

# Carbon Majors Readme

## Table of Contents

About .....	2
Methodology.....	3
Data Dictionary .....	8
Glossary.....	11
Support .....	12
Terms and Conditions.....	13

## About

Carbon Majors is a database of historic production data from 122 of the world's largest oil, gas, coal, and cement producers. This data is used to quantify the direct operational emissions and emissions from the combustion of marketed products that can be attributed to these entities. These entities include:

- 75 Investor-owned Companies, 36 State-owned Companies, 11 Nation States,
- 82 Oil Producing Entities, 81 Gas Entities, 49 Coal Entities, 6 Cement Entities

The data spans back to 1854 and contains over 1.42 trillion tonnes of CO<sub>2</sub>e covering 72% of anthropogenic emissions since the start of the Industrial Revolution.

## Methodology

The Carbon Majors database aims to trace greenhouse gas emissions from fossil fuels and cement produced by companies, historically from as early as 1854 to the present. This section gives an overview of the methodology that Carbon Majors uses to achieve this. For a more detailed description of this methodology, including discussion around the accounting protocol, calculation of emissions factors, historical attribution, uncertainties, etc., please refer to Rick Heede's 2014 paper, [Carbon Majors: Methods & Results Report](#)<sup>1</sup>.

### Entity Selection

Greenhouse gas emissions data has historically primarily been collected at the country level. The Carbon Majors database was created to instead link these emissions to fossil fuel production companies, or "carbon majors". Carbon Majors originally selected extant companies from a variety of sources that met an 8 MtCO<sub>2</sub> per year emissions threshold. Some entities in the database do not meet this threshold, for example, companies that met the threshold when the Carbon Majors project was started but have since shrunk, or smaller companies acquired by larger ones. However, this guideline still applies to ensure a manageable number of entities. The number of entities assessed may vary over time due to mergers and acquisitions, as well as additions to the database.

The assessed entities are divided into three entity types: investor-owned companies, state-owned companies, and nation-state producers. Investor-owned companies include both publicly listed and privately held producers. Nation-state producers are used primarily in the coal sector and are included only when investor-owned or state-owned companies haven't been established or played a minor role in the relevant country. Examples include North Korea and former Soviet states (the former Soviet Union and separately the Russian Federation, Kazakhstan, Ukraine, etc.). While current production is available for some Chinese coal entities, historic production data is unavailable and it has not been possible to verify the ownership structure of these entities, many of which are reportedly operated or directed by provincial government. Hence, China's coal production has been aggregated and reported as a nation state. State-owned companies are often partially owned by institutional or individual shareholders. These are considered state owned if more than fifty percent of shares are controlled by the state.

The database also tracks mergers and acquisitions. In such cases, the acquired companies' emissions are attributed to the surviving company. Divestitures are inherently accounted for, as the production from divested assets will not be included in subsequent company disclosures (see Production Data below for further

---

<sup>1</sup> The data storage and processing methods as well as the output formats have changed. The new data structure is accessible and is explained at on the website.

explanation). Assets that have been nationalized or expropriated are also monitored to the extent that equity-owned production is reported accurately by the relevant entities. Breakups of companies are also accounted for. For example, the multiple smaller companies into which the Standard Oil Trust was broken up have evolved to become some of the most recognizable companies in the database today. Some are direct descendants of Standard Oil, like ExxonMobil, with both Exxon and Mobil as descendants separately, and Chevron. Others have resulted from mergers with descendants of Standard Oil, such as BP and ConocoPhillips.

## Production Data

Carbon Majors obtains production data for each entity for each year. Due to the importance of transparency in the Carbon Majors approach, self-reported production data by the producing entity is always preferred to and used instead of any other sources. This includes annual reports, company histories, SEC filings, operation reviews, online datasets on production, etc. However, in some cases, reputable third-party sources are used when self-reported data is unavailable, particularly for national companies that irregularly, inaccurately, or do not publish production data. Alternative, third-party sources include the U.S. Energy Information Administration (EIA), the Keystone Coal Industry Manual, the Oil & Gas Journal (especially the annual OGI100/150 issues), and others.

For historical data, complete production records are sought and used where applicable. However, for some entities, production data is unavailable from the establishment of the entity, resulting in minor cases of underreporting, as early production is often overshadowed by later company expansions. In such cases, the entity's missing early production data are left empty. This is especially true before the U.S. Securities Act of 1933, which required companies to provide full and accurate financial and operational information. Other data gaps sometimes occur, often due to missing annual reports. In the absence of alternative available data, such gaps are filled through interpolation of surrounding data.

Net production data is preferred to gross, as gross production often includes output from joint ventures, production-sharing partnerships, or a state resource owner. While reporting gross production was common in the 1960s and early 1970s, it tends to overestimate emissions. In such cases, net production is estimated by applying a net-to-gross ratio. State-owned oil and gas companies typically report total production rather than their equity share. This practice can lead to a potential issue of double counting, where production is recorded both as overseas equity production by multinational oil and gas companies and as production by state-owned entities. To address this, collating data from third-party sources is utilized to adjust self-attributed production. This involves reducing total national production by a percentage corresponding to the portion of production owned by the state.

Due to variations in how companies report production, the reported production data is standardized to a common commodity type, each with a standard unit: Oil & Natural Gas Liquids (million barrels), Natural Gas (billion cubic feet), and Coal (million tonnes).

To improve data accuracy, coal production is further categorized by rank, such as bituminous or anthracite, or by utilization, such as thermal or metallurgical. Preferably, coal rank data reported by the producing entity is used. However, coal rank is frequently reported in generic terms, but often with data on heat content. Using this information, along with the geographical locations of coal mines, enables coal rank categorization when entities fail to do so themselves. If this information is only available for specific years, this coal rank split may be applied to production data outside these years as an estimation.

## Emissions Calculations

### Fossil Fuel Emissions

Emission factors for each fuel type are used to estimate the carbon content released when these fossil fuels are combusted. These emissions factors were mostly derived from Tier 1 defaults from the Intergovernmental Panel on Climate Change's (IPCC) [Guidelines for National Greenhouse Gas Inventories](#). Other sources, including the International Energy Agency (IEA), United Nations, EIA, US Environmental Protection Agency (EPA), and Carbon Dioxide Information Analysis Center (CDIAC), were also consulted.

These emissions factors are then modified by deducting net non-energy uses of each fuel. This is due to some proportion of the fuel produced being refined into products that effectively store carbon, such as various petrochemicals. Non-energy uses vary by a wide variety of factors, however, like other global emissions databases, a common factor must be applied for non-energy uses associated with each fuel type. While this factor is likely reasonably accurate on a global scale, it may not always precisely represent each specific entity's non-energy uses.

Applying this factor to the standardized production results in the emissions from the combustion of marketed products, comprising nearly 90% of total emissions tracked by the database. These are Scope 3 Category 11 emissions, corresponding to "use of sold products", however this has been modified to quantify emissions from each fossil fuel company's net production of oil, gas, or coal as opposed to sold products. This was done to avoid double counting, and deliberately excludes emissions from crude oil purchased from other producers, natural gas purchased for re-sale, or coal sold on behalf of other producers.

## Emission factors for the combustion of oil & natural gas liquids, natural gas, and coal<sup>2</sup>

Fuel type	Carbon factor	CO <sub>2</sub> factor
Oil & Natural Gas Liquids	101.4 kgC/bbl	371.4 kgCO <sub>2</sub> /bbl
Natural Gas	14.6 kgC/kcf	53.4 kgCO <sub>2</sub> /kcf
Lignite Coal	328.4 kgC/tonne	1,203.2 kgCO <sub>2</sub> /tonne
Sub-Bituminous Coal	495.1 kgC/tonne	1,814.1 kgCO <sub>2</sub> /tonne
Bituminous Coal	665.5 kgC/tonne	2,438.6 kgCO <sub>2</sub> /tonne
Anthracite Coal	715.4 kgC/tonne	2,621.5 kgCO <sub>2</sub> /tonne
Metallurgical Coal	727.4 kgC/tonne	2,665.4 kgCO <sub>2</sub> /tonne
Thermal Coal	581.0 kgC/tonne	2,128.9 kgCO <sub>2</sub> /tonne

Four further direct operational Scope 1 emission types are then estimated:

- Flaring of CO<sub>2</sub> at oil and gas facilities, including various upstream and midstream facilities, relevant to oil and gas production.
- Venting of CO<sub>2</sub> from natural gas processing plants, also relevant to oil and gas production.
- Fugitive methane emissions from coal mines, oil extraction and storage, and gas production, processing, and transportation systems, applicable to oil, gas, and coal production.
- CO<sub>2</sub> emissions resulting from entity's use of their own fuel, limited to gas production, primarily the difference between total gas produced and "gas available for sale".

## Emission factors for the vented, flared, and fugitive carbon dioxide and methane, and use of own fuel<sup>3</sup>

Fuel type	Combustion (kgCO <sub>2</sub> /tCO <sub>2</sub> )	Flaring (kgCO <sub>2</sub> /tCO <sub>2</sub> )	Venting (kgCO <sub>2</sub> /tCO <sub>2</sub> )	Fugitive methane (kgCH <sub>4</sub> /tCO <sub>2</sub> )	Fugitive methane (kgCO <sub>2</sub> /tCO <sub>2</sub> )	Own fuel use
Oil & Natural Gas Liquids	1,000	15.94	3.83	1.92	53.86	-
Natural Gas	1,000	1.74	28.53	9.88	276.59	57.26
Coal	1,000	-	-	4.03	112.97	-

<sup>2</sup> Emissions factors include deduction for non-energy uses.

<sup>3</sup> This analysis uses the IPCC AR5 100 year warming potential of 28 x CO<sub>2</sub> for methane.

## Cement Emissions

Estimation of CO<sub>2</sub> emissions for cement production differs from that for fossil fuel production. Cement-related emissions are estimated as a proportion of gross emissions reported by the major cement companies to the Cement Sustainability Initiative. This proportion of gross emissions estimates the process emissions from the calcining of limestone into clinker or Portland cement and excludes the emissions from fuel and electricity inputs, thus avoiding the double counting of fuels from fossil fuels producers already accounted for in Carbon Majors. From all these calculations, the database tracks the total emissions value in CO<sub>2</sub> equivalent units generated by each entity each year.

## Global Fossil Fuel & Cement Emissions

This research compares the emissions tracked by the Carbon Majors database to total fossil fuel and cement emissions since the beginning of the Industrial Revolution in 1751. Data from the Carbon Dioxide Information Analysis Center (CDIAC), and more recently the [Global Carbon Project](#), provides this total, amounting to 1,773 GtCO<sub>2</sub> from 1751 to 2022. The CO<sub>2</sub> emissions figures obtained from the above calculations (excluding fugitive methane CO<sub>2</sub> equivalent emissions) are compared to this total to calculate entities' relative contributions to total global fossil fuel and cement emissions.

## Data Dictionary

The Carbon Majors dataset is available for download as CSV files with 3 levels of granularity: low, medium, and high.

### Low granularity

column location	column name	description	unit	required	type	column location
emissions_low_granularity.csv	year	The year of the data point	-	yes	integer	emissions_low_granularity.csv
emissions_low_granularity.csv	parent_entity	The entity to whom the emissions are traced to	-	yes	string	emissions_low_granularity.csv
emissions_low_granularity.csv	parent_type	The type of the parent_entity. Can be one of: investor-owned company, state-owned entity, nation state.	-	yes	string	emissions_low_granularity.csv
emissions_low_granularity.csv	total_emissions_MtCO2e	The total emissions traced to the 'parent_entity' in the 'year'	Million tonnes of carbon dioxide equivalent (MtCO2e)	yes	float	emissions_low_granularity.csv

### Medium granularity

column location	column name	description	unit	required	type	column location
emissions_medium_granularity.csv	year	The year of the data point	-	yes	integer	emissions_medium_granularity.csv
emissions_medium_granularity.csv	parent_entity	The entity to whom the emissions are traced to	-	yes	string	emissions_medium_granularity.csv
emissions_medium_granularity.csv	parent_type	The type of the parent_entity. Can be one of: investor-owned company, state-owned entity, nation state.	-	yes	string	emissions_medium_granularity.csv
emissions_medium_granularity.csv	commodity	Specifies which commodity the production refers to: Oil & NGL, Natural Gas, Anthracite Coal, Bituminous Coal, Lignite Coal, Metallurgical Coal, Sub-Bituminous Coal, Thermal Coal, or Cement.	-	yes	string	emissions_medium_granularity.csv
emissions_medium_granularity.csv	production_value	The quantity of production	-	yes	float	emissions_medium_granularity.csv
emissions_medium_granularity.csv	production_unit	The unit of production (Oil & NGL - million barrels, Natural Gas - billion cubic feet, Coal - million tonnes, Cement - million tonnes CO2 (see methodology for explanation))	Billion cubic feet per year (Bcf/yr), Million barrels per year (Million bbl/yr), or Million tonnes per year (Million tonnes/yr)	yes	string	emissions_medium_granularity.csv



emissions_medium_granularity.csv	total_emissions_MtCO2e	The total emissions traced to the 'parent_entity' in the 'year'.	Million tonnes of carbon dioxide equivalent (MtCO2e)	yes	float	emissions_medium_granularity.csv
----------------------------------	------------------------	--	--	-----	-------	----------------------------------

## High granularity

column location	column name	description	unit	required	type	column location
emissions_high_granularity.csv	year	The year of the data point	-	yes	integer	emissions_high_granularity.csv
emissions_high_granularity.csv	parent_entity	The entity to whom the emissions are traced to	-	yes	string	emissions_high_granularity.csv
emissions_high_granularity.csv	parent_type	The type of the parent_entity. Can be one of: investor-owned company, state-owned entity, nation state.	-	yes	string	emissions_high_granularity.csv
emissions_high_granularity.csv	reporting_entity	The entity who reports the production. This can be different to the parent_entity if, for example, the parent_entity later acquires the reporting_entity.	-	yes	string	emissions_high_granularity.csv
emissions_high_granularity.csv	commodity	Specifies which commodity the production refers to: Oil & NGL, Natural Gas, Anthracite Coal, Bituminous Coal, Lignite Coal, Metallurgical Coal, Sub-Bituminous Coal, Thermal Coal, or Cement.	-	yes	string	emissions_high_granularity.csv
emissions_high_granularity.csv	production_value	The quantity of production	-	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	production_unit	The unit of production (Oil & NGL - million barrels, Natural Gas - billion cubic feet, Coal - million tonnes, Cement - million tonnes CO2 (see methodology for explanation))	Billion cubic feet per year (Bcf/yr), Million barrels per year (Million bbl/yr), or Million tonnes per year (Million tonnes/yr)	yes	string	emissions_high_granularity.csv
emissions_high_granularity.csv	product_emissions_MtCO2	Emissions from the combustion of marketed products. These are Scope 3 emissions, specifically Category 11 'use of sold products'.	Million tonnes of carbon dioxide (MtCO2)	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	flaring_emissions_MtCO2	Emissions from the flaring of CO2 at oil and gas facilities, including various upstream and midstream facilities, relevant to oil and gas production. These are Scope 1 emissions.	Million tonnes of carbon dioxide (MtCO2)	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	venting_emissions_MtCO2	Emissions from the venting of CO2 from natural gas processing plants, relevant to oil and gas production. These are Scope 1 emissions.	Million tonnes of carbon dioxide (MtCO2)	yes	float	emissions_high_granularity.csv

emissions_high_granularity.csv	own_fuel_use_emissions_MtCO2	CO2 emissions resulting from entity's use of their own fuel, limited to gas production, primarily the difference between total gas produced and "gas available for sale". These are Scope 1 emissions.	Million tonnes of carbon dioxide (MtCO2)	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	fugitive_methane_emissions_MtCH4	Fugitive methane emissions from coal mines, oil extraction and storage, and gas production, processing, and transportation systems, applicable to oil, gas, and coal production. These are Scope 1 emissions.	Million tonnes of methane (MtCH4)	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	fugitive_methane_emissions_MtCO2e	Fugitive methane emissions from coal mines, oil extraction and storage, and gas production, processing, and transportation systems, applicable to oil, gas, and coal production. These are Scope 1 emissions.	Million tonnes of carbon dioxide equivalent (MtCO2e)	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	total_operational_emissions_MtCO2e	This is the sum of the four scope 1 emissions (flaring, venting, own fuel use, fugitive methane. This value is given in units million tonnes of carbon dioxide equivalent.	Million tonnes of carbon dioxide equivalent (MtCO2e)	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	total_emissions_MtCO2e	The total emissions traced to the 'parent_entity' in the 'year'. This value is given in units million tonnes of carbon dioxide equivalent.	Million tonnes of carbon dioxide equivalent (MtCO2e)	yes	float	emissions_high_granularity.csv
emissions_high_granularity.csv	source	The source of the data point	-	yes	string	emissions_high_granularity.csv

## Glossary

### Units

*Note - Metric prefixes are used, e.g., k for the multiplier  $10^3$  or 1,000.*

MtCO<sub>2</sub> – Million tonnes of carbon dioxide.

MtCO<sub>2</sub>e – Million tonnes of carbon dioxide equivalent.

CH<sub>4</sub> - Methane

Bbl/bbls - Barrel or barrels.

Mbbl – Million barrels, the standard unit for Oil & NGL.

Bcf – Billion cubic feet, the standard unit for Natural Gas.

Kcf - Thousand cubic feet.

Mt – Million metric tonnes, the standard unit for Coal.

## Support

Please direct any questions and comments around the Carbon Majors data to [info@carbonmajors.org](mailto:info@carbonmajors.org). All other inquiries, including media, should be directed to the appropriate contact listed on the [Contact](#) page.

## Terms and Conditions

The Carbon Majors dataset is open source available for download as CSV files for non-commercial use. InfluenceMap's [Terms and Conditions](#) apply.