Manufacturing sector: Steel Manufacturing Emissions

Andrew Isabriye and Ashank Sinha
Authors affiliated with TransitionZero and Climate TRACE



1. Introduction

Steel production is both energy and emissions intensive, representing about 8% of total energy demand and 7% of global energy sector carbon emissions (IEA, 2020). The emissions intensity of steel production is due to its reliance on coal. At the same time, steel is vitally important to the global economy, used in buildings, infrastructure, weapons, vehicles and furniture, for example. Due to its importance to modern society, steel demand is expected to grow for the foreseeable future (IEA, 2020). Steel production is also relatively centralised, with China accounting for over half of global production, followed by the European Union (EU) and United Kingdom (UK) making up 9%, India 6%, Japan 5%, the US 5%, Russia 4% and South Korea 4% (IEA, 2020).

Steelmaking has two main metallic inputs: iron ore and recycled steel scrap. Around 70% of the total metallic input to steel production globally is derived from iron ore, with scrap making up the rest. Primary steel is produced using blast furnaces, with this route accounting for the majority of this sector's emissions. Secondary (or scrap-based) production is carried out in electric furnaces and is around one-eighth as energy-intensive as production from iron ore, using electricity – as opposed to coal – as the main energy input (IEA, 2020).

Most importantly, steel production can be used as a proxy for emissions (i.e., kg CO₂/ton of steel) and therefore it is desirable to ascertain real time and in-depth production values to guide future climate policy. With any such globally traded commodity however, it is characterised by fierce competition amongst producers and, as a consequence, facility level production data is rarely made publicly available. In most cases it is only possible to obtain steel production/emissions quantities at the country level, and often there is a substantial delay (~years) in obtaining the data (E-PRTR contributors, 2022; UNFCCC contributors, 2022; US EPA contributors, 2022).

In this work we look to address both the temporal and spatial limitations in traditional reporting of steel production/emissions to provide more timely and accurate facility level data. More specifically, we have aimed to deliver steel production and emissions estimates on a monthly basis (with 1 month lag) for all sources identified on the Global Energy Monitor Global Steel Plant Tracker database (GEM contributors, 2022) (see section 2.1.1). Our approach improves upon certain capacity-based approaches which may only proportionately infer production between facilities or assign some average utilisation rates of installed capacities. The approach achieves this

using satellite-derived mapping data which can capture variation in activity for certain types of steel facilities - and use this data to infer steel production in a given time frame.

2. Materials and Methods

During steel production, direct emissions are released during the manufacturing of crude steel, with estimates varying by different production routes, namely blast furnace-basic oxygenation furnace (BF-BOF) and electric arc furnace (EAF). A simplified overview of the steel production value chain from raw materials to final steel products is available in Appendix 1, section 8.2.

We investigate each major processing route independently using a combination of satellite imagery, publicly reported data, and academic papers. This section provides a high-level overview of the datasets and associated pipelines used to derive emission estimates for the iron and steel sub-sector.

Given the lack of source-level emission data publicly available, a standardised "bottom-up" approach is used to quantify the emissions. This process is characterised by first estimating production levels for each plant (in tons of crude steel produced for a steel plant, for instance) before then subsequently applying a calculated emissions factor (tons of CO₂ per ton of crude steel produced) to estimate emissions.

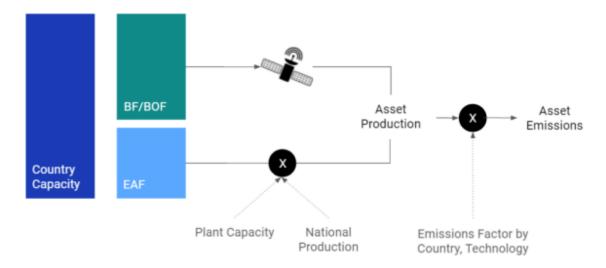


Figure 1 Example of the modelling approach to quantify source-level emissions for the Iron & Steel industry, segregated by steel processing route.

There are primarily two approaches to estimating source level production. As a priority, satellite-derived production estimates are used whenever a facility releases enough heat to be captured by satellite imagery (Zhou *et al.*, 2018; Marchese *et al.*, 2019; Liangrocapart, Khetkeeree

and Petchthaweetham, 2020). This is the case for BF-BOF facilities which have several units that function at temperatures higher than 1,200°C. These hotspots include signals from coke plants, sinter plants, BFs, slag pits, and BOFs (Figure 3). For all the other plants where hotspots were not captured by satellite imagery (EAFs for instance), we used a basic disaggregation method. This is determined by calculating each plant's share of national capacity, before multiplying this number by the country's production to derive the plant's contribution for the given timeframe. For country-level statistics these point-source estimates are then aggregated. Figures 1 and 2 summarises the modelling approach, however, M'Barek and Gray (2021) provide a more detailed methodological breakdown.

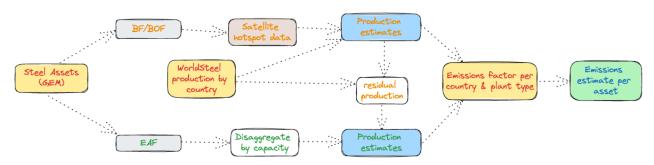


Figure 2 Flowchart of the methodology to calculate plant level emissions for the steel plants.

2.1. Materials

2.1.1. Steel inventory dataset

Global Energy Monitor (GEM) provided facility level information such as GPS coordinates, owner, capacity, age, product type and technology type via the Global Steel Plant Tracker that includes every iron and steel plant currently operating with a capacity of more than 0.5 million tonnes. The GSPT also includes all plants meeting the capacity threshold that have been proposed or under construction since 2017 and retired or mothballed since 2020. In total there are 1,016 unique steel facilities with an annual production capacity of 3.1 billion tonnes across 89 countries. Since we use remote sensing data, we need facility geolocation data accurate to within a few tens of metres. While some of the listed geolocations from the GEM dataset provides this information, it is in some cases, approximative. We supplement the source geolocation data using the Google Maps API (Google Maps, 2022) and OpenStreetMap (OpenStreetMap, 2022), before manually validating all geolocations. Figure 3 shows how steel sources are distributed globally with high concentration of sources in China followed more distantly by India, the US and Japan (facility count details in section 8.5). The map also highlights countries with little to no steel production, mainly concentrated in Sub-Saharan Africa, Latin America and Southeast Asia.-monitored

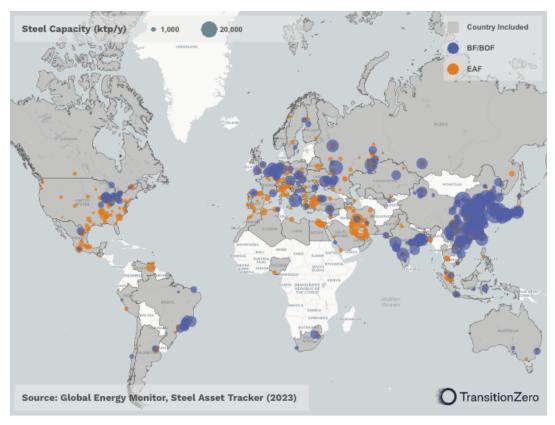


Figure 3 Global map showing steel facilities. Countries included in steel emissions monitoring are shaded in grey - with EAF (orange) and BF-BOF (blue) steel processing routes identified.

2.1.2. Remote sensing datasets

Satellite based production estimates make use of multispectral image from two different collections:

- The European Space Agency (ESA) Copernicus Sentinel mission with a resolution of 20m and a combined 5-day at the equator revisit with two satellites:
 - o Sentinel 2A: with historical images available since 2017 (ESA, 2022).
 - Sentinel 2B: with historical images available since 2017 (ESA, 2022).
- The U.S. Geological Survey (USGS) and National Aeronautics and Space Administration (NASA) Landsat program with a resolution of 30m and a combined 8-day revisit:
 - Landsat-8: with historical images available since 2015 (NASA and USGS, 2022a).
 - Landsat-9: with historical images available from 2022 (NASA and USGS, 2022b).

All images were sourced and processed using Google Earth Engine (Google Earth Engine, 2022a, 2022b, 2022c). For each image, we relied on the surface reflectance products and computed the normalised band ratio between the two short wave infrared bands of each satellite, called the Normalised Heat Index (NHI). For the respective satellite collections, we infer the NHI through the following calculations:

- Sentinel-2A/B: $NHI = \frac{(B12 B11)}{(B12 + B11)}$ Landsat-8/9: $NHI = \frac{(B07 B06)}{(B07 + B06)}$

Where B# refers to the band number for the specific satellite. This ratio was used to identify thermal anomalies within the temperature range of industrial processes, while eliminating most of the noise from reflectance interference (adapted from Kato (2021) for industrial applications). Pixel values between the Sentinel 2A/B MultiSpectral Instrument and Landsat Operational Land Imager Instrument were harmonised using NASA's band pass adjustments (NASA, 2018), allowing the two image collections to be used as if they were a single collection. The harmonised dataset ensured higher revisit for time-series surface applications. Partial images (coverage of the facility's boundaries less than 80%) and cloudy images (more than 20% clouds) are excluded. Figure 4 shows an example of NHI with identified hotspot locations at a steel plant.

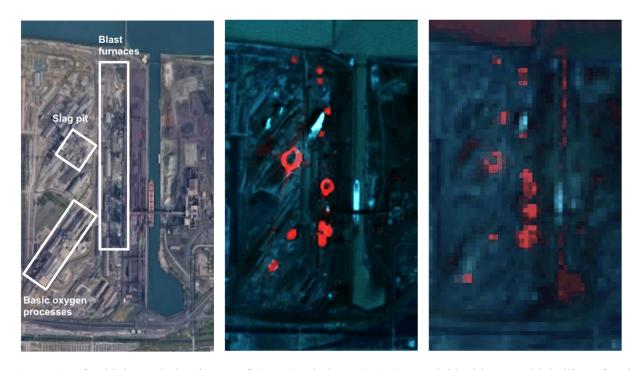


Figure 4 Left - high resolution image of Gary Steel plant, (U.S.A), overlaid with manual labelling of major heat-releasing units; Middle - composite Sentinel-2 image acquired on May 30th, 2021, using the NHI which shows thermal anomalies (red for hot pixels, blue for cooler pixels); Right - composite Landsat 8/9 image acquired the same day, also using the NHI to show thermal anomalies. Sources: modified Copernicus and USGS data, OpenStreetMap, and Terrametrics.

2.1.3. Production datasets

Aggregated steel production used in our models is primarily sourced from Worldsteel (Worldsteel, 2022) for national statistics, except for China, where more granular province level data is available through the Chinese National Bureau of Statistics (CNBS) (CNBS, 2022). This improved spatial granularity in China is particularly critical to improve modelling granularity as China accounts for approximately half of the global capacity at 1,112 Mt/a with 349 sources available in the GEM database. All aggregated production statistics are sourced on a monthly basis. In a select few countries, such as Albania, Morocco and North Korea, only annual production data were available and monthly level data were approximated through disaggregation of these values. Lastly, in 12 countries, production data were available however there were no sources associated with the given geography. This temporal and spatial mismatch likely occurs due to limitations in the GEM database which currently only has records on sources with capacity higher than 0.5 Mt/a. Future updates to this database may look to include smaller sources.

2.1.4. Emissions factors dataset

Emissions factors from Global Efficiency Intelligence were provided by steel processing route and country, and these are used to convert facility level production estimates to emissions (GEI contributors, 2022). Countries include Brazil, Canada, China, France, Germany, India, Italy, Japan, Mexico, Poland, Russia, South Korea, Spain, Turkey, Ukraine, United States, and Vietnam. For other countries we used the median emission factor by production process. Use of the median emissions factor was preferable to avoid bias through the inclusion of extreme values demonstrated in some countries. Power generation related emissions are excluded from these emission factors (when applicable) as these emissions are already accounted for in the Climate TRACE Electricity Generation sector.

2.1.5. Validation datasets

In this work facility level production and emissions models are developed based upon country-level target data. As a consequence, alternative sources are required to validate our model results. Our results are validated against source level emissions published by US Environmental Protection Agency (US EPA) and the European Pollutant Release and Transfer Register (E-PRTR) publish facility level data (E-PRTR, 2022; US EPA, 2022).

2.2. Methods

2.2.1. Production methodology

The first step to quantify emissions for each steel source is estimating the associated activity, expressed in tons of crude steel produced in a given time period. We rely on three modelling methods to estimate individual production levels, with more details given in the following sections. An example of the resulting country-level coverage of steel production by modelling method is available in Appendix 3.

2.2.1.1. Satellite-derived estimates

At BF-BOF steel plants, once hotspot activities were identified from the satellite NHIs, country level production data was used to calibrate the relationships between these activities and production at the plant level, visualised in Figure 5.

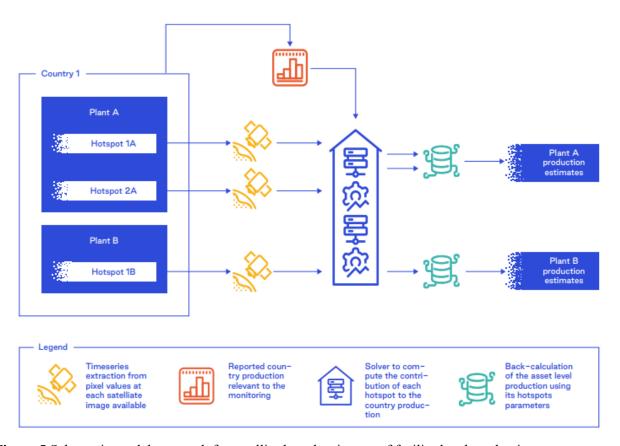


Figure 5 Schematic model approach for satellite-based estimates of facility level production.

Specifically, we begin by extracting the shapes of all hotspots at each plant. For each of these hotspots, a time-series of pixel values was extracted before signal processing was applied to normalise these values between 0 and 1 (Appendix 2). Each normalised hotspot time-series was fed into an optimizer function which developed a linear relationship (or weight) between hotspot NHI and reported country level production. Model constraints were also implemented during this optimization phase that ensured individual plant predictions should not exceed their maximum capacity. The source production estimates were then back calculated using the source's hotspot signals and their corresponding weight.

2.2.1.2. Capacity-based estimates

For sources which are not monitored via satellite (i.e., EAFs), a basic disaggregation method is applied: for each facility, we compute its share of national capacity, before multiplying this number

by the country's remaining production (after accounting for satellite-derived source production estimates) to derive the facility's contribution for the given timeframe. An illustrative example of a country with two plants A and B (of capacities C_A and C_B respectively) and a total production P_m for a country/region and a given month m, the capacity-based estimates for these two plants are (respectively $P_{A,m}$ and $P_{B,m}$):

$$\bullet \quad P_{A,m} = \frac{C_A}{C_A + C_B} \times P_m$$

$$\bullet \quad P_{B,m} = \frac{C_B}{C_A + C_B} \times P_m$$

2.2.1.3. Non-spatially allocated source estimates

In some cases, there is a shortfall between the sum of satellite and capacity-based source production estimates and the country-level production targets. This circumstance may occur when the capacity of listed sources is lower than the aggregated production figures (country or province-level). This missing production is then allocated to a non-spatially allocated source for the purposes of more accurately estimating country-level emissions.

2.2.2. Emissions methodology

With the production estimates derived for each source, the emission estimates were derived by multiplying the production by the relevant emission factor (Hasanbeigi, 2022). Electricity-related emissions were excluded from these emission factors (when applicable) as these emissions are already accounted for in the Climate TRACE Electricity generation sector emissions. For non-spatially allocated source production, the average emissions factor for a country was used and weighted by the share of processing capacity of each route within the country.

2.3. Coverage

Based on 2022 production numbers, source level estimates for 79% (1,380 Mt/a) of world crude steel production are provided. The remaining 21% are provided through the country level estimates as non-spatially allocated sources. While satellite-monitored sources only represent 28% of total sources (274 out of 973), these sources contribute to 60% of the total source level emissions estimates (1,343 out of 2,266 MtCO₂/a) in 2022. A more detailed breakdown of coverage is available in section 8.5.

3. Results and analysis

Table 1 and 2 shows the biggest producers and emitters in the steel industry at country and plant level, respectively. China accounts for more than half of the global steel production and emissions whereas the biggest and most emitting steel plants are located in South Korea owned by POSCO.

Table 1 Top producers and emitters in the steel industry - by country

Country	CO ₂ emissions (MtCO ₂)	Production (Mt)
China	1,287	899.5
India	237.2	121.5
Japan	137.2	89.2
S Korea	87.6	65.9
Russia	74.9	71.7
Brazil	44.8	34.1
Germany	44.4	36.8
USA	43.3	80.5
Vietnam	29.8	20.2
Taiwan	25.9	20.8

Table 2 Top producers and emitters in the steel industry - by plant

Plant Name	Country	CO ₂ emissions (MtCO ₂)	Production (Mt)	Capacity (Mt)
POSCO Gwangyang steel plant	S Korea	29.6	17.3	23
POSCO Pohang steel plant	S Korea	29.2	17.1	17.4
JFE West Japan Works (Fukuyama) steel plant	Japan	21.5	11.2	13
Ma'anshan Iron & Steel Co., Ltd.	China	19.2	12.6	16.8
Wuhan Iron and Steel Co., Ltd.	China	17.4	11.2	15.9
NLMK Lipetsk steel plant	Russia	17.3	11.1	13.2
Hyundai Steel Dangjin steel plant	S Korea	17.3	11.8	12.6
JFE West Japan Works (Kurashiki) steel plant	Japan	16.9	8.8	10
Liuzhou Iron & Steel Co., Ltd.(Liuzhou Base)	China	16.6	10.7	12.5
Nippon East Japan Works (Kimitsu) steel plant	Japan	16.6	8.6	10

Our methodology provides production and, by extension, emissions estimate at a monthly level. In this section our results are provided through statistical analysis. Accuracy is assessed by first gathering production and emissions data for sources before calculating performance indicators - mean absolute error (MAE) and mean absolute percentage error (MAPE).

3.1. Comparison of source level emissions estimates

In this section our CO₂ emissions estimates are compared to reported data by the US EPA and E-PRTR (US EPA, 2022; E-PRTR, 2022). Data is published at an annual resolution in both cases, so our own monthly emissions estimates will be aggregated on an annual basis for comparison. The results of this analysis are displayed in Figure 6. In total there are 98 sources (774 samples) of which 36 are from the EU (261 samples) and 62 from the US (513 samples). Overall, MAE (MAPE) accuracy across the samples is 0.53 MtCO₂/a (32%). Of the 98 sources, 75 are capacity based (591 samples) and 23 are satellite-based (183 samples) with accuracy of 0.09 MtCO₂/a (34%) and 1.5 MtCO₂/a (25%), respectively. Here a lower emissions factor in combination with generally smaller capacity for capacity-based sources results in a substantially lower MAE. By comparing MAPE, it is seen that the satellite-based approach is 9% more accurate on average.

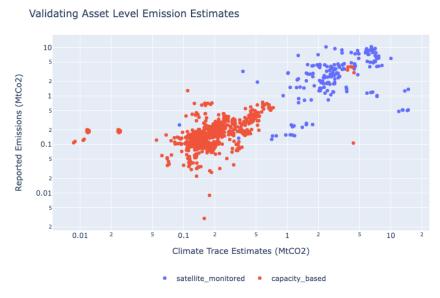


Figure 6 Comparison of source level CO_2 emissions estimates to reported estimates from the US-EPA and E-PRTR for periods January 2015 to December 2022. Red dots = satellite-based production estimates, blue dots = capacity-based estimates.

3.2. Aggregated emissions estimates for EU, US

Source level emissions estimates are the target of this work. However, it is also useful to validate our results at a more aggregated level. Figure 7 highlights the results based on aggregated annually reported emissions for selected plants (as stated in Section 3.2) in the EU and US, respectively. In

both cases our results capture the overall trend of emissions in each region. For the US, our emissions estimate closely reflects the trend of reported emissions throughout the monitoring period.

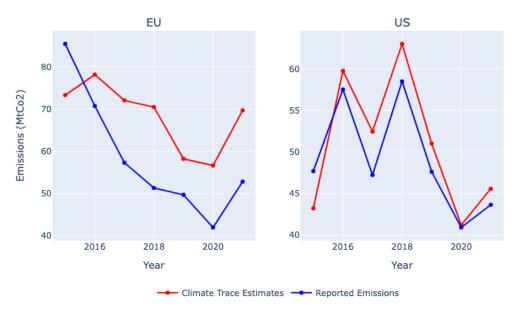


Figure 7 Comparison of reported emissions and Climate TRACE emissions estimates for the EU (left) and US (right).

4. Discussions

Our steel emissions estimates modelling approach consists of first, estimating facility level production, and second, applying a relevant emissions factor to yield facility emissions. The approach for estimating steel production at the facility level is dependent upon the facility type. BF-BOF facilities are estimated via our satellite-based approach as they contain numerous elements such as the coke plants, sinter plants, BFs, slag pits, and BOFs which can operate at temperatures in excess of 1,200°C. Such high temperatures present a strong signal which can be used as a proxy for inferring activity (or production) for a given facility at a given time. Other facilities that do not present such high temperature elements, such as EAF type facilities, are instead estimated via the capacity-based approach. The satellite-based approach is advantageous in that it can provide an estimation of dynamic activity that more realistically represents the facility itself as opposed to the capacity-based approach which reflects national production. Limitations in this approach may arise from either the incorrect detection of relevant hotspots or lack of enough training data. In both cases a larger study period is expected to improve the accuracy of our model estimates.

Following our production estimates, country or region-specific emissions factors are used for each facility type to estimate facility level emissions. These emissions estimates are the main goal of

this work. Overall accuracy is generally strong, however, was reduced due to a select few cases which did not perform well under the given methodology. Three main hypotheses are presented for differences in emissions estimates: (1) different reporting scope, (2) country/region specific emissions factors, (3) poor production estimates, and there could additionally be many causes for these discrepancies for which it is not always trivial to identify in each facility. A difference in reporting scope of the emissions between the reported data and the emission factors is suspected in some cases. In our own models we estimate direct process emissions, however, reported values were not always clear on their exact scope. An example may be that a specific facility would include the associated emissions of electricity in addition to process emissions. Another reason can be linked to the fact that the emission factors we use are country and production method specific. For countries with numerous facilities these values may not always be representative for each plant. Emissions intensity is likely to vary substantially based upon the age of the facility, fuel mix, feedstocks and use of energy-efficient technologies (Hasanbeigi, 2022). Lastly, our production estimates may themselves be incorrect and, as emissions are directly proportional, would result in poor estimates for the emissions themselves. A case here would be for a capacity-based facility which is not producing consistently with the national average and would therefore not be represented well.

5. Conclusions

In this work a new approach has been implemented to estimate steel production and emissions at a facility level and on a monthly basis for all 1,016 steel plant sources, covering 89 countries, in the GEM database. In total these sources cover 79% (1,379 out of 1,753 Mt/a) of worldwide crude steel production.

The methodology was dependent on the production route, BF-BOF and EAF. Facilities with strong heat signatures, such as BF-BOF processing routes, were approximated using satellite-derived NHI measurements. Such an approach allowed for the use of machine learning to infer a relationship between facility hotspots and steel production - ultimately resulting in more accurate and timely estimates. For facilities where steel processing routes did not provide useful satellite information, such as EAF, a capacity-based approach was used to infer production estimates. Based on the 2022 production numbers, our satellite-monitored sources account for only 28% of total sources, however, contribute 60% of total source CO₂ emissions (1,343 out of 2,266 MtCO₂/a). Comparing capacity and satellite-based production estimates to reported values, there is marked improvement of 7% MAPE in accuracy for satellite-based sources. Overall, MAE (MAPE) accuracy was at 0.36 Mt/a (18%) for all samples. The remaining 18% of production not accounted for belonged to small sources (<0.5 Mt/a) which, for the purpose of estimating emissions at the country level, were introduced as non-spatially allocated sources.

To calculate emissions, we applied country and process-specific emission factors to our production estimates which were subsequently compared to reported facility and country-level figures (Hasanbeigi, 2022). The final results show that our methods were effective at estimating steel production, and emissions in most cases. The few cases with lower emissions accuracy highlight that a wide variation in emissions factors may be present for a given facility within a country which may not necessarily be fully captured by just process type alone.

Ultimately, this work has shown that timely and accurate facility level production and emissions can be obtained without the need to rely upon factory published figures, often years out of date, and may open pathways to assist in future climate policy. Future works may benefit from greater availability of source data to train in modelling, allowing subsequent versions of our model to benefit from a greater scope of historical data that may further improve our model results. Furthermore, as GEM begins to target smaller sources for inclusion in their database, we will be able to update our own results accordingly.

6. Supplementary metadata section

Steel sector CO2 emissions are reported for individual sources for the period 2015 - 2022. The emissions described here represent a subset of specific country-level emissions estimates from the Climate TRACE manufacturing sector. All data is freely available on the Climate TRACE website (https://climatetrace.org/). A detailed description of what is available is described in Tables 3 and 4.

Table 3 Metadata for Emissions for steel sector.

General Description	Definition
Sector definition	Emissions from steel production
UNFCCC sector equivalent	2.C.1 Iron and steel production
Temporal Coverage	2015 – 2022
Temporal Resolution	Monthly
Data format	CSV
Coordinate Reference System	None. ISO3 country code provided
Number of sources available for download	1,016 steel sources covering 89 countries
What emission factors were used?	Regional emission factors for each production route
What is the difference between a "0" versus "NULL/none/nan" data field?	"0" values are for true non-existent emissions. If we know that the sector has emissions for that specific gas, but the gas was not modelled, this is represented by "NULL/none/nan"

total	CO2e	_100yrGWI	P and
total_	_CO2e_	_20yrGWP	conversions

Climate TRACE uses IPCC AR6 CO2e GWPs. CO2e conversion guidelines are here:

https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_W GI_FullReport_small.pdf

Table 4 Steel s	Table 4 Steel sector description for confidence and uncertainty in emissions.						
Data attribute	Confidence Definition	Uncertainty Definition					
type	 Very low: Based on highly speculative or obsolete information. Very low level of confidence in the accuracy of steel plant classification. Low: Limited or somewhat outdated data. Low level of confidence in the classification's correctness. Medium: A mix of historical and more recent data. A medium level of confidence in its accuracy. High: Grounded in comprehensive and recent data. A high level of confidence in the precise classification of the steel plant. Very high: Extensive, up-to-date, and verified data. A very high level of confidence in the accurate and detailed identification of the steel plant. 	Not used; N/A					
capacity_description	 Very low: Limited or outdated data, and significant uncertainties exist. Low: Outdated and/or incomplete data. Medium: A mix of historical and recent data. High: Comprehensive and recent data updates. High level of certainty. Very high: Extensive, up-to-date, and verified data. Very high level of certainty. 	Not used; N/A					
capacity_factor_description	 Very low: Data is sparse or highly unreliable. Considerable uncertainty in capacity factor estimations. Low: Moderate uncertainty in capacity factor calculations. Medium: Data is sufficiently available, though not comprehensive. No absolute accuracy in capacity factor estimations. High: High confidence in the accuracy of capacity factor calculations. Very high: Derived from thorough and validated data sources. Very high 	Not used; N/A					

	precision of capacity factor estimations.	
activity_description	 Very low: Largely speculative or based on outdated information. A very low level of confidence in activity assessments. Low: Limited or somewhat outdated sources. A low level of confidence in the activity assessments. Medium: A mix of historical and more recent data. Medium level of confidence in activity insights. High: Detailed and current operational data ensures a high level of confidence in the accuracy of activity assessments. Very high: Extensive, verified, and up-to-date data. A very high level of confidence in their accuracy. 	10% of source production (based on IPCC)
CO2_emissions_factor	 Very low: Highly uncertain due to insufficient or unreliable data. Low: Estimated from incomplete data. Low confidence level in its precision. Medium: A mix of historical and more recent data. Medium level of confidence in their accuracy. High: Derived from comprehensive and recent data. A high level of confidence in their precision. Very high: Based on extensive and validated data, providing a very high level of confidence in their precision. 	10% of source emission factor (based on IPCC)
CH4_emissions_factor	Not used; N/A	Not used; N/A
N2O_emissions_factor	Not used; N/A	Not used; N/A
other_gas_emissions_factor	Not used; N/A	Not used; N/A
CO2_emissions	 Very low: Based on very rough estimations or outdated information. A very low level of confidence in its accuracy. Low: Estimated from incomplete data. Low confidence level in its precision. Medium: A mix of historical and more recent data. Medium level of confidence in their accuracy. High: Derived from comprehensive and recent data. A high level of confidence in their precision. 	20% of source emissions

	Very high: Based on extensive and validated data, providing a very high level of confidence in their precision.	
CH4_emissions	Not used; N/A	Not used; N/A
N2O_emissions	Not used; N/A	Not used; N/A
other_gas_emissions	Not used; N/A	Not used; N/A
total_CO2e_100yrGWP	 Very low: Based on very rough estimations or outdated information. A very low level of confidence in its accuracy. Low: Estimated from incomplete data. Low confidence level in its precision. Medium: A mix of historical and more recent data. Medium level of confidence in their accuracy. High: Derived from comprehensive and recent data. A high level of confidence in their precision. Very high: Based on extensive and validated data, providing a very high level of confidence in their precision. 	20% of source emissions
total_CO2e_20yrGWP	 Very low: Based on very rough estimations or outdated information. A very low level of confidence in its accuracy. Low: Estimated from incomplete data. Low confidence level in its precision. Medium: A mix of historical and more recent data. Medium level of confidence in their accuracy. High: Derived from comprehensive and recent data. A high level of confidence in their precision. Very high: Based on extensive and validated data, providing a very high level of confidence in their precision. 	20% of source emissions

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

Citation format: Isabriye, A., Sinha, A. (2023). *Manufacturing sector - Steel Sector Emissions*. TransitionZero, UK, Climate TRACE Emissions Inventory - https://climatetrace.org [Accessed date]

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 <u>contact</u> <u>us</u>.

7. References

- 1. CNBS (2022) 'Steel production data retrieved from: CNBS'. China National Bureau of Statistics. Available at: http://www.stats.gov.cn/english/.
- 2. E-PRTR contributors (2022) 'Steel emissions data retrieved from: E-PRTR'. European Pollutant Release and Transfer Register (E-PRTR). Available at: http://prtr.ec.europa.eu/.
- 3. European Space Agency (ESA) (2022) 'Satellite images data retrieved from: Copernicus Sentinel-2 ESA'. Available at: https://www.esa.int/Applications/Observing the Earth/Copernicus/Sentinel-2.
- 4. GEI contributors (2022) 'Steel emission factor data retrieved from: GEI'. Global Efficiency Intelligence (GEI). Available at: https://www.globalefficiencyintel.com/.
- 5. GEM contributors (2022) 'Steel inventory data retrieved from: GEM Global Steel Plant Tracker'. Global Energy Monitor (GEM). Available at: https://globalenergymonitor.org/projects/global-steel-plant-tracker/.
- 6. Google Earth Engine (2022a) *Landsat 8 Datasets in Earth Engine*, *Google Developers*. Available at: https://developers.google.com/earth-engine/datasets/catalog/landsat-8 (Accessed: 26 July 2022).
- 7. Google Earth Engine (2022b) *Landsat 9 Datasets in Earth Engine*, *Google Developers*. Available at: https://developers.google.com/earth-engine/datasets/catalog/landsat-9 (Accessed: 26 July 2022).
- 8. Google Earth Engine (2022c) *Sentinel-2 Datasets in Earth Engine*, *Google Developers*. Available at: https://developers.google.com/earth-engine/datasets/catalog/sentinel-2 (Accessed: 26 July 2022).
- 9. Google Maps contributors (2022) 'Steel inventory data retrieved from: Google Maps'. Google. Available at: https://www.google.co.uk/maps.
- 10. Hasanbeigi, A. (2022) 'An International Benchmarking of Energy and CO2 Intensities', p. 31.
- 11. IEA (2020) 'Iron and Steel Technology Roadmap Towards more sustainable steelmaking', p. 190.
- 12. Kato, S. (2021) Automated classification of heat sources detected using SWIR remote sensing Elsevier Enhanced Reader. Available at: https://doi.org/10.1016/j.jag.2021.102491.
- 13. Liangrocapart, S., Khetkeeree, S. and Petchthaweetham, B. (2020) 'Thermal Anomaly Level Algorithm for Active Fire Mapping by Means of Sentinel-2 Data', in *2020 17th International*

- Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON).
- 14. Marchese, F. *et al.* (2019) 'A Multi-Channel Algorithm for Mapping Volcanic Thermal Anomalies by Means of Sentinel-2 MSI and Landsat-8 OLI Data', *Remote Sensing*, 11(23), p. 2876. Available at: https://doi.org/10.3390/rs11232876.
- 15. M'Barek, B. and Gray, M. (2021) 'The Spotlight Effect'. Transition Zero. Available at: https://www.transitionzero.org/reports/the-spotlight-effect.
- 16. National Aeronautics and Space Administration (NASA) (2018) 'Band Pass Adjustment'. Available at: https://hls.gsfc.nasa.gov/algorithms/bandpass-adjustment/ (Accessed: 25 July 2022).
- 17. National Aeronautics and Space Administration (NASA) and United States Geological Survey (USGS) (2022a) 'Satellite images data retrieved from: Landsat-8 NASA/USGS'. Available at: https://www.usgs.gov/landsat-missions/landsat-8.
- 18. National Aeronautics and Space Administration (NASA) and United States Geological Survey (USGS) (2022b) 'Satellite images data retrieved from: Landsat-9 NASA/USGS'. Available at: https://www.usgs.gov/landsat-missions/landsat-9.
- 19. OpenStreetMap contributors (2022) 'Steel inventory data retrieved from: OpenStreetMap'. OpenStreetMap. Available at: https://www.openstreetmap.org/.
- 20. UNFCCC contributors (2022) 'Steel emissions data retrieved from: UNFCCC'. United Nations Framework Convention on Climate Change (UNFCCC). Available at: https://unfccc.int/process-and-meetings/transparency-and-reporting/greenhouse-gas-data/ghg-d ata-unfccc/ghg-data-from-unfccc.
- 21. US EPA contributors (2022) 'Steel emissions data retrieved from: US EPA'. United States Environmental Protection Agency (US EPA). Available at: https://www.epa.gov/.
- 22. Worldsteel contributors (2022) 'Steel production data retrieved from: Worldsteel'. Worldsteel.
- 23. Zhou, Y. *et al.* (2018) 'A Method for Monitoring Iron and Steel Factory Economic Activity Based on Satellites', *Sustainability*, 10(6), p. 1935. Available at: https://doi.org/10.3390/su10061935.

8. Appendices

8.1. List of acronyms

Table 5 List of acronyms.

BF	Blast Furnace
BOF	Basic Oxygen Furnace
CNBS	China National Bureau of Statistics
E-PRTR	European Pollutant Release and Transfer Register
EAF	Electric Arc Furnace
GEI	Global Efficiency Intelligence
GEM	Global Energy Monitor
GSPT	Global Steel Plant Tracker
IEA	International Energy Agency
MAE	Mean absolute error
MAPE	Mean absolute percentage error
Mt	Million metric ton
NASA	National Aeronautics and Space Administration
NHI	Normalised Heat Index
US EPA	United States Environmental Protection Agency
USGS	United States Geological Survey

8.2. Appendix 1: Overview of the steel production process

The principal inputs to steelmaking today are iron ore, energy, limestone and scrap. Iron ore and scrap are used to provide the metallic charge, with scrap having a much higher metallic concentration than iron ore. Energy inputs are used to provide heat to melt the metallic input, and in the case of iron ore, to chemically remove oxygen. Limestone is used at various stages of the steelmaking process to help remove impurities. Indirect carbon emissions vary widely based on the production route. Figure 8 highlights these main process routes.

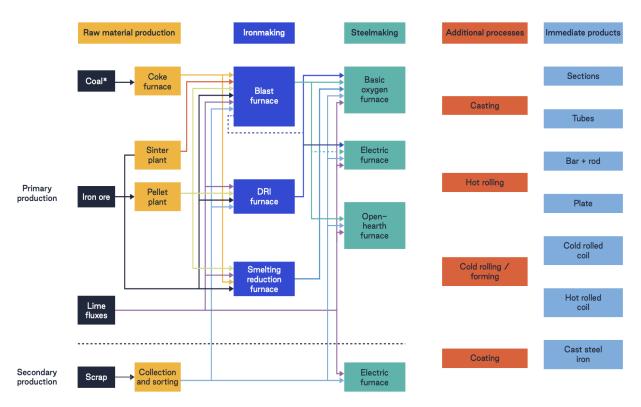


Figure 8 Adapted from Cullen, et al. (2021) and IEA (2020).

8.3. Appendix 2: Example of production estimates derived from satellite data - Gary Steel, Indiana, USA

This section illustrates the processing steps to derive production estimates from satellite imagery. We use the example of Gary Steel works, Indiana, USA, where we identify 12 hotspots in the plant (i.e., blast furnaces, slag pit etc.; see Figure 3).

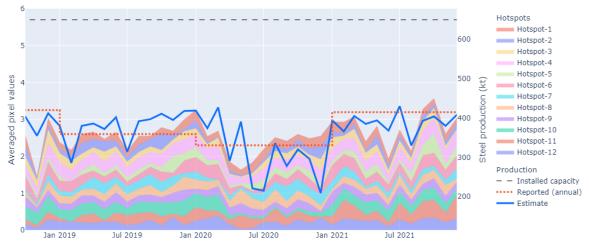


Figure 9 NHI values and production data for Gary Steel, Indiana, USA. Time series of averaged pixel values at each hotspot on the left axis. Reported production (red dashed line), NHI estimated production (solid blue line) and installed capacity (black dashed line) are shown on the right axis.

- 1. We start by extracting every pixel value from each image within the hotspots. Figure 9 shows the monthly average pixel values for each hotspot. At this stage it already shows a certain level of correlation with reported production.
- 2. We reduce dimensionality by applying a principal component analysis on the hotspots. The optimal number of principal components is derived through hyperparameter training (in the case of Gary steel, the optimal PCA is of one component).
- 3. We then normalise the reduced feature space between 0 and 1 by applying a MinMaxScaler.
- 4. This processed time series is then fed into the optimiser described in Figure 5, to derive the weights allowing the optimal conversion to tons of crude steel, while respecting the constraints (estimated production does not exceed installed capacity).

The advantages of this production estimation process become evident in Figure 9 whereby production values are provided at a monthly frequency (blue line) compared to the annual (red dashed line) frequency (in addition to being provided with a lag of only 1 month). This improved granularity allows for greater insight into facility production rates. Gary Steel works could be described as functioning at approximately half capacity for 2019 and 2021 with fairly consistent production throughout the year. For 2020 however, a year marred by the pandemic many steel production facilities across the country operated at reduced capacity. This substantial drop in production is captured and reflected in our estimation.

8.4. Appendix 3: Example of the country-level coverage of steel production by modelling method - Germany

Germany is a country with a relatively large steel production capacity (50.5 Mt/a) and benefits from monthly country-level production data. According to the GEM database it has 18 sources. Of the sources, 8 are BF-BOF (Total of 37.7 Mt/a) allowing for production estimates to be approximated via satellite monitoring, 9 are EAF and 1 DRI-EAF (Total of 12.8 Mt/a), and production is subsequently estimated via the capacity-based approach. In order to estimate facility level production, a model is first developed for the satellite-based sources and trained against a scaled (in accordance with capacity) country level reported data. The significance of each hotspot within each source is a learned outcome, and their NHI variation over time is indicative of the production activity at a particular facility. Where satellite-based methodology allows for capturing the variation in production for a particular facility and may vary accordingly, the remaining capacity-based source production is estimated based on the difference in known country level production from the satellite production estimates. As such the production activity between capacity-based sources is proportional to one another. Lastly, the shortfall in country production by our estimates is attributed to the omission of small sources (<0.5 Mt/a) in the GEM database. We account for this missing production via the introduction of a non-spatially allocated source. The production estimates for each of these methodologies is expressed for Germany in Figure 10.

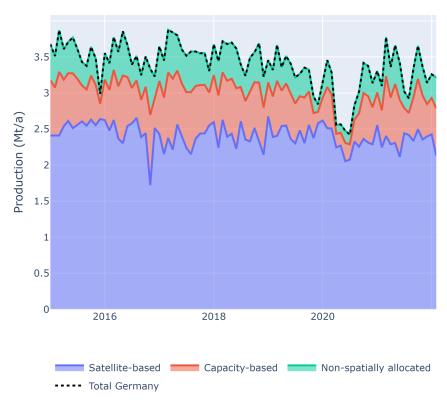


Figure 10 Source type production estimates for Germany (2015-2021).

8.5. Appendix 4: Summary of the coverage by geographies

Table 6 Summary of the data coverage by country, with a breakdown of the steel plants monitored by satellite (expressed by number of sources, share of installed capacity, production and emissions). Values representative for 2022.

Country	Total sources	Satellite- monitored Sources	Total source production (Mt)	Satellite- monitored production (Mt)	Total source emissions (MtCO ₂)	Satellite- monitored emissions (MtCO ₂)	Total source capacity (Mt)	Satellite- monitored capacity (Mt)
China	350	159 (45%)	899.53	469.6 (52%)	1287.03	729.8 (56%)	1417.24	698.0 (49%)
India	57	16 (28%)	121.52	35.2 (28%)	237.19	82.7 (35%)	171.41	72.2 (42%)
Japan	45	15 (33%)	89.24	62.6 (70%)	137.23	120.5 (88%)	132.92	86.3 (65%)
South Korea	17	3 (18%)	65.85	44.3 (67%)	87.62	75.6 (86%)	96.93	53.0 (55%)
Russia	31	8 (26%)	71.75	41.8 (57%)	74.94	60.9 (81%)	95.94	55.1 (56%)
Brazil	25	7 (28%)	34.09	20.2 (59%)	44.79	34.2 (76%)	53.40	32.1 (60%)
Germany	20	8 (40%)	36.85	26.9 (73%)	44.45	39.3 (88%)	57.23	37.7 (66%)

		•						
United States	75	9 (12%)	80.53	18.3 (23%)	43.32	28.0 (65%)	119.79	36.4 (30%)
Vietnam	16	2 (12%)	20.18	8.7 (43%)	29.78	17.2 (57%)	30.75	13.1 (43%)
Taiwan	10	2 (20%)	20.80	4.5 (22%)	25.94	7.4 (28%)	26.89	6.0 (22%)
Turkey	26	3 (12%)	35.03	10.2 (28%)	20.95	15.8 (75%)	52.08	13.3 (26%)
France	10	2 (20%)	12.12	7.7 (64%)	16.83	13.8 (82%)	20.29	10.8 (53%)
Indonesia	8	0 (0%)	14.30	0 (0%)	16.15	0 (0%)	14.30	0 (0%)
Canada	10	3 (30%)	12.10	5.8 (48%)	13.23	8.8 (67%)	17.40	8.0 (46%)
Ukraine	11	6 (55%)	6.05	4.8 (80%)	11.63	9.9 (85%)	42.74	26.5 (62%)
Italy	24	1 (4%)	21.60	4.5 (21%)	11.53	7.5 (65%)	36.97	11.5 (31%)
Mexico	17	2 (12%)	18.14	1.9 (10%)	11.26	3.0 (27%)	26.44	6.0 (23%)
Netherlands	1	1 (100%)	6.65	6.6 (100%)	10.43	10.4 (100%)	7.50	7.5 (100%)
Iran	31	2 (6%)	30.59	2.1 (7%)	10.31	3.4 (33%)	50.10	5.1 (10%)
Poland	7	1 (14%)	7.41	2.7 (36%)	9.95	5.9 (59%)	13.40	5.0 (37%)
Austria	2	2 (100%)	6.31	6.3 (100%)	9.90	9.9 (100%)	7.57	7.6 (100%)
United Kingdom	7	2 (28%)	6.17	5.7 (93%)	9.70	9.4 (97%)	16.22	8.2 (51%)
Belgium	5	1 (20%)	7.03	4.4 (63%)	8.48	7.0 (82%)	10.15	5.0 (49%)
Australia	5	2 (40%)	5.67	3.8 (66%)	7.58	6.1 (81%)	7.36	4.4 (60%)
Czechia	4	2 (50%)	4.55	4.2 (93%)	7.00	6.6 (95%)	10.61	6.2 (57%)
Slovakia	3	1 (33%)	4.52	2.7 (61%)	6.38	4.3 (68%)	8.31	4.5 (54%)
Spain	14	1 (7%)	11.51	0.7 (6%)	6.05	0.9 (16%)	19.44	1.2 (6%)
Pakistan	5	1 (20%)	6.01	0 (0%)	5.34	0 (0%)	10.00	3.0 (30%)
South Africa	5	2 (40%)	4.40	1.4 (31%)	4.92	2.3 (46%)	12.03	6.4 (53%)
Sweden	6	2 (33%)	4.40	2.9 (65%)	4.78	4.5 (94%)	6.81	3.8 (56%)
Kazakhstan	4	1 (25%)	4.15	1.0 (25%)	4.37	1.7 (39%)	10.08	6.0 (60%)
Malaysia	11	0 (0%)	6.90	0 (0%)	4.16	0 (0%)	15.45	0 (0%)

Argentina	5	1 (20%)	5.09	1.8 (36%)	3.99	3.0 (75%)	7.15	3.2 (45%)
Romania	4	1 (25%)	3.38	2.4 (70%)	3.91	3.7 (95%)	5.35	3.2 (60%)
Finland	4	1 (25%)	3.54	2.0 (56%)	3.81	3.2 (83%)	5.27	2.6 (49%)
Saudi Arabia	6	0 (0%)	9.07	0 (0%)	3.37	0 (0%)	12.68	0 (0%)
Bangladesh	6	0 (0%)	5.50	0 (0%)	2.73	0 (0%)	7.06	0 (0%)
Serbia	3	1 (33%)	1.70	1.5 (87%)	2.52	2.4 (96%)	2.88	2.2 (76%)
Algeria	3	0 (0%)	4.20	0 (0%)	2.38	0 (0%)	7.70	0 (0%)
Egypt	8	0 (0%)	9.79	0 (0%)	2.37	0 (0%)	14.00	0 (0%)
North Korea	3	0 (0%)	1.25	0 (0%)	2.04	0 (0%)	7.72	0 (0%)
Hungary	3	1 (33%)	1.20	1.2 (97%)	1.85	1.8 (100%)	2.84	1.6 (56%)
Chile	3	1 (33%)	1.15	0.7 (57%)	1.47	1.1 (75%)	2.59	1.4 (56%)
Thailand	9	0 (0%)	5.32	0 (0%)	1.29	0 (0%)	8.36	0 (0%)
Bosnia and Herzegovina	2	1 (50%)	0.89	0.7 (79%)	1.19	1.1 (96%)	1.94	1.1 (59%)
Uzbekistan	2	0 (0%)	0.70	0 (0%)	1.14	0 (0%)	1.01	0 (0%)
United Arab Emirates	2	0 (0%)	3.21	0 (0%)	1.03	0 (0%)	3.80	0 (0%)
Switzerland	2	0 (0%)	1.20	0 (0%)	1.02	0 (0%)	1.82	0 (0%)
Portugal	3	0 (0%)	1.95	0 (0%)	1.01	0 (0%)	2.54	0 (0%)
Philippines	2	0 (0%)	0.90	0 (0%)	0.86	0 (0%)	1.86	0 (0%)
Peru	3	0 (0%)	1.10	0 (0%)	0.84	0 (0%)	2.85	0 (0%)
Belarus	1	0 (0%)	2.08	0 (0%)	0.50	0 (0%)	3.00	0 (0%)
Oman	3	0 (0%)	2.00	0 (0%)	0.48	0 (0%)	4.25	0 (0%)
Kuwait	2	0 (0%)	1.30	0 (0%)	0.41	0 (0%)	1.30	0 (0%)
Luxembourg	2	0 (0%)	1.87	0 (0%)	0.37	0 (0%)	3.00	0 (0%)
Qatar	2	0 (0%)	1.08	0 (0%)	0.35	0 (0%)	2.82	0 (0%)
Greece	5	0 (0%)	1.54	0 (0%)	0.30	0 (0%)	6.60	0 (0%)

Norway	2	0 (0%)	0.70	0 (0%)	0.30	0 (0%)	0.86	0 (0%)
Singapore	2	0 (0%)	0.58	0 (0%)	0.28	0 (0%)	1.05	0 (0%)
New Zealand	2	0 (0%)	0.57	0 (0%)	0.24	0 (0%)	0.79	0 (0%)
Libya	1	0 (0%)	0.69	0 (0%)	0.17	0 (0%)	1.75	0 (0%)
Slovenia	2	0 (0%)	0.60	0 (0%)	0.16	0 (0%)	0.77	0 (0%)
Moldova	1	0 (0%)	0.56	0 (0%)	0.14	0 (0%)	1.00	0 (0%)
Morocco	2	0 (0%)	0.50	0 (0%)	0.12	0 (0%)	1.80	0 (0%)
Bulgaria	1	0 (0%)	0.48	0 (0%)	0.09	0 (0%)	1.20	0 (0%)
Albania	1	0 (0%)	0.05	0 (0%)	0.08	0 (0%)	0.70	0 (0%)
Guatemala	1	0 (0%)	0.27	0 (0%)	0.07	0 (0%)	0.50	0 (0%)
North Macedonia	1	0 (0%)	0.25	0 (0%)	0.06	0 (0%)	0.55	0 (0%)
Azerbaijan	1	0 (0%)	0.20	0 (0%)	0.05	0 (0%)	1.00	0 (0%)
Croatia	1	0 (0%)	0.17	0 (0%)	0.03	0 (0%)	0.35	0 (0%)
Nigeria	1	0 (0%)	0.10	0 (0%)	0.02	0 (0%)	1.30	0 (0%)
Venezuela	1	0 (0%)	0.03	0 (0%)	0.01	0 (0%)	5.10	0 (0%)
	2	0 (0%)	0.01	0 (0%)	0.00	0 (0%)	2.20	0 (0%)
Syria Trinidad and Tobago	1	0 (0%)	0.00	0 (0%)	0.00	0 (0%)	1.00	0 (0%)

Table 7 Summary of the coverage in Chinese provinces, with a breakdown of the steel plants monitored by satellite in the coverage (expressed by number of sources, share of installed capacity, production and emissions). Values representative for 2022.

Chinese province	Total sourc es	Satellite Sources	Total production (Mt)	Satellite production (Mt)	Total emissions (MtCO ₂)	Satellite emissions (MtCO ₂)	Total capacity (Mt)	Satellite capacity (Mt)
Hebei	57	34 (60%)	207.08	127.9 (62%)	308.45	198.7 (64%)	297.46	176.3 (59%)
Liaoning	16	10 (62%)	69.66	37.7 (54%)	101.38	58.6 (57%)	102.97	57.9 (56%)
Jiangsu	30	14 (47%)	65.81	45.0 (68%)	96.54	69.9 (72%)	108.79	56.7 (52%)
Shandong	14	9 (64%)	65.4	26.7 (41%)	92.79	41.5 (45%)	106.69	45.6 (43%)
Shanxi	24	17 (71%)	54.42	29.8 (55%)	80.75	46.3 (56%)	87.55	45.9 (52%)

Anhui	16	7 (44%)	37.46	20.8 (56%)	52.43	32.3 (62%)	52.64	29.2 (56%)
Guangxi	18	3 (17%)	37.81	16.0 (42%)	51.3	24.9 (49%)	55.61	21.1 (38%)
Hubei	18	3 (17%)	36.44	14.7 (40%)	48.62	22.8 (47%)	51.88	23.6 (45%)
Inner Mongolia	11	6 (55%)	30.36	10.3 (34%)	43.48	15.9 (37%)	49.93	25.0 (50%)
Jiangxi	7	4 (56%)	27.68	16.1 (57%)	39.73	25.0 (63%)	41.57	20.9 (50%)
Fujian	16	5 (31%)	27.98	11.4 (41%)	38.77	17.7 (46%)	45.6	16.9 (37%)
Sichuan	15	3 (20%)	28.77	10.6 (37%)	37.93	16.4 (43%)	43.41	15.8 (36%)
Hunan	5	3 (60%)	24.96	13.4 (54%)	36.33	20.8 (56%)	36.37	19.2 (53%)
Henan	15	9 (60%)	22.91	11.8 (51%)	33.14	18.3 (55%)	45.32	22.4 (50%)
Yunnan	12	6 (50%)	21.96	10.4 (47%)	31.53	16.2 (51%)	31.58	14.8 (47%)
Tianjin	5	3 (60%)	14.77	7.1 (48%)	21.47	11.0 (51%)	33.44	14.9 (45%)
Shaanxi	7	3 (43%)	14.72	6.7 (45%)	20.39	10.3 (51%)	22.92	11.6 (51%)
Jilin	9	4 (44%)	13.55	7.1 (53%)	19.1	11.1 (57%)	26.74	12.4 (46%)
Zhejiang	9	2 (22%)	13.76	7.4 (54%)	18.47	11.5 (62%)	21.5	9.0 (42%)
Xinjiang	8	5 (62%)	12.67	6.4 (51%)	17.46	10.0 (56%)	23.34	12.1 (52%)
Chongqing	4	1 (25%)	9.15	6.6 (72%)	13.3	10.2 (77%)	15.8	8.4 (53%)
Heilongjiang	4	2 (50%)	8.53	4.3 (50%)	12.42	6.6 (53%)	13.38	6.2 (46%)
Ningxia	3	1 (33%)	6.1	0.9 (15%)	8.85	1.4 (16%)	9.19	1.8 (20%)
Gansu	4	1 (25%)	6.04	2.0 (33%)	8.44	3.1 (37%)	10.84	4.0 (37%)
Guizhou	8	0 (0%)	4.71	0 (0%)	4.73	0 (0%)	13.18	0 (0%)
Shanghai	2	0 (0%)	1.86	0 (0%)	2.44	0 (0%)	18.81	0 (0%)
Hebei	57	34 (60%)	207.08	127.9 (62%)	308.45	198.7 (64%)	297.46	176.3 (59%)
Liaoning	16	10 (62%)	69.66	37.7 (54%)	101.38	58.6 (57%)	102.97	57.9 (56%)