Completeness of Bottom-up Emissions Estimates and Associated Metadata

Dan Moore[1,3], Zoheyr Doctor[1,3], Krsna Raniga[1,3], Christy Lewis[1,3], Lekha Sridhar[1,3], Peter Thomas[1,3], Ishan Saraswat[1,3], Lauren Betz[1,3], Amy Piscopo[1,3], Gary Collins[2,3], April Nellis[2,3], Nicole Brown[2,3], Michael Pekala[2,3], Justin Rokisky[2,3], Elizabeth Reilly[2,3], Marisa Hughes[2,3], and Gavin McCormick[1,3]

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

1. Introduction

Spatial, temporal, and sectoral completeness of the Climate TRACE emissions inventory is critical to reducing greenhouse gas and pollutant emissions: A complete and bottom-up accounting of emissions provides a full view of when and where emissions *could* be reduced. Furthermore, metadata for these emissions – from the magnitude of emissions-causing activity to the confidence level of estimates – provide insight into *how* emissions can be reduced and how effective those changes might be. However, the raw measurements and inferences performed for each sub-sector may not be complete, meaning that missing values or other minor inconsistencies must be addressed post-hoc. In this document, we describe the suite of imputation techniques we use to ensure every reported source of emissions has estimates for all reported times in the Climate TRACE inventory, and that each of these estimates has associated metadata.

While each subsector's emissions may be modeled independently, they can all be cast into a sector-agnostic mathematical framework (US EPA, 2024). This enables us to use a common set of interventions to clean the data and impute any missing values, regardless of the sector, location, and gas. These interventions generally fall into three categories of imputation: fully constrained imputation, under-constrained imputation, or over-constrained imputation. In the fully constrained case, the existing information about an emitting source or "asset" is enough to constrain the values of any missing information. In the under-constrained case, only some information is provided, and missing information must either be interpolated from other Climate TRACE data or provided by an external source. Lastly, in the over-constrained case, information may exist for all or some of the emissions and metadata quantities, but some of the values are in contradiction to each other.

In Section 2, we describe the raw Climate TRACE data which may have missing or inconsistent information and then the set of interventions applied to achieve completeness. In Section 3, we detail the outcome of these interventions. Finally, in Sections 4 and 5, we discuss the implications of the interventions and outline future improvements.

2. Materials and Methods

2.1 Datasets Employed

The primary input dataset was the bottom-up inventory of emissions and metadata for each Climate TRACE asset for a given start and end time and gas as provided by Climate TRACE coalition members for each sector. These inputs came either in a catalog form or in rasterized form. In the rasterized case, we aggregated each quantity to predetermined spatial boundaries, namely the Database of Global Administrative Areas (GADM, 2022) and Global Human Settlement Functional Urban Areas (Schiavina et al 2019), and converted them to a catalog form which was concatenated with the rest of the data set. Table 2 in the Supplemental Materials shows the counts of the number of rows for each sector and gas from these data over all times starting from 2021.

The following metrics are required for each asset for comprehensiveness and comparability. However, depending on the sub-sector, some have all the metrics provided whereas others may have missing (NULL) metrics:

- **Emissions metrics:**
    - Emissions quantity E: The mass of emissions emitted for a specific gas or pollutant.
    - Activity A: The amount of sub-sector-specific activity occurring that is producing emissions.
    - Capacity C: The maximum amount of activity possible for the source.
    - Capacity factor κ: The ratio between activity and capacity. In most cases this is the fractional utilization of the total capacity, but in select cases is a unit-ful quantity that also transforms from capacity to activity units.
    - Emissions factor ϵ: The amount of emissions per unit activity for a given gas or pollutant.

- **Robustness metrics:**
    - Confidence: A categorical confidence score associated with each of the above quantities. It can take values “very low”, “low”, “medium”, “high”, or “very high” indicating the robustness of the associated emissions metric. Each sub-sector may have a different approach to classifying estimates into these confidence categories.
    - Uncertainty: A numerical value representing the width of uncertainty around each of the provided emissions metrics. This width may be differently defined depending on the sub-sector: It may be e.g. a 1-sigma credible interval or confidence interval, or a range that brackets the possible true value around the emissions metric.

By definition, each Climate TRACE asset has (or must have) the following emissions metrics that must obey the following equation:

$$E = \epsilon A = \epsilon \kappa C. \tag{1}$$

The time-segments (i.e. the windows of time over which the estimates are reported with start and end times) associated with each of the above quantities may be of varying granularity depending on the sub-sector, but all estimates are given on an annual, quarterly, or monthly segment. The equation above also defines our notion of emissions data being fully, under, or over-constrained: Fully constrained means that a missing variable could be solved for by two other known variables in Equation 1. Underconstrained would mean that one or no variables can solve for a missing value, and overconstrained means that the existing variables do not obey Equation 1.

Table 3 in the Supplemental Materials section summarizes the state of input emissions metrics to our imputation procedures. For each of the emissions metric quantities, we report the percentage of entries that were missing. We also report the percentage of over-constrained values by computing fractional differences between the reported activities and emissions and their expected values given the reported capacities times capacity factors and activities times emissions factors, respectively. For the purposes of identifying highly mismatched metrics, we choose to threshold on 0.05 as a fractional difference that constitutes an over-constraint:

$$\frac{|\kappa C - A|}{A} = \frac{\Delta A}{A} > 0.05 \tag{2}$$

$$\frac{|\epsilon A - E|}{E} = \frac{\Delta E}{E} > 0.05 \tag{3}$$

For example, a mining asset with a capacity of 2 tonnes of ore, a capacity factor 0.5, and an activity of 2 tonnes of ore would be identified as overconstrained, because the product of capacity and capacity factor is not within 5% of the reported activity. Note that this 5% threshold is not used within the data processing and is just to illustrate in Table 3 where there are larger rounding errors.

2.2 Model Development

At a high level, we process each Climate TRACE sub-sector's data separately, with the raw asset-level data flowing through a set of sequential interventions, each of which further cleans or imputes the data. Below, we describe the input data, external datasets that are used, and the set of interventions.

2.2.1 Emissions metrics

We examined the raw data submitted by each sector to assess which general interventions may be needed to fill in any missing emissions metrics and ensure adherence to Equation 1. Through this process, we identified a set of common interventions to apply to all sectors by default. If the data from a sector had no missing values or mathematical inconsistencies, it would pass through each intervention step without any change. We describe these interventions below and then summarize the order in which they are applied in Table 1.

2.2.1.1 Fully Constrained Imputation

The most frequent intervention was for imputation in the fully-constrained case, which we refer to as "NULLtiplication," because NULL (i.e. missing) values were filled in via multiplication (or division) to satisfy Equation 1. In a NULLtiplication step, first any missing capacity, capacity factor, and then activity values were each solved for using the other two values, if the other two were not missing. For example, if A was NULL, we would fill it in with $A=\kappa C$ if κ and C were not missing. If C was missing, we filled it with A/κ, and so on. We then performed the analogous operation on the activity, emissions factor and emissions quantity. Finally, we attempted another solve for the capacity and capacity factor in case the activity was solved via emissions quantity and emissions factor in the previous step.

2.2.1.2 Under-Constrained Imputation

If a value could not be filled in via NULLtiplication, we imputed it using other times and/or assets for the same sector (and the same gas for emissions factors or quantities). Each time an unconstrained imputation operation was performed, a NULLtiplication step followed immediately thereafter to ensure that any missing values that were now fully constrained would be solved for. This also prevented introducing over-constrained entries.

A simple but common case was where emissions quantities or emissions factors were NULL because they were meant to be zero. Through communication with sector leads, we determined which of these NULL values should be zero and collated the list for use in our imputation pipeline. In some cases, entire timestamps were missing, for which we used similar logic to decide whether to set to zero or impute a non-zero value. See Table 2 in Supplemental Materials for details.

If an asset had emissions metrics for some time-segments but not others, we first backward filled the missing values in time, and then attempted forward fill any remaining NULLs, thereby always preferring to fill with the most recent data. These backward/forward fills operated on each type of emissions metric independently, in order of emissions factor, capacity factor, capacity, and activity (emissions quantity was filled in automatically by a NULLtiplication step once activity and emissions factor were calculated).

For assets that were missing some emissions metrics for all reported times and could not be filled in via NULLtiplication and backward/forward fills, we attempted a "regional" imputation. We filled missing emissions factors and capacity factors with the median of those metrics from assets that had data for the same ISO3 country, sector, and time-segment (and gas for emissions factor). We did the same using mean values for capacity and activity, since capacity and activity quantities add to each other whereas the factors do not. We followed the regional imputation with a global analog for any remaining missing values.

As a final stopgap for remaining missing emissions metrics, we pulled from a hand-curated list of emission factors and if needed set capacity factors to 1. The only case requiring the use of a default emission factor was CO_2 in the coal-mining sector. We used the value of 0.0175 tonnes of CO_2 per tonne of coal extracted, following Table 4-5 of Skone (2018).

2.2.1.3 Over-Constrained Imputation

In some cases, the emissions metrics were over-constrained (>5% relative error in the expected multiplicand). These few over-constrained cases were caused by rounding of some very small raw values. We performed two operations to bring these back into alignment with Equation 1. First, if one or more of the metadata metrics was zero but the emissions quantity was non-zero, we turned those metadata metrics with zeros to NULL values and imputed them with the methods described in the previous sections, e.g. using a regional imputation to fill in NULL values. Then, as a very final step after all imputations, we "forced" consistency with Equation 1 by recomputing all emissions factors and capacity factors. This affected only the few asset entries where rounding of very small values in the input data had broken Equation 1.

2.2.1.4 Summary and Order of Default Imputations

In Table 1, we summarize the default imputations and the order we performed them in. We applied these imputations to all sectors and gases, except CO_2 equivalents, which were re-calculated post-hoc after imputation on the constituent gases as described in Section 2.2.1.6.

Table 1 Summary and order of default imputations applied to all sectors and gases

	Imputation
1	NULLtiplication
2	Turn zero-valued metadata into NULLs if emissions quantity was non-zero
3	Turn NULL emissions quantities and emissions factors to 0 for assets that are known to have no emissions
4	Backward/forward fill in time by asset
5	Regional imputation
6	Global imputation
7	Fill NULL emissions factors with default values
8	Set NULL capacity factors = 1, and as a last resort, capacity = 1
9	Force capacity factors and emissions factors to obey Equation 1

2.2.1.5 Sector-Specific Imputations

2.2.1.5.1 Forestry Sectors

Some recent time-segments do not have reported data in the forestry sectors (forest-land-degradation, forest-land-clearing, forest-land-fires, shrubgrass-fires, wetland-fires, and removals) because those data are not yet recorded. All entries for these timestamps are set to zero.

2.2.1.5.2 Mining Sectors

Copper-mining and bauxite-mining sectors also had some assets that could not be imputed from other assets for the same time period. We therefore used regional and global imputations on emissions factors and capacity factors that used data from all time-segments. Additionally, some small activity and capacity values had been rounded in the input data causing some imputed capacity factors to be larger than one, even though capacity factors were defined to be between zero and one in these sectors. Any imputed capacity factors that exceeded one were set to exactly one and Equation 1 was used to recompute the capacity based on the activity.

2.2.1.5.3 "Other" Sectors

Several manufacturing sectors were submitted with only emissions quantities and no associated metadata, as described in Moore et al (2025). For each asset in these sectors (other-chemicals, other-metals, other-manufacturing), the activity and capacity were simply defined as "1 factory." Therefore, the resulting capacity factor was set to 1 in every case, and the emission factor for each gas was the emission quantity per manufacturing facility.

2.2.1.6 CO_2 Equivalents

CO_2-equivalent global warming potentials for 20-year and 100-year periods are provided in the Climate TRACE data as well following IPCC AR6 (Forster et al, 2021). While these equivalents for a given time-segment and asset may have been present in the input data, we recomputed these from scratch after all imputations had been performed on the CO_2 equivalent constituent gases (CO_2, N_2O, CH_4). For CH_4, we use different CO_2-equivalent global warming potentials (GWP) for fossil-fuel-based methane and non-fossil methane. For fossil-fuel methane, we use GWPs of 82.5 and 29.8 for 20-year and 100-year periods, respectively. For non-fossil-fuel methane, we use GWPs of 79.7 and 27 for 20-year and 100-year periods, respectively. For N_2O, we use 273 for both 20-year and 100-year periods.

2.2.1.7 Spatio-Temporal (Dis)aggregation of Estimates

In order for the Climate TRACE inventory to have uniform resolution in time, emissions metrics were resampled to a common time grid as described in Raniga et al (2024). For emissions aggregated to specific spatial regions such as GADM or GHS-FUA boundaries, the metadata were aggregated as well by aggregating the activity and capacity and then solving for an average capacity factor and emissions factor via Equation 1. These spatio-temporal operations were performed after missing input data were imputed with the methods described in the previous sections.

2.2.2 Confidence

We took a similar approach to imputing missing confidences as we did with emissions metrics. An ordered set of operations were performed to fill out confidence scores for every asset, gas, and time. First, we performed the equivalent of a NULLtiplication operation but on confidences. If confidences existed for two emissions metrics, but a confidence was missing for a third one that could be constrained by the first two, we filled the missing confidence with the lower confidence of the two existing estimates. For example, if an emissions quantity was missing a confidence, it was filled with the lower of the confidences of the emissions factor and activity. A similar operation was performed for CO_2 equivalent confidences: the lower of the confidences of the constituent gases was used for imputation.

In cases where missing confidences were unconstrained by other confidence values, we filled them with the lowest confidence for that sector, gas, and value. Next, there were some cases where confidences were missing for emissions factors and emissions quantities because that gas was not expected to be emitted in that sector. In those cases, we set the confidence to "very high," with the corresponding emissions metrics being zero. Note that many of these very high estimates were downgraded to "high" from temporal downscaling described below. The final missing-confidence imputation operation was to fill any confidences that were still NULL with "very low."

After the NULL confidences were filled, we re-sampled all the confidences to the time granularity of the final emissions metric data (monthly). Any confidences that were not already originally at monthly (and therefore were coarser estimates in time) were downgraded by one confidence level indicating the uncertainty associated with temporal disaggregation. Estimates for any times that were missing coverage in the input data entirely were assigned "very low" confidence.

2.2.3 Uncertainty

A single, generic approach was taken to impute missing uncertainty information and standardize the information to the common temporal granularity, regardless of the sector and gas. Uncertainty values provided in the input data from coalition members were roughly treated as the standard deviation of a Gaussian distribution around the corresponding emissions metric estimate. The emissions metrics were treated as independent variables for simplicity, although Equation 1 does imply correlations among different quantities (Note that large uncertainties compared to their associated emissions metrics were not interpreted to imply negative emissions, and instead interpreted as a distribution that truncated at zero).

For assets that had uncertainties for some times but not others, we imputed each asset's missing uncertainties using the median uncertainties for each emissions metric for that asset over all times. If uncertainties existed for some but not all assets in a given sector, the median relative variance of all assets/times for a given gas was applied to any missing assets/times. If uncertainties did not exist for a given sector, the median relative standard deviation reported over all assets and all times for a given gas was applied to the missing values.

To resample the provided uncertainties to the common temporal granularity, we first compared the emissions metrics between the (cleaned and imputed) data with and without monthly resampling. If the emissions metrics were the same, the provided uncertainty was left unchanged. If instead a metric changed by some factor f through the monthly resampling, the variance (square of the standard deviation) was scaled by f as well. This operation preserved the input variance as being the sum of variances of estimates at finer temporal granularities.

If no uncertainty value σ_{orig} was provided in the input data for emissions quantity, activity, or capacity, the imputed uncertainty σ_{imp} was calculated using

$$\sigma_{imp}^{2} / M_{imp} = median(\sigma_{orig}^{2} / M_{orig}), \tag{4}$$

where M_{imp} and M_{orig} were the imputed and original emissions metrics. The median was taken over either all times for the same asset, or if that data was not available, all assets in the sector. Emissions-factor and capacity-factor uncertainties were imputed similarly, with one key difference: if uncertainties did not exist in the input data, we took the smallest of the following possible values:

- Equation 4 with the median over all times for a given asset and gas
- Equation 4 with the median over all assets for a given sector and gas
- The larger of:
    - the standard deviation of M_{orig} in that sector, gas, and country, or
    - 5% of M_{imp}
- 50% of M_{imp}

This scheme automatically probed cases where a very specific emissions factor or capacity factor was used in a sector and country, suggesting it should have low (but not lower than 5%) uncertainty, while also controlling the uncertainty from becoming excessively large.

If the requisite data for σ_{orig} or M_{orig} did not exist for any of the above calculations, a blanket 50% uncertainty was applied. A 50% uncertainty was also applied to all metrics in the scraped-asset sectors described in Moore et al (2024), but capacity factor was given an uncertainty of 0.1. Capacity factor for lime, however, was kept as a 50% uncertainty. The choices herein were made so that uncertainties could be simply imputed across all sectors using the available information, while not introducing uncertainties that were unreasonably small or large.

3. Results

After performing the imputations from Section 2 on the input data, all missing and over-constrained values are eliminated. Table 2 shows the effect of our imputations on global emissions quantities. Only sectors where the absolute value of the percent change is at least 0.1 are shown. The mining sectors showed the largest increase from imputations because many assets' emissions quantities were missing in that sector. All other sectors' totals for 2023 changed less than 10%. The apparent negative changes shown in Table 2 are artifacts from our use of different global warming potentials as described in Section 2.2.1.6, not the underlying constituent gases. While imputing missing values had a small effect on emissions quantities overall, it nonetheless produced a complete dataset, giving users of the data an estimate for every quantity for each asset and time reported.

Table 2 Absolute difference and percentage change of total 2024 CO_2 equivalent (100 year global warming potential) emissions in millions of tonnes (Mt) for reported assets between the input data and data after imputation for cases where change is > 1% and CO_2 equivalent is initially reported.

sector	difference (Mt)	% change
copper-mining	40.7	74.8
iron-mining	19.7	52.2
bauxite-mining	2.1	51.3
oil-and-gas-refining	192.1	24.5
coal-mining	291.7	15.8
iron-and-steel	175.6	7.7
aluminum	18.5	4.6

4. Discussion

The imputations we perform to achieve completeness of the input data make only a modest increase to total emissions in the inventory. However, a considerable number of metadata are imputed as shown in Tables 3 and 5. Many of the missing emissions metrics are immediately filled through NULLtiplication or through setting missing values to zero which are known to be a-priori. These imputations will enable stakeholders to easily investigate future changes in activity and emissions factors affected by pollution control measures that mitigate asset-level emissions and achieve climate goals.

In terms of the confidences, many emissions quantity and emissions factor confidences are filled with “very high” because those values are known to be zero. Otherwise, most missing confidences are filled with “very low” to be conservative. Future work could refine these confidences further.

The uncertainty widths we imputed and re-sampled are a starting point for assessing the numerical uncertainty of Climate TRACE asset data and provide some numerical estimate to accompany confidences. However, we note that the input uncertainties may not be comparable between sectors and the full multivariate distributions of emissions metrics are not modeled here. Therefore, future work is needed to provide more robust uncertainty metrics.

Temporal disaggregation of both the metadata and the emissions estimates makes the Climate TRACE inventory both complete and also high-resolution in time. The confidences and uncertainties reflect the extra operations performed on the data for imputation and standardization.

Overall, the scheme outlined in this methodology has the advantage of being simple and sector-agnostic, making any changes to the data transparent and reproducible. However, the sequential nature of the imputations does not account for correlations among different variables. We may explore multivariate techniques in future work.

5. Conclusion

In sum, we standardized the Climate TRACE dataset as reported directly from coalition members by filling missing data and re-sampling to a common monthly time resolution. We used a set of simple, sequential operations to achieve completeness of the data over time and over assets. The effect on total emissions quantities was little to none in most sectors.

6. Acknowledgements

We would like to acknowledge Dr. Aaron Davitt for helpful comments on this manuscript.

7. Supplementary materials

Table 3 Climate TRACE sub-sectors by gas breakdown for the following: Percentage of emissions metrics with missing values for columns Emissions, E; Emissions factor, ϵ; Activity, A; Capacity factor, κ; and Capacity, C. Percentage of activity inconsistencies ($\Delta A/A > 0.05$) and emissions inconsistencies ($\Delta E/E > 0.05$), and total count (#) of assets per sector and gas. Note that "oil-and-gas-production" and "oil-and-gas-transport" metadata are proprietary information that we do not obtain nor impute at this time.

sub-sector	gas	E	ϵ	A	κ	C	$\Delta A/A >$ 0.05	$\Delta E/E >$ 0.05	#
bauxite-mining	N_2O	70.8	70.8	26.3	42.4	53.2	0.1	0.0	1370
bauxite-mining	CO_2	26.3	26.3	26.3	42.4	53.2	0.1	0.4	1370
bauxite-mining	CH_4	70.8	70.8	26.3	42.4	53.2	0.1	0.0	1370
cement	N_2O	100.0	100.0	0.0	0.0	0.0	0.0	0.0	117942
cement	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	117942
cement	CH_4	100.0	100.0	0.0	0.0	0.0	0.0	0.0	117942
coal-mining	N_2O	0.0	0.0	1.7	1.7	1.7	0.0	0.0	43630
coal-mining	CO_2	100.0	100.0	1.7	1.7	1.7	0.0	0.0	43630
coal-mining	CH_4	2.2	0.5	1.7	1.7	1.7	0.0	0.0	43630
copper-mining	N_2O	66.7	66.8	21.8	38.2	67.4	0.0	0.0	4570
copper-mining	CO_2	21.8	21.9	21.8	38.2	67.4	0.0	0.0	4570
copper-mining	CH_4	66.7	66.8	21.8	38.2	67.4	0.0	0.0	4570
cropland-fires	N_2O	0.0	100.0	100.0	100.0	100.0	0.0	0.0	516194
cropland-fires	CO_2	0.0	100.0	100.0	100.0	100.0	0.0	0.0	516194
cropland-fires	CH_4	0.0	100.0	100.0	100.0	100.0	0.0	0.0	516194
domestic-aviation	N_2O	0.0	24.9	0.0	24.9	0.0	0.0	0.0	619904
domestic-aviation	CO_2	0.0	24.9	0.0	24.9	0.0	0.0	0.0	619904
domestic-aviation	CH_4	0.0	24.9	0.0	24.9	0.0	0.0	0.0	619904
electricity-generation	N_2O	100.0	100.0	0.0	0.0	0.0	0.0	0.0	758560
electricity-generation	CO_2	0.0	1.2	0.0	0.0	0.0	0.0	0.0	758560
electricity-generation	CH_4	100.0	100.0	0.0	0.0	0.0	0.0	0.0	758560
international-aviation	N_2O	0.0	43.2	0.0	43.2	0.0	0.0	0.0	281600
international-aviation	CO_2	0.0	43.2	0.0	43.2	0.0	0.0	0.0	281600
international-aviation	CH_4	0.0	43.2	0.0	43.2	0.0	0.0	0.0	281600
iron-mining	N_2O	76.7	76.7	43.8	64.2	84.1	0.0	0.0	3770

sub-sector	gas	*E*	ϵ	*A*	κ	*C*	Δ*A*/*A* > 0.05	Δ*E*/*E* > 0.05	#
iron-mining	CO_2	43.8	43.8	43.8	64.2	84.1	0.0	0.0	3770
iron-mining	CH_4	76.7	76.7	43.8	64.2	84.1	0.0	0.0	3770
oil-and-gas-refining	N_2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8884
oil-and-gas-refining	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8884
oil-and-gas-refining	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8884
rice-cultivation	N_2O	19.6	0.0	1.1	1.1	1.1	0.0	0.0	260097
rice-cultivation	CO_2	19.6	0.0	1.1	1.1	1.1	0.0	0.0	260097
rice-cultivation	CH_4	1.1	1.1	1.1	1.1	1.1	0.0	0.2	260097
road-transportation	N_2O	0.0	0.5	0.0	0.5	0.0	0.0	0.0	3354792
road-transportation	CO_2	0.0	0.5	0.0	0.5	0.0	0.0	0.0	3354792
road-transportation	CH_4	0.0	0.5	0.0	0.5	0.0	0.0	0.0	3354792
solid-waste-disposal	N_2O	100.0	100.0	0.0	51.7	51.7	0.0	0.0	774312
solid-waste-disposal	CO_2	100.0	100.0	0.0	51.7	51.7	0.0	0.0	774312
solid-waste-disposal	CH_4	0.1	0.1	0.0	51.7	51.7	0.0	0.0	774312
synthetic-fertilizer-application	N_2O	0.0	11.3	0.0	11.3	0.0	0.0	0.0	3368960
synthetic-fertilizer-application	CO_2	0.0	11.3	0.0	11.3	0.0	0.0	0.0	3368960
synthetic-fertilizer-application	CH_4	0.0	11.3	0.0	11.3	0.0	0.0	0.0	3368960
wetland-fires	N_2O	100.0	100.0	57.6	57.5	91.2	0.0	0.0	983005
wetland-fires	CO_2	0.1	91.2	57.6	57.5	91.2	0.0	0.0	983005
wetland-fires	CH_4	100.0	100.0	57.6	57.5	91.2	0.0	0.0	983005
aluminum	N_2O	100.0	100.0	0.0	0.0	0.0	0.0	0.0	13844
aluminum	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	13844
aluminum	CH_4	100.0	100.0	0.0	0.0	0.0	0.0	0.0	13844
forest-land-clearing	N_2O	100.0	100.0	55.2	55.1	63.6	0.0	0.0	1025735
forest-land-clearing	CO_2	0.7	63.6	55.2	55.1	63.6	0.0	0.0	1025735
forest-land-clearing	CH_4	100.0	100.0	55.2	55.1	63.6	0.0	0.0	1025735
pulp-and-paper	N_2O	100.0	100.0	0.0	0.0	0.0	0.0	0.0	16824
pulp-and-paper	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16824
pulp-and-paper	CH_4	100.0	100.0	0.0	0.0	0.0	0.0	0.0	16824
chemicals	N_2O	100.0	100.0	0.0	0.0	0.0	0.0	0.0	20796
chemicals	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20796
chemicals	CH_4	100.0	100.0	0.0	0.0	0.0	0.0	0.0	20796

sub-sector	gas	*E*	ϵ	*A*	κ	*C*	ΔA/A > 0.05	ΔE/E > 0.05	#
forest-land-degradation	N_2O	100.0	100.0	57.1	57.0	83.3	0.0	0.0	992985
forest-land-degradation	CO_2	0.7	83.3	57.1	57.0	83.3	0.0	0.0	992985
forest-land-degradation	CH_4	100.0	100.0	57.1	57.0	83.3	0.0	0.0	992985
other-manufacturing	N_2O	36.2	100.0	100.0	100.0	100.0	0.0	0.0	3582
other-manufacturing	CO_2	36.2	100.0	100.0	100.0	100.0	0.0	0.0	3582
other-manufacturing	CH_4	36.2	100.0	100.0	100.0	100.0	0.0	0.0	3582
forest-land-fires	N_2O	100.0	100.0	56.6	56.6	77.3	0.0	0.0	998973
forest-land-fires	CO_2	0.1	77.3	56.6	56.6	77.3	0.0	0.0	998973
forest-land-fires	CH_4	100.0	100.0	56.6	56.6	77.3	0.0	0.0	998973
domestic-shipping	N_2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1110078
domestic-shipping	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1110078
domestic-shipping	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1110078
international-shipping	N_2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1016215
international-shipping	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1016215
international-shipping	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1016215
net-forest-land	N_2O	100.0	100.0	0.2	0.0	2.1	0.0	0.0	529207
net-forest-land	CO_2	0.2	3.7	0.2	0.0	2.1	0.0	0.0	529207
net-forest-land	CH_4	100.0	100.0	0.2	0.0	2.1	0.0	0.0	529207
net-shrubgrass	N_2O	100.0	100.0	0.1	0.0	2.4	0.0	0.0	530959
net-shrubgrass	CO_2	0.1	4.2	0.1	0.0	2.4	0.0	0.0	530959
net-shrubgrass	CH_4	100.0	100.0	0.1	0.0	2.4	0.0	0.0	530959
net-wetland	N_2O	100.0	100.0	0.3	0.1	17.6	0.0	0.0	516311
net-wetland	CO_2	0.3	19.7	0.3	0.1	17.6	0.0	0.0	516311
net-wetland	CH_4	100.0	100.0	0.3	0.1	17.6	0.0	0.0	516311
removals	N_2O	100.0	100.0	0.0	0.0	2.3	0.0	0.0	541239
removals	CO_2	0.1	4.1	0.0	0.0	2.3	0.0	0.0	541239
removals	CH_4	100.0	100.0	0.0	0.0	2.3	0.0	0.0	541239
shrubgrass-fires	N_2O	100.0	100.0	56.8	56.7	79.8	0.0	0.0	997389
shrubgrass-fires	CO_2	0.1	79.8	56.8	56.7	79.8	0.0	0.0	997389
shrubgrass-fires	CH_4	100.0	100.0	56.8	56.7	79.8	0.0	0.0	997389
water-reservoirs	N_2O	100.0	100.0	0.0	0.0	0.0	0.0	0.0	900136
water-reservoirs	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	900136

sub-sector	gas	E	ϵ	A	κ	C	$\Delta A/A > 0.05$	$\Delta E/E > 0.05$	#
water-reservoirs	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	900136
enteric-fermentation-cattle-pasture	N_2O	0.0	11.2	0.0	11.2	0.0	0.0	0.0	601600
enteric-fermentation-cattle-pasture	CO_2	0.0	11.2	0.0	11.2	0.0	0.0	0.0	601600
enteric-fermentation-cattle-pasture	CH_4	0.0	11.2	0.0	11.2	0.0	0.0	0.0	601600
manure-left-on-pasture-cattle	N_2O	0.0	11.2	0.0	11.2	0.0	0.0	0.0	601600
manure-left-on-pasture-cattle	CO_2	0.0	11.2	0.0	11.2	0.0	0.0	0.0	601600
manure-left-on-pasture-cattle	CH_4	0.0	11.2	0.0	11.2	0.0	0.0	0.0	601600
residential-onsite-fuel-usage	N_2O	0.0	0.1	0.0	0.1	0.0	0.0	0.0	3068160
residential-onsite-fuel-usage	CO_2	0.0	0.1	0.0	0.1	0.0	0.0	0.0	3068160
residential-onsite-fuel-usage	CH_4	0.0	0.1	0.0	0.1	0.0	0.0	0.0	3068160
non-residential-onsite-fuel-usage	N_2O	0.0	13.0	0.0	13.0	0.0	0.0	0.0	3068160
non-residential-onsite-fuel-usage	CO_2	0.0	13.0	0.0	13.0	0.0	0.0	0.0	3068160
non-residential-onsite-fuel-usage	CH_4	0.0	13.0	0.0	13.0	0.0	0.0	0.0	3068160
oil-and-gas-production	N_2O	100.0	100.0	100.0	100.0	100.0	0.0	0.0	80524
oil-and-gas-production	CO_2	0.0	100.0	100.0	100.0	100.0	0.0	0.0	80524
oil-and-gas-production	CH_4	0.0	100.0	100.0	100.0	100.0	0.0	0.0	80524
oil-and-gas-transport	N_2O	100.0	100.0	100.0	100.0	100.0	0.0	0.0	80524
oil-and-gas-transport	CO_2	0.0	100.0	100.0	100.0	100.0	0.0	0.0	80524
oil-and-gas-transport	CH_4	0.0	100.0	100.0	100.0	100.0	0.0	0.0	80524
manure-applied-to-soils	N_2O	0.0	12.2	0.0	11.5	0.1	0.0	0.0	2887680
manure-applied-to-soils	CO_2	0.0	12.2	0.0	11.5	0.1	0.0	0.0	2887680
manure-applied-to-soils	CH_4	0.0	12.2	0.0	11.5	0.1	0.0	0.0	2887680
crop-residues	N_2O	0.0	13.4	0.0	11.5	4.3	0.0	0.0	2887680
crop-residues	CO_2	0.0	13.4	0.0	11.5	4.3	0.0	0.0	2887680
crop-residues	CH_4	0.0	13.4	0.0	11.5	4.3	0.0	0.0	2887680
Domestic-wastewater-treatment-and-discharge	N_2O	0.0	4.8	0.0	0.0	0.0	0.0	0.0	8754771
Domestic-wastewater-treatment-and-discharge	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8754771
Domestic-wastewater-treatment-and-discharge	CH_4	0.0	4.8	0.0	0.0	0.0	0.0	0.0	8754771
Industrial-wastewater-treatment-and-discharge	N_2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	163328
Industrial-wastewater-treatment-and-discharge	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	163328

sub-sector	gas	E	ϵ	A	κ	C	$\Delta A/A > 0.05$	$\Delta E/E > 0.05$	#
Industrial-wastewater-treatment-and-discharge	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	163328
petrochemical-steam-cracking	N_2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3094
petrochemical-steam-cracking	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3094
petrochemical-steam-cracking	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3094
iron-and-steel	N_2O	100.0	100.0	0.0	0.0	0.0	0.0	0.0	48792
iron-and-steel	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	48792
iron-and-steel	CH_4	100.0	100.0	0.0	0.0	0.0	0.0	0.0	48792
glass	N_2O	37.6	37.6	0.0	0.0	0.0	0.0	0.0	4338
glass	CO_2	0.4	0.4	0.0	0.0	0.0	0.0	0.0	4338
glass	CH_4	49.8	49.9	0.0	0.0	0.0	0.0	0.0	4338
lime	N_2O	91.1	91.1	0.4	0.0	0.4	0.0	0.0	14724
lime	CO_2	0.1	0.5	0.4	0.0	0.4	0.0	0.0	14724
lime	CH_4	91.1	91.1	0.4	0.0	0.4	0.0	0.0	14724
food-beverage-tobacco	N_2O	0.4	0.5	0.1	0.0	0.1	0.0	0.0	135423
food-beverage-tobacco	CO_2	0.6	0.6	0.1	0.0	0.1	0.0	0.0	135423
food-beverage-tobacco	CH_4	0.4	0.5	0.1	0.0	0.1	0.0	0.0	135423
other-metals	N_2O	40.9	100.0	100.0	100.0	100.0	0.0	0.0	3762
other-metals	CO_2	40.9	100.0	100.0	100.0	100.0	0.0	0.0	3762
other-metals	CH_4	40.9	100.0	100.0	100.0	100.0	0.0	0.0	3762
other-chemicals	N_2O	33.5	100.0	100.0	100.0	100.0	0.0	0.0	3951
other-chemicals	CO_2	33.5	100.0	100.0	100.0	100.0	0.0	0.0	3951
other-chemicals	CH_4	33.5	100.0	100.0	100.0	100.0	0.0	0.0	3951
textiles-leather-apparel	N_2O	32.5	32.5	8.4	0.0	8.4	0.0	0.0	96624
textiles-leather-apparel	CO_2	32.5	32.5	8.4	0.0	8.4	0.0	0.0	96624
textiles-leather-apparel	CH_4	32.5	32.5	8.4	0.0	8.4	0.0	0.0	96624
enteric-fermentation-cattle-operation	N_2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4523248
enteric-fermentation-cattle-operation	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4523248
enteric-fermentation-cattle-operation	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4523248
manure-management-cattle-operation	N_2O	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4523248
manure-management-cattle-operation	CO_2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4523248
manure-management-cattle-operation	CH_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4523248

Table 4 Mapping of which NULL emissions quantities and emissions factors (by gas) and activities and capacities should be interpreted as zeros, defined as “TRUE”. Where there is a “FALSE”, NULL emissions quantities and emissions factors are not necessarily to be interpreted as zero. The “missing” columns refer to similar handling but applied to cases where timestamps are missing entirely.

sub-sector	CH_4 -> 0	CO_2 -> 0	N_2O -> 0	NULL A -> 0	NULL C -> 0	Missing A -> 0	Missing C -> 0
aluminum	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
bauxite-mining	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
cement	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
chemicals	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
coal-mining	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
copper-mining	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
cropland-fires	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
crop-residues	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
domestic-aviation	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE
domestic-shipping	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE
domestic-wastewater-treatment-and-discharge	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
electricity-generation	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	FALSE
enteric-fermentation-cattle-operation	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE
enteric-fermentation-cattle-pasture	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE
enteric-fermentation-other	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE
food-beverage-tobacco	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
forest-land-clearing	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
forest-land-degradation	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
forest-land-fires	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
glass	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
heat-plants	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
industrial-wastewater-treatment-and-discharge	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
international-aviation	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE
international-shipping	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE

sub-sector	CH_4 -> 0	CO_2 -> 0	N_2O -> 0	NULL A -> 0	NULL C -> 0	Missing A -> 0	Missing C -> 0
iron-and-steel	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE
iron-mining	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
lime	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
manure-applied-to-soils	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
manure-left-on-pasture-cattle	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
manure-management-cattle-operation	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
manure-management-other	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
net-forest-land	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
net-shrubgrass	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
net-wetland	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
non-broadcasting-vessels	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE	FALSE
non-residential-onsite-fuel-usage	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
oil-and-gas-production	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
oil-and-gas-refining	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
oil-and-gas-transport	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-chemicals	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-energy-use	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-fossil-fuel-operations	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-manufacturing	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-metals	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-mining-quarrying	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-onsite-fuel-usage	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
other-solid-fuels	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
petrochemical-steam-cracking	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
pulp-and-paper	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
removals	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
residential-onsite-fuel-usage	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
rice-cultivation	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE
road-transportation	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
shrubgrass-fires	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE

sub-sector	CH_4 -> 0	CO_2 -> 0	N_2O -> 0	NULL A -> 0	NULL C -> 0	Missing A -> 0	Missing C -> 0
soil-organic-carbon	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
solid-waste-disposal	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
synthetic-fertilizer-application	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE
textiles-leather-apparel	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
water-reservoirs	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE
wetland-fires	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE

Table 5 Percentage of NULL confidence values for each data field by sub-sector and gas, and the total row count (#). We do not include cases here where there are no initial confidences whatsoever, or where all percentages are 0.

sub-sector	gas	C	κ	A	ϵ	E	#
bauxite-mining	N_2O	78	57	22	100	100	37
bauxite-mining	CO_2	78	57	22	22	22	37
bauxite-mining	CH_4	78	57	22	100	100	37
cement	N_2O	0	0	0	100	100	120102
cement	CH_4	0	0	0	100	100	120102
coal-mining	N_2O	0	0	0	100	100	43630
coal-mining	CO_2	0	0	0	100	100	43630
copper-mining	N_2O	68	69	22	100	100	914
copper-mining	CO_2	68	69	22	22	22	914
copper-mining	CH_4	68	69	22	100	100	914
cropland-fires	N_2O	100	100	31	31	31	213
cropland-fires	CO_2	100	100	31	100	100	213
cropland-fires	ch4	100	100	31	31	31	213
electricity-generation	N_2O	0	0	0	100	100	755840
electricity-generation	CH_4	0	0	0	100	100	755840
iron-mining	N_2O	84	64	44	100	100	753
iron-mining	CO_2	84	64	44	44	44	753
iron-mining	CH_4	84	64	44	100	100	753
rice-cultivation	N_2O	100	100	1	100	100	80081
rice-cultivation	CO_2	100	100	1	100	100	80081

sub-sector	gas	*C*	κ	*A*	ϵ	*E*	#
rice-cultivation	CH_4	100	100	1	1	1	80081
synthetic-fertilizer-application	N_2O	3	100	3	3	3	3728378
synthetic-fertilizer-application	CO_2	3	100	3	100	100	3728378
synthetic-fertilizer-application	CH_4	3	100	3	100	100	3728378
wetland-fires	N_2O	89	89	89	100	100	775936
wetland-fires	CO_2	89	89	89	100	48	775936
wetland-fires	CH_4	89	89	89	100	100	775936
aluminum	N_2O	0	0	0	100	100	28648
aluminum	CH_4	0	0	0	100	100	28648
forest-land-clearing	N_2O	53	53	53	100	100	775936
forest-land-clearing	CO_2	53	53	53	100	12	775936
forest-land-clearing	CH_4	53	53	53	100	100	775936
pulp-and-paper	N_2O	0	0	0	100	100	16824
pulp-and-paper	ch4	0	0	0	100	100	16824
chemicals	N_2O	0	0	0	100	100	51516
chemicals	CH_4	0	0	0	100	100	51516
forest-land-degradation	N_2O	78	78	78	100	100	775936
forest-land-degradation	CO_2	78	78	78	100	38	775936
forest-land-degradation	CH_4	78	78	78	100	100	775936
other-manufacturing	N_2O	100	100	100	100	0	3184
other-manufacturing	CO_2	100	100	100	100	0	3184
other-manufacturing	CH_4	100	100	100	100	0	3184
forest-land-fires	N_2O	72	72	72	100	100	775936
forest-land-fires	CO_2	72	72	72	100	31	775936
forest-land-fires	CH_4	72	72	72	100	100	775936
net-forest-land	N_2O	2	2	2	100	100	460143
net-forest-land	CO_2	2	2	2	100	2	460143
net-forest-land	CH_4	2	2	2	100	100	460143
net-shrubgrass	N_2O	3	3	3	100	100	460143
net-shrubgrass	CO_2	3	3	3	100	3	460143
net-shrubgrass	CH_4	3	3	3	100	100	460143

sub-sector	gas	C	κ	A	ϵ	E	#
net-wetland	N_2O	19	19	19	100	100	460143
net-wetland	CO_2	19	19	19	100	19	460143
net-wetland	CH_4	19	19	19	100	100	460143
removals	N_2O	89	89	89	100	100	460143
removals	CO_2	89	89	89	100	89	460143
removals	CH_4	89	89	89	100	100	460143
shrubgrass-fires	N_2O	75	75	75	100	100	775936
shrubgrass-fires	CO_2	75	75	75	100	34	775936
shrubgrass-fires	CH_4	75	75	75	100	100	775936
water-reservoirs	N_2O	0	0	0	100	100	907386
oil-and-gas-production	N_2O	100	100	100	100	100	79588
oil-and-gas-production	CO_2	100	100	100	100	0	79588
oil-and-gas-production	CH_4	100	100	100	100	0	79588
oil-and-gas-transport	N_2O	100	100	100	100	100	79552
oil-and-gas-transport	CO_2	100	100	100	100	0	79552
oil-and-gas-transport	CH_4	100	100	100	100	0	79552
manure-applied-to-soils	N_2O	0	100	0	0	0	601780
manure-applied-to-soils	CO_2	0	100	0	100	100	601780
manure-applied-to-soils	CH_4	0	100	0	100	100	601780
crop-residues	N_2O	0	100	0	0	0	601530
crop-residues	CO_2	0	100	0	100	100	601530
crop-residues	CH_4	0	100	0	100	100	601530
iron-and-steel	N_2O	0	0	0	100	100	49667
iron-and-steel	CO_2	0	0	0	100	100	49667

Data citation format: Moore, D., Doctor, Z, Raniga, K., Lewis, C., Sridhar, L., Thomas, P., Saraswat, I., Betz, L., Piscopo, A., Collins, G., Nellis, A., Brown, N., Pekala, M., Rokisky, J., Reilly, E., Hughes, M. and McCormick, G. (2025). *Completeness of Bottom-up Emissions Estimates and Associated Metadata.* WattTime and Johns Hopkins University Applied Physics Laboratory, 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. Forster, P., et al. (2021). *The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity*. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 923–1054, doi: 10.1017/9781009157896.009.
2. GADM (2022). *GADM Maps and Data* [Online]. Available at: gadm.org (Accessed: 8 November 2024)
3. Moore, D., et al (2024). *Bottom-up Estimation of Emissions in Manufacturing Sub-sectors* [Online]. Available at: https://github.com/climatetracecoalition/methodology-documents/tree/main/2024/Post%20Processing%20for%20Global%20Emissions%20and%20Metadata%20Completeness (Accessed: 15 November 2024)

4. Raniga, K., et al (2024). *Temporal Disaggregation of Emissions Data for the Climate TRACE Inventory* [Online]. Available at: https://github.com/climatetracecoalition/methodology-documents/tree/main/2024/Post%20Processing%20for%20Global%20Emissions%20and%20Metadata%20Completeness (Accessed: 15 November 2024)
5. Schiavina, M., Moreno-Monroy, A., Maffenini, L., Veneri, P. (2019): GHS-FUA R2019A - GHS functional urban areas, derived from GHS-UCDB R2019A (2015). European Commission, Joint Research Centre (JRC) [Dataset] doi: 10.2905/347F0337-F2DA-4592-87B3-E25975EC2C95 PID: http://data.europa.eu/89h/347f0337-f2da-4592-87b3-e25975ec2c95
6. Skone, T.J., Schivley, G., Jamieson, M., Marriott, J., Cooney, G., Littlefield, J., Mutchek, M., Krynock, M., and Shih, C.Y. (2018). *Life Cycle Analysis: Sub-Critical Pulverized Coal (SubPC) Power Plants* [Online]. Office of Scientific and Technical Information. United States. Available at: https://doi.org/10.2172/1542447 (Accessed: 11 November 2024)
7. United States Environmental Protection Agency (2024). *Basic Information of Air Emissions Factors and Quantification* [Online]. Available at: https://www.epa.gov/air-emissions-factors-and-quantification/basic-information-air-emissions-factors-and-quantification (Accessed: 8 November 2024)