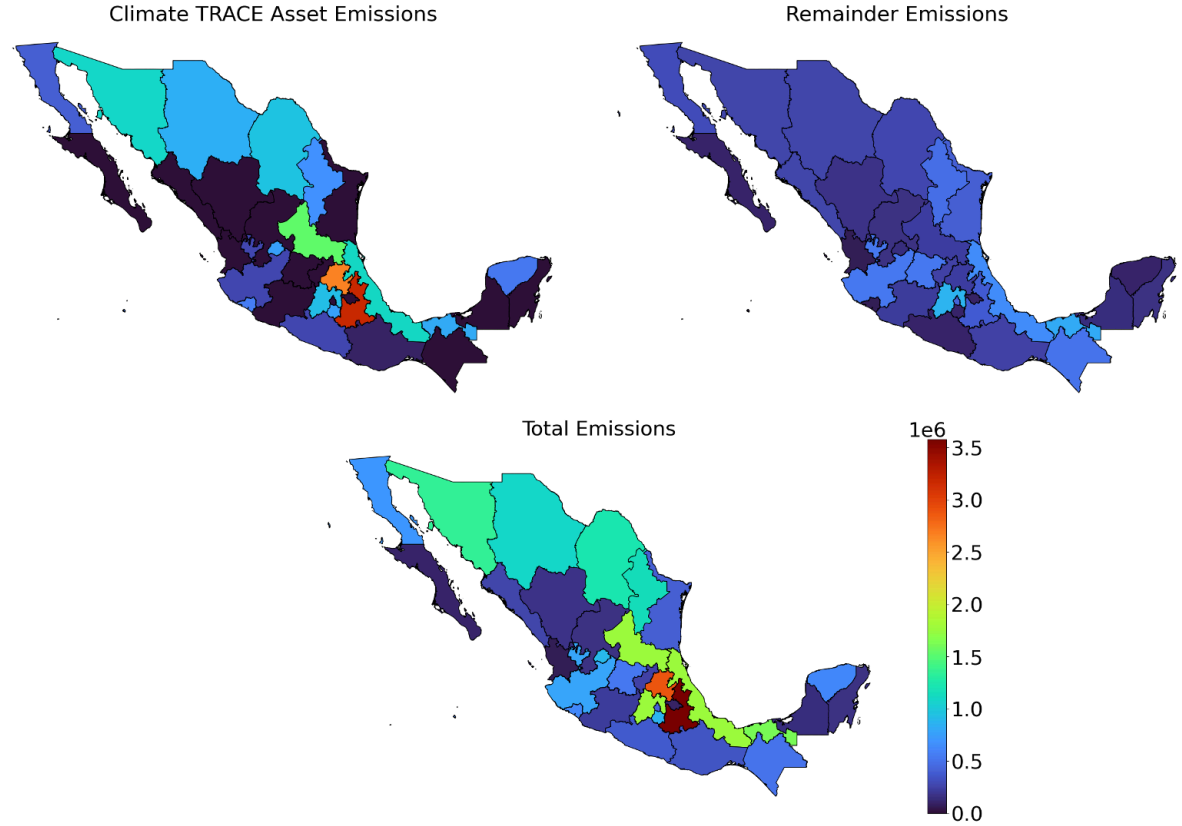
Figure 5. Asset, remainder, and total emissions for Mexico from cement at GADM Administration 1. The Asset Emissions subfigure shows the total emissions from cement accounted for in the Climate TRACE database, the remainder emissions dataset contains the unaccounted for emissions after disaggregation, and the third plot is of the sum of the first two.
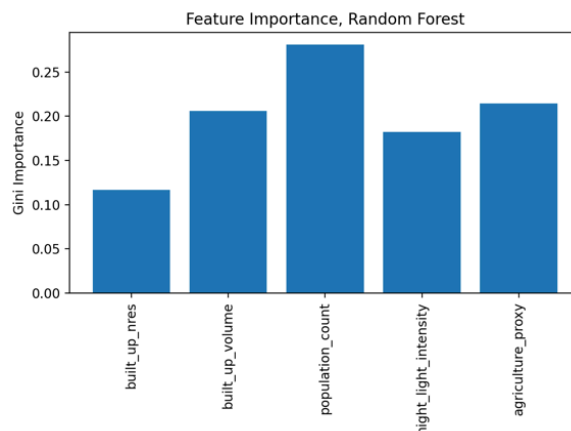
3.2. Ensemble Proxy

We piloted the ensemble proxy approach discussed in 2.3.2 for the rice cultivation sector, which was chosen because Climate TRACE has very good data for some major rice-producing countries like China and India, but poor data for others like Nigeria, which produced over 5.5 million metric tons of rice from 2023-2024 (Foreign Agriculture Service. 2024). In addition, rice cultivation stood out as a sector where industrial signals like night lights and building volume would not be appropriate, but agricultural signals like GHS-SMOD R2023A low-density and very-low-density regions also did not clearly correlate with emissions. The correlation between Climate TRACE emissions data and each proxy is shown in Figure 6, where *built_up_nres, built_up_volume, population_count, night_light_intensity*, and *agriculture_proxy* are defined as in Table 1, and emissions_quantity represents the methane emissions due to rice cultivation. Rice

cultivation results mainly in methane emissions, so we focus on methane in the rest of this section.
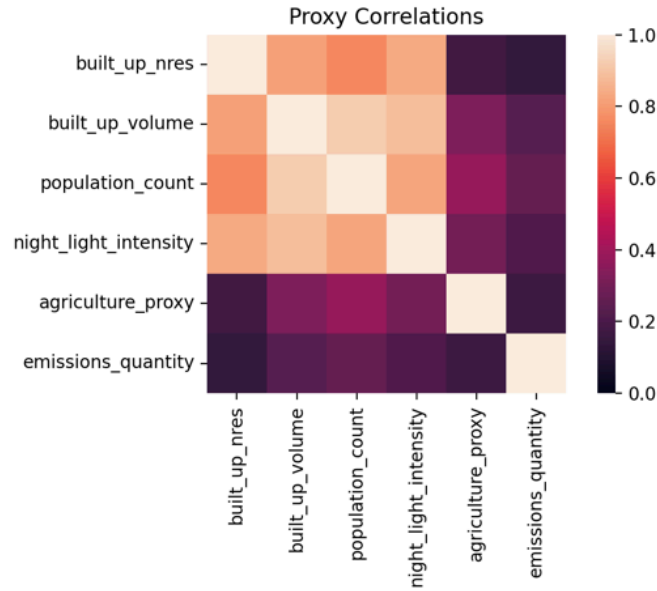
Figure 6. Correlation Between Common Proxies and Methane from Rice Cultivation

We further assessed the relationship between the proxies by evaluating the mean squared error between each proxy and the Climate TRACE emissions distribution and comparing that to the results of the RF-produced ensemble proxy. After evaluating the suite of proxies that we assembled for the ensemble, we found that no single proxy uniformly outperformed the other proxies in all countries in the training set. This supports the idea that an ensemble proxy is needed for a more generalizable and accurate proxy, at least when discussing rice cultivation. This is further supported by the resultant RF feature importance analysis shown in Figure 7. Notice that no single proxy contributes overwhelmingly to the RF ensemble model.

Figure 7. Random Forest Feature Importance

Furthermore, the trained RF produced an ensemble proxy that was able to outperform all the considered individual proxies, including the agriculture proxy which showed the most initial promise (since rice is an agricultural product). When compared specifically to the agriculture proxy, the error in the ensemble proxy was lower than the error from the agricultural proxy in 33 out of 34 countries where Climate TRACE has near-complete data. These 33 countries encompass 99.5% of the methane emissions attributed to countries where Climate TRACE has high completeness (>99.9%) in their bottom-up data. The difference in the mean squared errors for the agricultural and ensemble proxies are displayed in Figure 8. Furthermore, the ensemble proxy places a higher percentage of emissions in the correct GADM region. While 74.5% of methane emissions were accurately attributed to the correct GADM region using the agricultural proxy, the ensemble proxy was able to accurately place 85.7% of global methane emissions. This improvement is significant, though of course there is room for more improvement in future iterations of the ensemble.
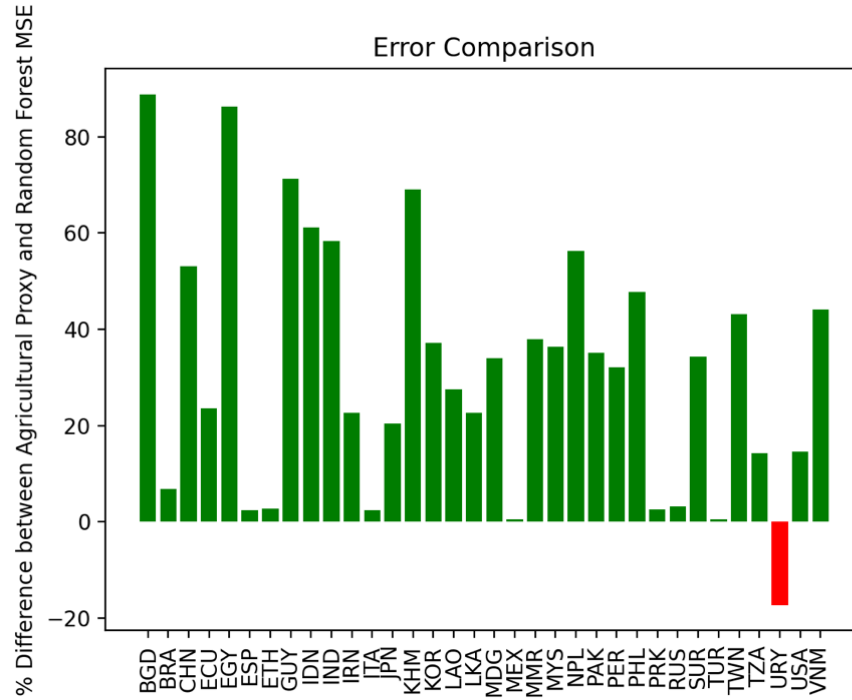
Figure 8. Percent Difference in the Error Between RF and Agricultural Proxies

The emissions distribution predicted by the proxy for a portion of Asia is shown in Figure 9 and can be compared to the Climate TRACE emissions data depicted in Figure 10. The random forest proxy is able to allocate emissions to GADM regions where Climate TRACE also records emissions. However, one drawback of all the proxies we studied, including the trained ensemble, was that they were not able to accurately capture outlier emissions, such as the 241,727 tons of methane emissions recorded for Aksu prefecture in the Xinjiang Uygur region of China (GADM CHN.28.1_1).
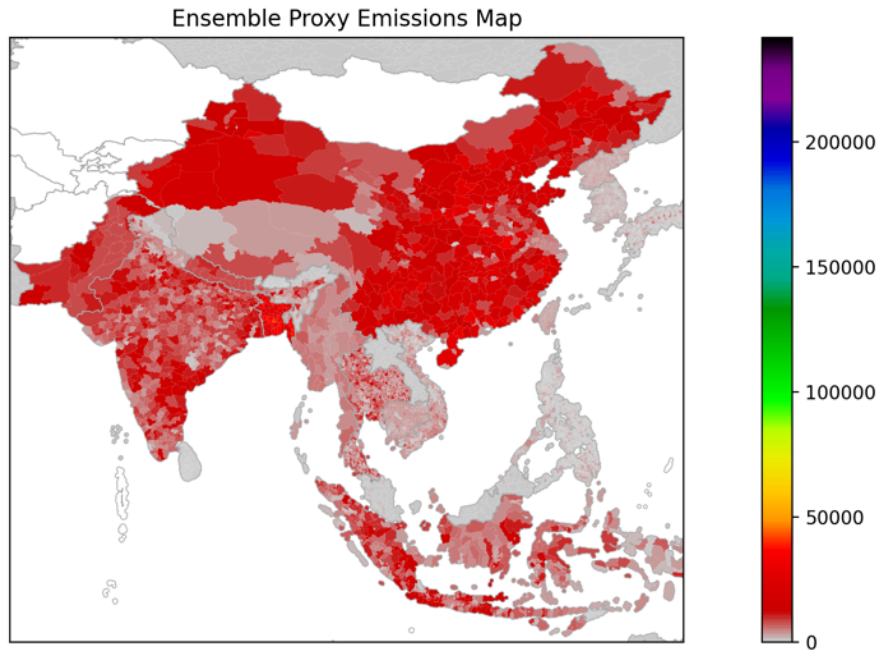
Figure 9. Ensemble Proxy Emissions Distribution
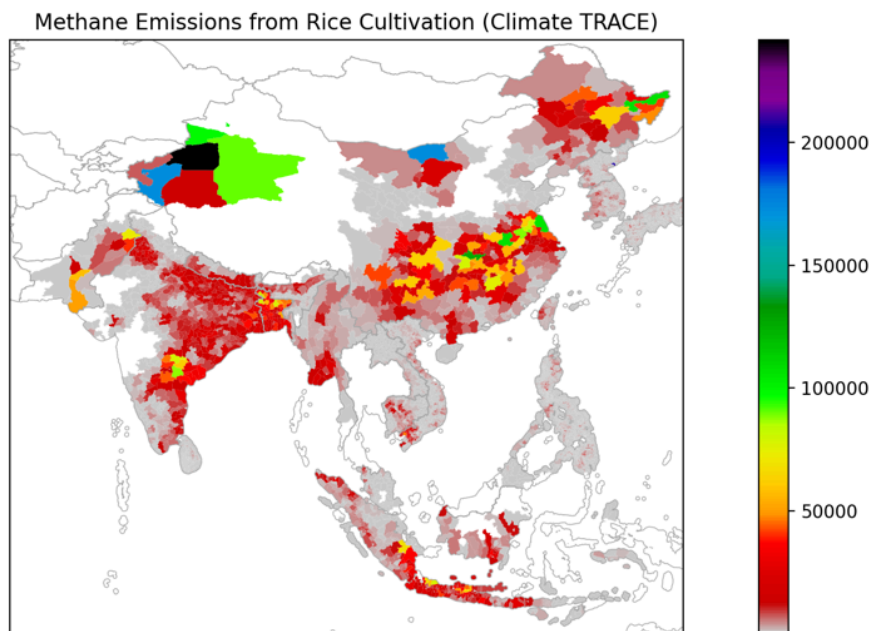
Figure 10. Climate TRACE Database Emissions Distribution

4.   Conclusions

Spatially distributed products from Climate TRACE use a combination of location-specific asset-level information as well as proxied remainder data on modified supports. The

modified proxies take traditional proxies (like nightlights and human population count) and remove mass (GHGs) from regions where asset-level GHG emissions are known.

We take a data-driven, rather than heuristic, approach to the use of proxies for spatial disaggregation of remainder emissions. This is made possible because the Climate TRACE database contains near-complete bottom-up rice cultivation emissions data for 34 countries which account for 77.6% of estimated global methane emissions due to rice cultivation. With this additional information, we are able to produce new insights regarding the accuracy of current commonly-used proxies and propose an ensemble proxy which combines features of multiple single proxies.

First, we find that of the commonly used emissions proxies, no single proxy has been found to accurately reflect the distribution of methane emissions from rice cultivation. However, an ensemble proxy trained on asset-level data has the potential to provide a more tailored, data-informed approach to spatial distribution proxying. In fact, we find that the inclusion of high accuracy point source emissions data provides a detailed picture of emissions distributions that commonly used proxies are not able to accurately recreate. For example, in a majority of countries analyzed, the true GADM2-level spatial distribution of rice cultivation methane emissions was better approximated by a trained ensemble proxy than by an agricultural proxy which is the current preferred proxy used by entities like EDGAR and the Community Emissions Data System for Historical Emissions (CEDS), as discussed in Crippa et al. 2024 and Hoesly et al. 2018.

While the random forest regressor is flexible and learns well, it does not take the proximity of nearby GADM regions into account. To address this, we have begun development on a more advanced approach to emissions distributions. Emissions and proxies are discretized on a global grid with 1 degree by 1 degree resolution. The distributions can then be analyzed using methods developed for image analysis. Specifically, we train a convolutional neural network (CNN) on the gridded proxies and evaluate it against gridded Climate TRACE emissions data. By maintaining the spatial relationships between regions, training on the CNN should be able to improve the accuracy of the ensemble proxies.

5. Acknowledgements

6. Supplementary Materials

Table 2. Remainders for 2022 listed by sector in order from largest to smallest remainder emissions.

Sector	Climate TRACE Total Emissions (MT)	Remainder Emissions (MT)	%Missing
All Sectors	*77581.62*	*15474.44*	*19.95*
other-energy-use	1804.4	1804.4	100.00
other-manufacturing	1724.75	1707.36	98.99
fluorinated-gases	1537.12	1537.12	100.00
other-agricultural-soil-emissions	1374.87	1374.87	100.00
other-chemicals	1200.05	1110.63	92.55
enteric-fermentation-cattle-operation	1014.39	908.05	89.52
enteric-fermentation-other	792.57	792.57	100.00
cement	2272.78	771.88	33.96
solid-fuel-transformation	672.71	672.71	100.00
other-fossil-fuel-operations	597.66	597.66	100.00
iron-and-steel	3509.9	597.51	17.02
domestic-wastewater-treatment-and-discharge	743.41	593.11	79.78
solid-waste-disposal	1337.74	463.25	34.63
electricity-generation	13237.93	318.7	2.41
manure-management-other	225.1	225.1	100.00
other-onsite-fuel-usage	202.66	202.66	100.00
crop-residues	201.77	201.77	100.00
manure-applied-to-soils	167.85	167.85	100.00
other-transport	162.65	162.65	100.00
industrial-wastewater-treatment-and-discharge	166.37	156.46	94.04
aluminum	436.39	135.63	31.08
food-beverage-tobacco	263.76	109.42	41.49
railways	102.83	102.83	100.00
heat-plants	91.36	91.36	100.00

Sector	Climate TRACE Total Emissions (MT)	Remainder Emissions (MT)	%Missing
international-shipping	677.67	87.07	12.85
lime	505.25	76.98	15.24
manure-management-cattle-operation	96.31	74.56	77.42
incineration-and-open-burning-of-waste	66.26	66.26	100.00
rice-cultivation	762.92	64.71	8.48
pulp-and-paper	106.6	63.48	59.55
iron-mining	119.38	50.69	42.46
other-mining-quarrying	32.6	32.6	100.00
oil-and-gas-transport	2009.54	26.35	1.31
biological-treatment-of-solid-waste-and-biogenic	24.63	24.63	100.00
other-metals	58.22	20.33	34.92
residential-onsite-fuel-usage	3370.14	20.2	0.60
wood-and-wood-products	15.6	15.6	100.00
copper-mining	87.58	13.36	15.25
glass	46.72	7.78	16.65
textiles-leather-apparel	52.95	7.22	13.63
oil-and-gas-refining	969.73	6.17	0.64
bauxite-mining	12.18	2.87	23.61
road-transportation	6440.17	1.7	0.03
non-residential-onsite-fuel-usage	460.63	1.34	0.29
rock-quarrying	1.25	1.25	100.00
petrochemical-steam-cracking	284.94	1.13	0.40
forest-land-fires	5650.01	1.1	0.02
shrubgrass-fires	3209.72	0.71	0.02
sand-quarrying	0.56	0.56	100.00
wetland-fires	43.4	0.15	0.35

Sector	Climate TRACE Total Emissions (MT)	Remainder Emissions (MT)	%Missing
synthetic-fertilizer-application	542.05	0.04	0.01
domestic-shipping	387.77	0.02	0.00
domestic-aviation	297.43	0.01	0.00
international-aviation	414.87	0.01	0.00
cropland-fires	1392.76	0.01	0.00
manure-left-on-pasture-cattle	373.85	0	0.00
forest-land-clearing	5295.84	0	0.00
removals	125.52	0	0.00
coal-mining	1848.5	0	0.00
forest-land-degradation	367.4	0	0.00
net-wetland	163.88	0	0.00
net-forest-land	936.91	0	0.00
net-shrubgrass	922.64	0	0.00
chemicals	602.05	0	0.00
enteric-fermentation-cattle-pasture	1130.02	0	0.00
oil-and-gas-production	3757.5	0	0.00
water-reservoirs	76.65	0	0.00

Table 3 Proxies used for each sector for remainder disaggregation. Night lights refers to the VIIRS night light maps (Elvidge *et al.* 2021), population count is from the GHSL population count for 2020, and the agriculture proxy product is from GHSL (Schiavina *et al.* 2023).

Category	Sector Name	Proxy
Fossil fuel, Manufacturing, and Transport Sectors	aluminum	Night Light Intensity
	bauxite-mining	
	cement	

Category	Sector Name	Proxy
	chemicals	
	coal-mining	
	copper-mining	
	domestic-aviation	
	domestic-shipping	
	electricity-generation	
	food-beverage-tobacco	
	fluorinated-gases	
	glass	
	heat-plants	
	international-aviation	
	international-shipping	
	iron-and-steel	
	iron-mining	
	lime	
	non-residential-onsite-fuel-usage	
	oil-and-gas-production	
	oil-and-gas-refining	
	oil-and-gas-transport	
	other-chemicals	
	other-energy-use	
	other-fossil-fuel-operations	
	other-manufacturing	
	other-metals	
	other-mining-quarrying	
	other-onsite-fuel-usage	

Category	Sector Name	Proxy
	other-transport	
	petrochemical-steam-cracking	
	pulp-and-paper	
	railways	
	residential-onsite-fuel-usage	
	road-transportation	
	rock-quarrying	
	sand-quarrying	
	solid-fuel-transformation	
	textiles-leather-apparel	
	wood-and-wood-products	
Waste Sectors	biological-treatment-of-solid-waste-and-biogenic	Population Count
	incineration-and-open-burning-of-waste	
	domestic-wastewater-treatment-and-discharge	
	industrial-wastewater-treatment-and-discharge	
	solid-waste-disposal	
Agriculture Sectors	crop-residues	GHSL *Very Low Density Rural* + GHSL *Rural Cluster*
	cropland-fires	
	enteric-fermentation-cattle-operation	
	enteric-fermentation-cattle-pasture	
	enteric-fermentation-other	
	forest-land-clearing	
	forest-land-degradation	
	forest-land-fires	
	manure-applied-to-soils	
	manure-left-on-pasture-cattle	

Category	Sector Name	Proxy
	manure-management-cattle-operation	
	manure-management-other	
	net-forest-land	
	net-shrubgrass	
	net-wetland	
	other-agricultural-soil-emissions	
	shrubgrass-fires	
	synthetic-fertilizer-application	
	water-reservoirs	
	wetland-fires	
	removals	
	rice-cultivation	Ensemble Proxy

Permissions and Use: All Climate TRACE data is freely available under the Creative Commons Attribution 4.0 International Public License, unless otherwise noted below.

Data citation format: Collins, G., Nellis, A., Raniga, K., Brown, N., Pekala, M., Doctor, Z., Moore, D., Reilly, E., and Hughes, M. (2024). *Spatial Disaggregation of Remainder Emissions.* The Johns Hopkins University Applied Physics Laboratory (JHU/APL) and WattTime, 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.

References

1. Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., Dijkstra, L., Van Damme, M., Clarisse, L., and Coheur, P.: Insights into the spatial distribution of global, national, and subnational greenhouse gas emissions in the Emissions Database for Global Atmospheric Research (EDGAR v8.0), Earth Syst. Sci. Data, 16, 2811–2830, doi:10.5194/essd-16-2811-2024, 2024.
2. Florczyk, A., C. Corbane, M. Schiavina, M. Pesaresi, L. Maffenini, M. Melchiorri, P. Politis. 2019. "GHS Urban Centre Database 2015, Multitemporal and Multidimensional Attributes, R2019A." European Commission, Joint Research Centre (JRC). https://data.jrc.ec.europa.eu/dataset/53473144-b88c-44bc-b4a3-4583ed1f547e.
3. Foreign Agriculture Service. 2024. *2023/2024 Rice Production*. United States Department of Agriculture. https://fas.usda.gov/data/production/commodity/0422110, 2024.
4. Global Administrative Areas. 2022. "GADM Database of Global Administrative Areas, Version 4.1." https://www.gadm.org.
5. Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369-408, https://doi.org/10.5194/gmd-11-369-2018, 2018.
6. Oda, T., Maksyutov, S., and Andres, R. J.: The Open-source Data Inventory for Anthropogenic CO2, version 2016 (ODIAC2016): a global monthly fossil fuel CO2 gridded emissions data product for tracer transport simulations and surface flux inversions, Earth Syst. Sci. Data, doi:10.5194/essd-10-87-2018, 2018.