Insights on the spatial distribution of global, national and sub-national GHG emissions in EDGARv8.0

Authors: Monica Crippa[2], Diego Guizzardi[1], Federico Pagani[2], Marcello Schiavina[6], Michele Melchiorri[1], Enrico Pisoni[1], Francesco Graziosi[1], Marilena Muntean[1], Joachim Maes[5], Lewis Dijkstra[1,5], Martin Van Damme[3,4], Lieven Clarisse[3], Pierre Coheur[3]

Formatted: Italian (Italy)

[1]European Commission, Joint Research Centre (JRC), Ispra, Italy
[2]Unisystems S.A., Milan, Italy

[3]Spectroscopy, Quantum Chemistry and Atmospheric Remote Sensing (SQUARES), Université libre de Bruxelles (ULB), Brussels, Belgium

[4]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

[5]European Commission, Directorate-General for Regional and Urban Policy, Brussels

[6]NTT DATA, Rue de Spa, 8, 1000 Bruxelles

Formatted: Italian (Italy)

Correspondence: enrico.pisoni@ec.europa.eu

Abstract

To mitigate the impact of greenhouse gas and air pollutant emissions, it is of the utmost importance understanding where emissions happen. Atmospheric pollutants are emitted by a variety of sources which can be represented by point source information (e.g. power plants, industrial facilities, etc.), but also diffuse sources (e.g. residential activities, agriculture, etc.). However, emission inventories are typically compiled making use of country level statistics by sector, which are then downscaled at gridcell level making use of spatial information. In this work, we develop high-spatial resolution proxies used to downscale national emission totals for all world countries as provided by the Emissions Database for Global Atmospheric Research (EDGAR).

The latest EDGAR v8.0 GHG emissions provide readily available emission data at different spatial granularity, obtained from a consistently developed GHG emissions database. This is achieved through the improvement and development of high-resolution spatial proxies which allow a more precise allocation of emissions over the globe. A key novelty of this work is the possibility to analyse sub-national GHG emissions over the European domain, but also over the US, China, India and other high-emitting countries. These data answer not only the need of atmospheric modellers but at aim at informing policy makers acting in the field of climate change mitigation. For example, the EDGAR GHG emissions at NUTS2 level over Europe contribute to the development of EU Cohesion policies, identifying the progress of each region towards the carbon neutrality target, as well as providing insights on the most emitting sectors. The data can be accessed at https://doi.org/10.2905/b54d8149-2864-4fb9-96b9-5fd3a020c224 specific for EDGARv8.0 (Crippa, 2023a) and doi:10.2905/D67EEDA8-C03E-4421-95D0-0ADC460B9658 for the sub-national dataset (Crippa et al., 2023b).

~~To mitigate the impact of greenhouse gas and air pollutant emissions, it is of utmost importance understanding where emissions happen. Atmospheric pollutants are emitted by a variety of~~

sources which can be represented by point source information (e.g. power plants, industrial facilities, etc.), but also diffuse sources (e.g. residential activities, agriculture, etc.). However, emission inventories are typically compiled making use of country level statistics by sector, which are then downscaled at gridcell level making use of spatial information. In this work, we develop high spatial resolution proxies used to downscale national emission totals for all world countries as provided by the Emissions Database for Global Atmospheric Research (EDGAR).

Knowing where emissions occur is essential for planning effective emission reduction measures and for atmospheric modelling. Emission inventories are typically compiled at national level and provide sector-specific emission estimates. Disaggregating national emissions on high-resolution grids requires spatial proxies that contain information on the location of different emission sources (e.g. point sources, linear and area sources). Knowing the correct allocation of emissions from point sources is essential to avoid the misallocating high emission levels. However, gathering information on point sources covering the entire globe and a wide temporal domain (1970 to present) is challenging due to limited data availability, accuracy of the reporting and completeness of data. The latest EDGAR v8.0 spatial proxies developed as part of the Emissions Database for Global Atmospheric Research (EDGARv8.0) provide the userreadily available emission data at different spatial granularity, with the possibility to work with different geographical details using obtained from aa consistently developed GHG emissions database. A key novelty of EDGARv8.0 is the possibility to analyse sub-national GHG emissions over the European domain, but also over the US, China, India and other high-emittingmain world countries. For example, the EDGAR GHG emissions at NUTS2 level over Europe contribute to the development of EU Cohesion policies, identifying the progress of each region towards the carbon neutrality target, as well as the most emitting sectors. The relevance of using updated spatial information is assessed on the basis of regional case studies. The data can be accessed at https://doi.org/10.2905/b54d8149-2864-4fb9-96b9-5fd3a020c224 specific for EDGARv8.0 (Crippa, 2023a) and doi:10.2905/D67EEDA8-C03E-4421-95D0-0ADC460B9658 for the sub-national dataset (Crippa et al., 2023b).

1 Introduction

Knowing where emissions are released happen is essential to support the design of effective mitigation actions and for atmospheric modelling purposes. Emission inventories are typically developed at the national level and provide sector-specific emission estimates. In order to disaggregate national emissions over high-resolution grids, information on the location of the different emission sources (e.g. point, linear and area sources) must be collected and 'spatial proxies' should be developed and applied to national sector specific emission totals to downscale them over gridmaps. The correct allocation of point source emissions is essential to avoid misplacing high emission levels. However, gathering information on point sources covering the entire globe and a wide temporal domain (1970 to present) is challenging due to limited data availability, accuracy in the reporting (real location vs. legal address, etc.) and completeness of data.

The Emissions Database for Global Atmospheric Research (EDGAR) provides global greenhouse gas (GHG) and air pollutant emissions over the global gridmap at 0.1x0.1 degree resolution, obtained through a downscaling process of national emissions using high-resolution, although the resolution of the underlying spatial data. information used to

~~downscale national totals may be higher (down to few hundred meters resolution).~~ The development and maintenance of the EDGAR gridmaps is essential since several regional and global databases rely on the EDGAR emission gridmaps to disaggregate national emissions to the grid. ~~weight national inventories.~~ This is the case of the Community Emissions Data System (CEDS) (Feng et al., 2020; Hoesly et al., 2018) or the EMEP Centre on Emission Inventories and Projections (CEIP) to support Parties to the LRTAP Convention ~~EU Member States~~ in their official gridded emission reporting requirements (CEIP, 2021). ~~This work is an update of previous EDGAR publications dealing with spatial data (Janssens-Maenhout et al., 2019; Crippa et al., 2021), and describes all the new developments for the spatialisation of the emissions from EDGARv8.0 onwards.~~

~~Knowing t~~~~The correct allocation of point source emissions is essential to avoid the misplacement of~~~~misplacing~~ ~~high emission levels. However, gathering information on point sources covering the entire globe and a wide temporal domain (1970 to present) is challenging due to limited data availability, accuracy in the reporting (real location vs. legal address~~~~site, etc.) and completeness of data. The latest spatial proxies developed within EDGAR will be presented in this work, focusing on high emitting sectors such as power plant and industrial activities, but also on more distributed sources such as residential activities.~~

~~This work is an update of previous EDGAR publications dealing with spatial data (Janssens-Maenhout et al., 2019; Crippa et al., 2021), and describes all the new developments for the spatialisation of the emissions from EDGARv8.0 onwards, focusing on high emitting sectors such as power plants and industrial activities, but also on more diffuse sources such as residential activities. High resolution spatial information has been gathered at the global level combining Global Energy Monitor data, official registries and satellite retrievals. The relevance of using updated spatial information is also assessed with regional case studies.~~

~~regardingpatterns,(2023; 2011; 2006)A,~~

This work is an update of previous EDGAR publications dealing with spatial data (Janssens-Maenhout et al., 2019; Crippa et al., 2021), and describes all the new developments for the spatialisation of the emissions from EDGARv8.0 onwards, focusing on high emitting sectors such as power plants and industrial activities, but also on more diffuse sources such as residential activities. High -resolution spatial information has been gathered at the global level combining Global Energy Monitor data, official registries and satellite retrievals. The relevance of using updated spatial information is also assessed with regional case studies.

The purpose of this publication is describing the EDGARv8.0 GHG gridded emission datasets, focusing on the updates of the spatial proxies included in this data release. The analysis of EDGARv8.0 emission time series (European Union, 2023; IEA-EDGAR CO2, 2023) and the methodology behind emission calculations is available in Crippa et al. (2023).

The m~~M~~ain novelties of this work are i) an update of emission point sources using global datasets (e.g. Global Energy Monitor), ii) the development of a gap-filling method for non-population based sources using built-up surface information for non-residential areas[1] from the

[1] This information is compliant with the definition of 'building' as per the 'Infrastructure for Spatial Information in Europe', INSPIRE directive, https://inspire.ec.europa.eu/id/document/tg/bu) for non-residential areas (i.e. industrial or commercial facilities, warehouses, etc.) from the Global Human Settlements Layer (GHSL).

Formatted: English (United Kingdom)

Global Human Settlements Layer (GHSL), iii) an update of population based proxies using the latest GHSL data including a weight for meteorological dependence of heating needs, and v) an update of international ship tracks and weights by vessel type. In addition, information at sub-national level (e.g. for Europe at NUTS2 level) is included when developing the new spatial proxies of EDGAR, thus allowing a more accurate allocation and analysis of sub-national emissions. The EDGARv8.0 GHG global emission maps can be accessed at doi:10.2905/D67EEDA8-C03E-4421-95D0-0ADC460B9658 for the subnational emissions, and at doi: 10.2905/B54d8149-2864-4FB9-96B9-5FD3A020C224 for v8.0 for the emission gridmaps at 0.1x0.1 degree resolution.

2 Overview on the methodology and data sources used for updating spatial information in EDGAR

Bottom-up global inventories , (such as EDGAR), compute emissions for each sector, pollutant and year at the national country level, making use of international statistics and official guidelines for emission computation (Janssens-Maenhout et al., 2019; Crippa et al., 2018). However, atmospheric modellers, policy makers, local authorities and scientists may need to analyse spatially distributed emissions at a higher resolution than country- -level data. Therefore, annual country specific emissions are distributed over the globe making use of spatial information, representing either the exact location of points sources (e.g. power plants, industrial facilities, etc.), or linear tracks (e.g. road network, ship and airplane tracks, etc.), andor area sources (e.g. populated areas, industrial areas, etc.). Within the EDGAR database, over 130 proxy datasets (f) varying over time are developed to weightdistribute the contribution of sector- specific emissions ($EM_{i,j,k}$) of each country (C) and pollutant (x) over time (t) to each grid cell ($em_{i,j,k}$) at 0.1°x0.1° resolution (about 10km at the equator) spatial resolution (WGS84, EPSG:4326), with the Heaviside function (i.e. unit step function whose value is zero for negative arguments and 1 for positive arguments), equalling 1 when the grid cell belongs to the country area, accordingly with the following formula:

$$emi_{i,j,k}\,(lon,lat,t,x) = EM_{i,j,k}(C,t,x)\cdot\frac{f_{i,j,k}(lon,lat,t)}{\sum_{lon,lat}(f_{i,j,k}(lon,lat,t)\cdot H_{i,j}(C,lon,lat))}$$

Where

Hi,j(C, lon, lat) = fraction/weight of gridcell within C,

i=sector,

j=fuel,

k=technology.

Table 1 summarises the data sources and the methodology used to update spatial information for each emitting sector in the EDGAR database, highlighting the most relevant and latest updates compared to previous EDGAR data releases. These updates apply from EDGARv8.0 onwards. Being a global database of emissions, the spatial data sources used are typically developed at the global level (e.g. satellite based retrievals, etc.), although but often relying on national data collections (e.g. national point- -source information reported to fulfill legal requirements). Therefore, the same data sources may be used by other inventory developers to

Formatted: Font: 12 pt

update their spatial disaggregation of the emissions. In the following sections, a detailed description on the data sources and the approach used for updating each emission sector is provided, distinguishing between point sources, area sources and linear sources. For all sectors not subjected to a recent revision in the EDGAR database, we ~~recommend~~ refer the reader to ~~rely on~~ the overview Table S1 and ~~there~~ references therein.

A key methodological advancement in the EDGAR gridding system is ~~also represented by~~ the inclusion of ~~the correct~~ sub-national attributes~~information~~ for each spatial proxy~~data~~ and in particular for each point source. This implies attaching to each point not only its exact location expressed in longitude and latitude, but also the related NUTS2 (Nomenclature of territorial units for statistics) code (EUROSTAT, 2021) for Europe or the Global ADMinistrative layer at level 1 (GADM version 4.1). The choice of including NUTS2 rather than NUTS3 information aims to ~~at~~ enhanc~~ing~~e the capability of a global database such as EDGAR to represent sub-national regional emissions in support of the development of regional policies (e.g. EU Cohesion Reports (European Commission, 2022) or the 2040 Climate Impact Assessment)~~,~~ ~~while compromising with the global dimension of the database~~. T~~In fact, t~~he attribution of subnational details is not only developed with an EU-~~-~~oriented focus, but also for other countries such as the United States, China, and India, by providing emissions at the state or province level.~~but also it allows approaching for example the Unites States, China and India not anymore at national level administrative boundary, but providing emissions on each US state, each Chinese province and Indian state~~ ~~Moving towards~~

~~province or city scale dimension starting from national emissions is~~ requires ~~not only subjected~~ associating ~~to the association of e.g. point sources to NUTS3 level but also relying on more disaggregated statistics. Therefore, considering the current purposes of EDGAR the NUTS2 level represents the right balance between accuracy of the final emissions and downscaling of national totals. The relevance of including not only country specific details, but also sub-regional information is essential when doing emission data extraction at sub-national level, thus avoiding border issues. Some inventory compilers (Kuenen et al., 2022), report point source information just as points without distributing them over a gridmap with a certain resolution. This approach is accurate since it provides the exact geographical coordinates of individual facilities; however, it does not reduce data extraction issues, since the allocation of a specific point to a certain gridcell cell may fall between the borders of e.g. two regions. Another challenge that we address with this new gridding approach is related with the harmonization of national and sub-national data. Local and regional inventories are often developed independently, therefore, undermining the possibility to collate together sub-national emissions to retrieve the national values. The challenge of using different and not coherent databases is overtaken by the EDGAR database, being able to consistently work both at the national and regional level, thus offering the user the possibility to work across different geographical scales.~~ The purpose of our work is to provide readily available emissions at sub-national level estimated in a consistent way for all countries. The EDGAR data may represent an approximation for those countries with developed statistical infrastructure (e.g. those including sub-national statistics and very precise spatial proxies), however, they provide a default if such data is not available, as it is the case for many countries in the world. In the results section, case studies on sub-national emissions are presented for the EU, US, China and India.

3 Point sources of emissions

Gathering information on point sources covering the globe and spanning a wide temporal domain (1970-Present~~nowadays~~) is challenging due to the limited data availability, accuracy and completeness in the reporting (real plant location vs. legal address~~site~~, etc.). The correct location of point sources is essential since they are often super emitters (e.g. power plants for CO_2 emissions). In EDGARv8.0, the location of the main industrial point sources (e.g. power plants, iron and steel industries~~ and other plants~~, coal mines, venting and flaring activities, etc.), which contribute for around half of global CO_2 emissions, has been updated using state of the art information making use of global databases, such as the Global Gas/Coal Plant Tracker of the Global Energy Monitor. A complete overview of the data sources and updates included in EDGARv8 is provided in Table 1. ~~In the following, we will describe sector by sector how the most up to date spatial data on point sources have been collected and implemented in the EDGAR database to downscale national emissions over the global gridmap.~~

However, point source databases are characterised by some limitations, in terms of completeness of the point sources, availability of time series of information, misplacement of data points compared to the real country belonging, etc. In EDGAR v8.0, quality checks procedures are applied to validate the correct location of each point source to the corresponding country or sub-national attribute. Moreover, missing information is completed using assumptions on the time life of power plants (i.e. 40 years) to indicatively attribute opening or closing years for each plant.

No consistency check between CO2 emissions estimated through independent methods has been here performed. However, Guevara et al. (2024) have proven the good agreement between national CO2 emissions from power plants as reported by EDGAR (which is based on international statistics) and plant level inventories.

Atmospheric modellers require information not only regarding the spatial patterns of the emissions, but also on the temporal and vertical distribution, as described in Ahsan et al. (2023), Bieser et al. (2011) and De Meij et. al. (2006). For example, De Meij et al. (2006) found that an important role is played by the vertical distribution of SO2 and NOx emissions in understanding the differences between emission inventories on calculated gas and aerosol concentrations. Accordingly, with the EMEP model, industrial point sources and power plants emissions are injected up to the third level (top up to 184 m), while shipping emissions happen in the first level (top up to 20 m). However, addressing the vertical distribution of the emissions in beyond the purpose of this work. In the following, we will describe sector by sector how the most up to date spatial data on point sources have been collected and implemented in the EDGAR database to downscale national emissions over the global gridmap.

3.1 Power plants

Power plants represent a major source of fossil CO_2 and GHG emissions globally, contributing nowadays for around 38% and 18%, respectively, to the corresponding global totals (Crippa et al., 2023). It is therefore of utmost importance ~~including the latest available information~~ to correctly spatially allocate these emissions at the global level and understand their evolution over time, in order to design and implement adequate emission mitigation measures.

In EDGARv8.0, fuel-specific spatial proxies have been developed using data from the Global Coal and Gas Plant Tracker of the Global Energy Monitor (for coal and gas) (Global Energy

Monitor, 2022d, a), the Global Power Plant Database v1.3.0 (World Resources Institute, 2018; WRI, 2021) for oil and biofuels, CARMAv3.0 for autoproducers (i.e. plants and industries producing power for their own use). In addition, information on autoproducers and biofuel-fired power plants in Europe has been integrated using the European Pollutant Release and Transfer Register (EPRTRv18) (EPRTR, 2020). For the US domain, the location of fossil fuel -fired power plants is taken from the US Energy Information Administration (US EIA, 2022b) as they represent the most updated source for the US. The time frame covered by the new power plant spatial proxy datasets developed in EDGARv8.0 is 1970-2022, which includes, for each plant, information on opening and closing years (also beyond 2022 for recently built power plants), capacity, main fuel type, etc. When only partial information is available for the years of operations, assumptions on the typical lifetime of power plants is ~~assumed~~ made (e.g. 40 years). The capacity of each power plant is used to relatively weigh~~t~~ within a country the fuel specific emissions from power plants. An additional adjustment is performed over the US domain, to ~~take into~~ account for the different sulphur content in the fuel used in the different US states based on EIA and FERC utility surveys.

The Global Energy Monitor is chosen as the main data source for updating power plant~~s~~ proxies since it relies on data from public and private data sources (including the Global Energy Observatory, CARMA, Platts World Energy Power Plant database, national-level trackers developed by environmental organisations, as well as various company and government sources). It is validated with i) government data on individual power plants;~~,~~ ii) country energy and resource plans, and government websites tracking coal plant permits and applications;~~,~~ iii) reports by state-owned and private power companies;~~,~~ iv) news and media reports; and~~,~~ v) local non-governmental organizations tracking coal plants or permits. Local experts are also involved in the review of coal and gas plant data. Regular bi-annual updates of these databases also guarantee the possibility to include further updates in future EDGAR releases. As of January 2019, the Global Coal Plant Tracker included exact locations for 95.3% of operating units (6411 out of 6725). Independent use and validation of the Global Coal and Gas Plant Trackers is also performed by Guevara et al. ~~(2023~~, 2024). Figure S1 shows the comparison between the geo-coverage of EDGARv8.0 and the previous EDGAR spatial data for power plants, while Fig. S2 provides a view ~~on~~ of the global coverage of power plants in EDGARv8.0 by fuel type.

Figure 1 shows the global coverage and intensity of CO_2 emissions from fossil fuel-~~-~~fired power plants from EDGARv8.0 for the years 1970 and 2022. As a general trend, the number of power plants ~~highly~~ strongly increased from 1970 to 2022 (see also Fig.2) due to the global industrialisation process ~~happened~~ over the past five~~5~~ decades ~~at the global level~~, although the number of power plants in 1970 is ~~subjected to higher uncertainty compared to nowadays situation.~~ more uncertain than that of the present day.

~~.~~The total number of power plants grew from around 8500 in 1970 to 13000 in 2022, with the sharpest increase occurring in China (4.5 times more) and North America (2 times more). However, the intensity of the emissions changed over the past 5 decades, depending on the region. As shown in Fig.2, despite the increase in the regional number of power plants, the shift towards cleaner fuels ~~is found~~ in historically industrialised regions (such as Europe and North America) together with increased energy efficiency, ~~which lead~~has led to stable and lower CO_2 emissions (e.g. 13% decrease in Europe between 1970 and 2022). On the contrary, emerging

Formatted: Indent: Left: 0 cm

regions are characterised by significantly higher emissions in 2022 and the use of high C-content fuels, such as coal. Over the past five decades, fossil CO_2 emissions from power plants increased up to 42 and 38 times in China and India, respectively. Country-specific trends of CO_2 and GHG emissions from power plants are presented in Crippa et al. (2023).

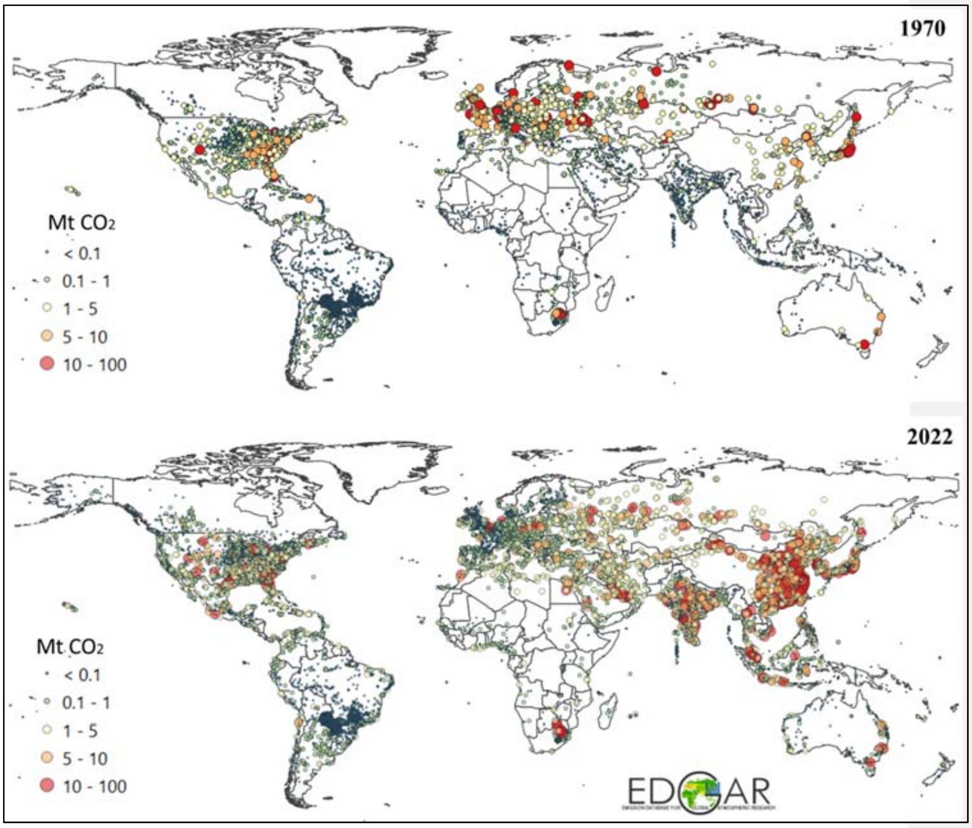

Figure 1 – CO_2 emissions from fossil fuel fired power plants in 1970 and 2022 from EDGARv8.0. The size of the circles is proportional to the magnitude of the emissions.

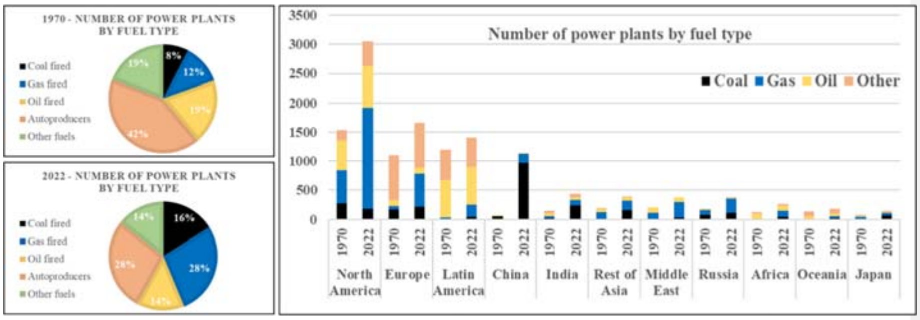

Figure 2 - Evolution of the total number of power plants (including fossil and bio fuels fired) from 1970 to 2022 by world region included in the updated EDGAR spatial proxies.

3.2 Industrial facilities and other point sources

Industrial activities cover a wide range of sectors encompassing ~~manufacturing,~~ the production of iron and steel, cement, glass, metals, ~~solvents,~~ chemicals, ~~or~~ fertilisers, -use of solvents, but also intensive animal farming (see section 3.4). Gathering information on industrial activities (e.g. production, capacity, location of the facilities, etc.) at the global level is challenging, also due to confidentiality and data protection issues. For this reason, we ~~partly~~ focussed not only on the update of information on industrial point sources (when available), but also on the ~~while we~~ improvement of ~~d~~ the gap-filling method for all industrial activities in case of incomplete or missing data (as discussed in detail in Sect. 3.5). In EDGARv8.0, we included the latest European Pollutant Release and Transfer Register (EPRTRv18) locations for all industrial facilities (with the exception of power plants, iron and steel facilities and coal mines, for which dedicated spatial proxies have been developed at the global level). Several manual adjustments were implemented to overcome data quality issues related with missing spatial information and inconsistencies. The analysis of the EPRTR dataset also inspired the idea of attributing only a fraction of the emissions to the reported point sources. This is also justified by the fact that industrial facilities have to report their emissions only if they fall above a certain threshold. The fraction of the emissions to be allocated to the available point sources is determined through the ratio between EPRTR emissions (typically of CO2) and the corresponding EDGAR emissions. When the ratio is 1, all emissions are allocated to the point sources; when the ratio is lower than 1, the complementary fraction is then attributed to the gap-filling grid (i.e. non-residential proxy as defined in Sect. 3.5).

In EDGARv8.0, we have also updated the global locations of iron and steel plants, which are among the most energy intensive industries. The Global steel plant tracker of the Global Energy Monitor (2022b) was used as a data source due to its global and temporal completeness (1970-present). The installed capacity was used to weigh the relative contribution of each iron and steel plant, although it may represent an approximation for the real capacity in use. A map of iron and steel production plants in 1970 and 2022 is presented in Fig.3. The number of iron and steel plants increased around tenfold over the last five decades (from 77 to 728) with the sharpest increase in China (fivefold), USA and India (2.7-fold).

Coal Mines are also a relevant source of fugitive emissions of GHGs and air pollutants (e.g. volatile organic compounds). In EDGARv8.0, we updated the information on coal mines at the global level using the Global Coal Mine Tracker of the Global Energy Monitor (2022c) complemented with the Energy Information Administration data for the US (US EIA, 2022a). For countries not covered by these data sources, we relied on the previous EDGAR spatial proxies including data from the US Geological Survey (USGS, 2019). More specifically, we included information on surface and underground mines, both for hard and brown coal.

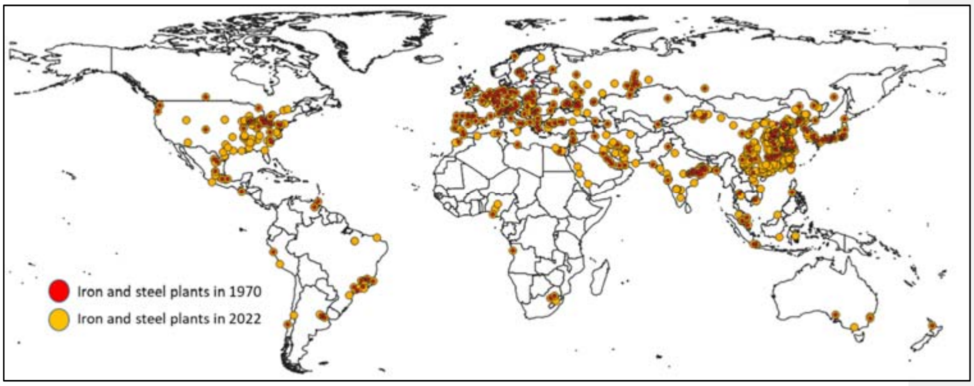

Figure 3 – Global location of iron and steel plants in 1970 and 2022.

3.3 Venting and flaring

Gas flaring is the burning of the natural gas ~~associated~~ that results ~~with~~ from oil extraction. Although this practice is highly polluting and represents a waste of resources, it is still in place due to economic constraints and the lack of appropriate legislation in several countries. Flaring takes place both as on-shore and off-shore activities and it is a source of GHG and air pollutant emissions.

Global CO2 emissions related with flaring account for 276 Mt in 2022, of which 76~~50~~% is emitted by 10 ~~only by four~~ countries, namely Russia (18% of the global total), Iraq (13%), Iran (12%) and Venezuela (7%), followed by~~. 76% of the global CO_2 emissions from flaring activities is produced by 10 top emitting countries with individual contribution higher than 2% of the global total (including~~ Algeria, USA,~~;~~ Mexico, Libya, Nigeria and China~~ in addition to the abovementioned top 4)~~. Although this emission source represents only 0.8% of global CO_2 emissions, it is particularly relevant for certain regions in the world, such as Venezuela (20% of the CO_2 country total), Iraq (18%), Libya (17%), Algeria (10%) and Nigeria (9%). Considering the relevance of venting emissions and the potential of control measures, it is essential to best quantify and attribute the correct georeference for this source. Flaring emissions can also be localised and quantified through space born measurements (Elvidge et al., 2017; NOAA, 2017). In EDGARv8.0, data from the World Bank Global Gas Flaring Tracker Report (2023) were used both for estimating the emissions and location of global flaring activities from 2012 to~~ill~~ 2022. These spatial data were also used as a best approximation to spatially distribute emissions from venting, which is the controlled release of natural gas without being burned, although the two activities may not overlap. The resulting CO_2 emission map in 2012 and 2022 is reported in Fig. 4.

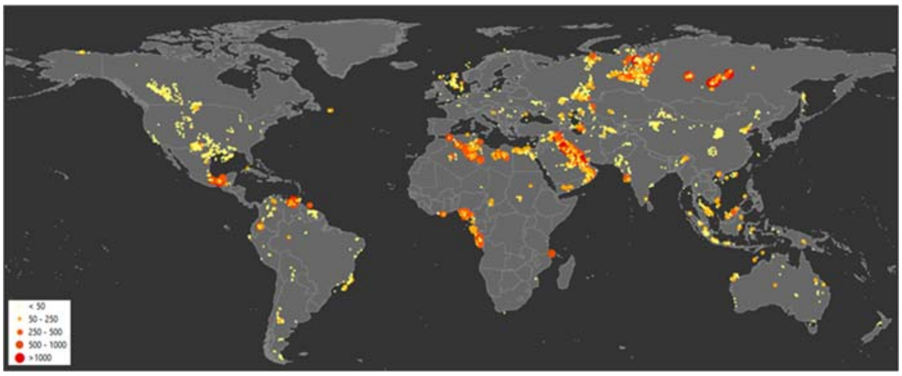

Figure 4 – Global map of CO2 emissions (kton) from flaring in 2022.

3.4 Intensive livestock and fertiliser industries

Agriculture includes a variety of activities that are typically distributed over large areas (e.g. crop areas, animal pastures, etc.). However, several agricultural activities can be defined as hot-spots or point sources and include intensive animal farming and manure management practices. In a broader sense, we allocate to this sector also fertiliser production industries which represent an important source of NH_3 and N_2O. In EDGARv8.0, the IASI satellite-derived NH_3 point source database (Van Damme et al., 2018; Clarisse et al., 2019) is included to map animal farming and fertiliser production emissions with yearly information -for the period 2008-2022. It includes 270 agricultural hot-spots and 251 production facilities of synthetic NH_3 worldwide. Since the NH_3 point source database includes only hot-spots we decided to allocate to these points only a fraction of the total emissions for that sector and country derived from approximate estimates of NH_3 emission flux from IASI measurements, while distributing the remaining fraction to livestock density maps formerly available in EDGAR. Similarly to what was done for other industries, for Europe, intensive livestock point sources and fertiliser production industries were taken from EPRTRv18. Similarly, the satellite-based information on fertiliser industries was integrated in the previous EDGAR proxy for this sector. This update represents a significant improvement in representing N related hot-spots ~~compared to former EDGAR releases~~ (Van Damme et al., 2018) compared to former EDGAR releases which mostly used animal density as proxy (see Table S1), although considering the uncertainty of IASI information of around 50%. A snapshot on N_2O emissions from manure management at global level and in Europe, where intensive livestock activities appear as emission hot-spots is shown in Fig. 5.

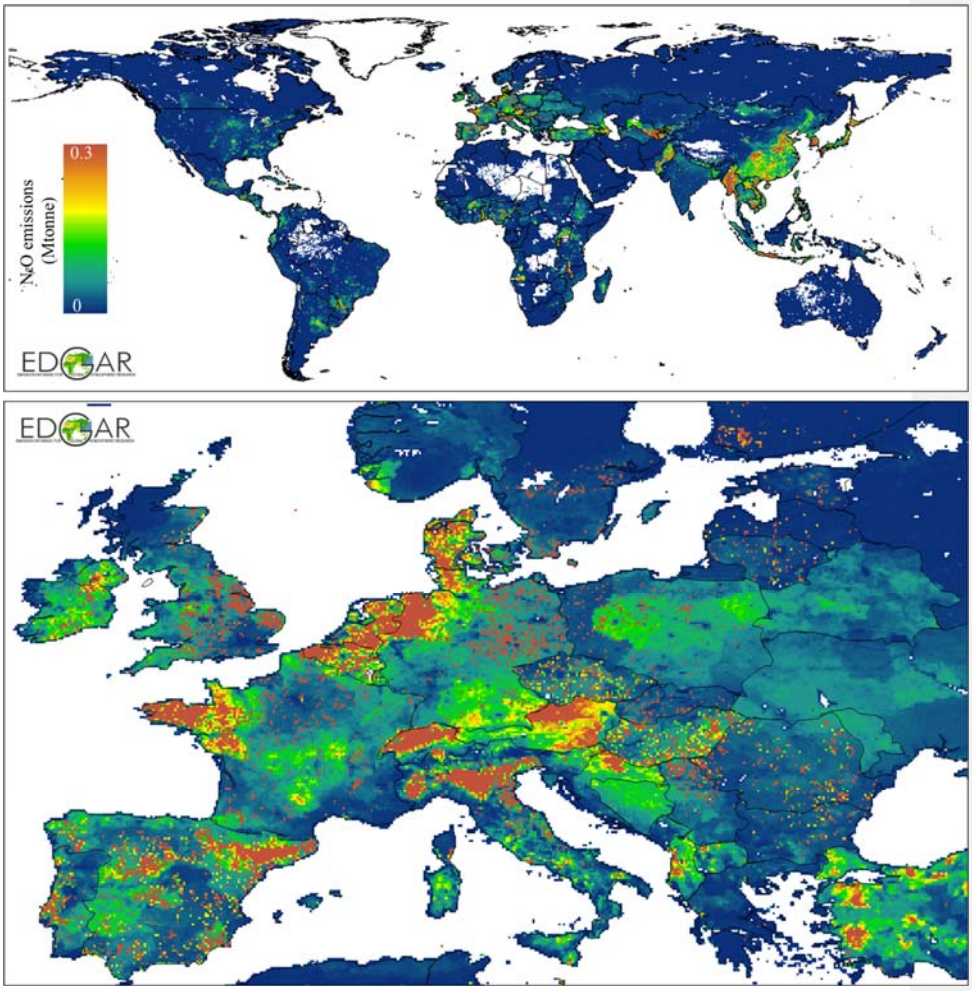

Figure 5 – N_2O emissions from manure management at global level and in Europe, where intensive livestock activities appear as emission hot-spots.

3.5 Gap-filling missing information of point sources

A significant improvement is represented by the development and use of a new spatial proxy to gap-fill missing information for all industrial related emissions. Until EDGARv7.0, population-related proxies were used as backup information when no spatial data was available to represent the emissions for a sector within a country (Crippa et al., 2021). However, here we decided to use the non-residential built-up surface information developed by the Global Human Settlements Layer (GHSL) (Pesaresi and Politis, 2023; European Commission, 2023) as a backup proxy to distribute the emissions of all the activities not related ~~with~~ to small-scale combustion for which no point source information was available (even for individual countries). This methodological assumption is a key novelty of this work due to its application at the global level. However, it is in line with methodologies already applied in regional inventories, such as in Europe (Kuenen et al., 2022), where the CORINE land-use dataset is used to spatially allocate emissions to areas with industrial activity, thus supporting the validity of this assumption.

For certain sectors and regions, this non-residential gap-filling proxy is also used to allocate a fraction of the emissions of a certain sectors (refer for example to the industrial facilities section for Europe). The overall effect of using this new proxy is a change in the industrial contribution over densely populated areas which was previously higher in EDGAR compared to other inventories in particular over Europe (Thunis et al., 2023). Figure 6 shows CO_2 emission maps from manufacturing industries obtained in EDGARv7.0 and EDGARv8.0. This comparison figure highlights the implications of using different gap-filling proxies for the industrial sector, and in particular contrasts those based on population (EDGARv7.0) and with the new ones based on non-residential built-up surface data used in EDGARv8.0.

Overall, using non-residential built-up information to allocate emissions of industrial activities to complement point source information leads to lower emission levels allocated to urban areas and a less densely distributed map over certain regions (e.g. China, India, etc.). Figure S3 shows the impact of this update on global fossil CO2 emissions from the industrial sector over global Functional Urban Areas (FUAs) in 2022. The share of CO2 industrial emissions to the national total ratio between these emissions over FUAs is typically higher, on average by around 30%, in EDGARv8.0 than in EDGARv7.0 for several developing countries (e.g. Africa, South America, India, etc.) due to the presence of industrial point sources and non-residential activities still close to urban areas. On the oppositeOn the other hand, lower (on average around 20% less) emissions from industries (on average around 20% less) –are found in many industrialised regions (e.g. Europe, USA, Oceania) due to the displacement of industrial activities in remote areas or outside the FUAs. This result represents the effect of using non-population based proxies for industrial emissions in EDGARV8.0 compared to previous EDGAR proxies.

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

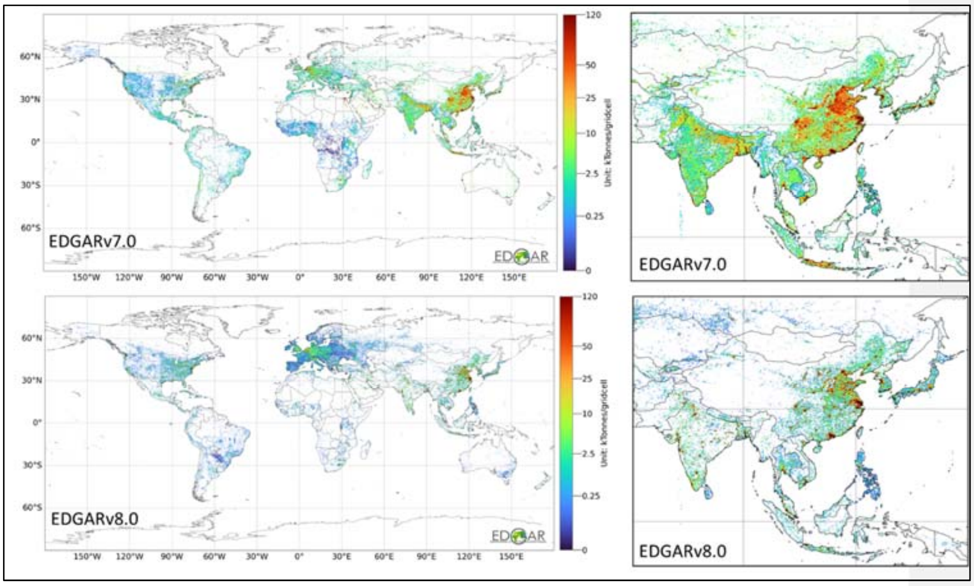

Figure 6 – CO2 emissions from industrial combustion in 2021 from EDGARv7.0 and v8.0, showing the impact of the gap-filling proxies used for industrial sources.

4 Linear sources of emissions: international shipping

Since EDGARv6.0, international shipping emissions are distributed using the STEAM3 (Ship Traffic Emission Assessment Model) model from the Finnish Meteorological Institute (Jalkanen et al., 2012; Johansson et al., 2017) and ~~the same spatial distribution is kept also~~this approach has remained unchanged in EDGARv8.0. Emissions are distributed on a yearly basis from 2000 to~~ill~~ 2018, including multi vessels information (cargo, container, fishing, passenger cruisers, service, tankers, vehicle carriers, miscellaneous). Compared to the previous EDGAR proxy, the use of the STEAM data allows a better representation of the evolution in time of the international shipping emissions, differentiating on ~~yearly~~ an annual basis the variation of the routes and their intensity for the different vessels consistently with the information available in EDGAR (see Fig. 7). Only data covering sea areas are included, since inland data over big rivers or lakes is not robust ~~yet~~ enough to be included in EDGAR. Information on Emission Control Areas (ECAs), and in particular on sulphur emission control areas (SECAs) and NO_x emission control areas (NECAs), are not yet included, ~~while it represents one of the~~although this may be considered for ~~is planned for~~ future updates of EDGAR. A comparison between international shipping intensities ~~as~~ that are available in EDGAR before and after this update is presented in Fig. S4 of the Supplement.

Figure 8 focus~~s~~es on three main vessel types, representing the largest fraction of GHG emissions from international shipping in 2022 and contributing specifically for around 22% (tankers), 24% (containers) and 28% (cargo) to total international shipping GHG emissions. The impact of using the STEAM data to develop the new spatial proxies for international shipping is shown in Fig. 8, where the comparison between EDGARv5 and EDGARv8 CO2 emissions from the three main vessel types over the different Oceans and Seas is presented.

Formatted: Default

EDGARv5 used an in-house EDGAR proxy based on Wang et al. (2008), improved with LRIT (Long-Range Identification and Tracking) information (Alessandrini et al., 2017) for European seas, as described in Janssens-Maenhout et al. (2019). EDGARv5 proxies were allocating most of the international shipping emissions over the Atlantic and Pacific Oceans, while the new proxies of EDGARv8 allocate the largest portion of these emissions (40%) over the Seas around China, Japan and Philippines. The relative share of tanker~~s~~ emissions over the Mediterranean Sea is also very different between the two versions, with the largest contribution (85%) among the three considered categories in EDGARv5. ~~Two times higher e~~Emissions ~~are also~~ allocated to the Gulf of Mexico and Arabian Sea are two times higher when using the STEAM based proxies in EDGARv8.

Formatted: Font: 12 pt

Formatted: Font: 12 pt

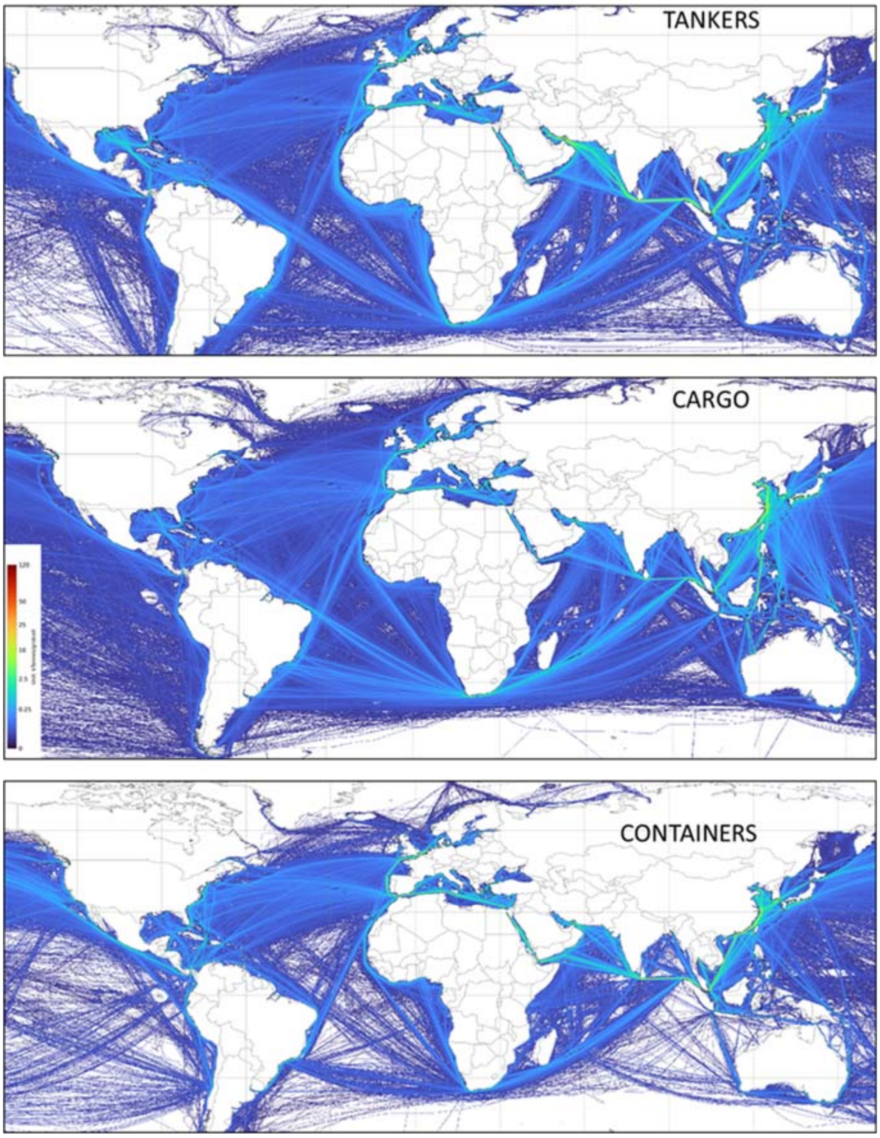

Figure 7 – International shipping GHG emissions (2021) with the ship tracks for tankers, containers and cargo vessels as in EDGARv8.0.

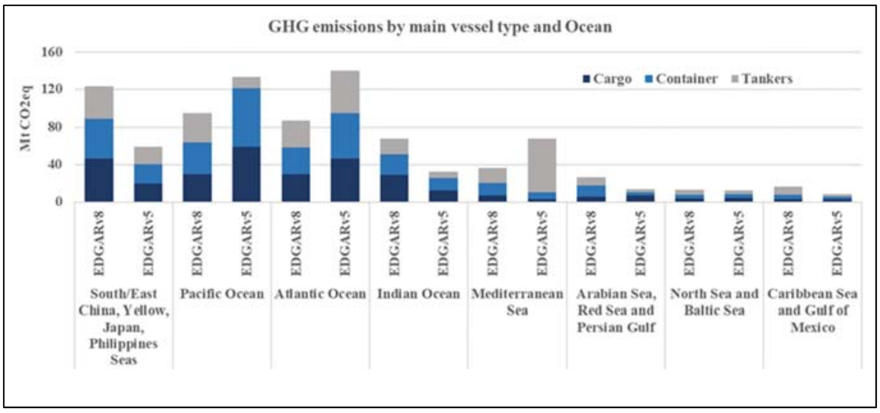

Figure 8 – Comparison of GHG emissions from international shipping (2022) by main vessel type and Ocean from EDGARv5 and EDGARv8. Fishing, services and passenger related emissions are excluded from this comparison.

5 Area sources of emissions

5.1 Residential activities

Small-scale combustion emissions are mostly related with non-industrial activities, such as those from the residential, commercial and agricultural/fishing sectors. Therefore, population based spatial proxies are often used to downscale national emissions. EDGARv8.0 aims to~~at~~ ~~coupling~~ couple population distribution with heating degree days since the amount of emissions is not only dependent on the number of people living over certain areas, but also on the meteorological conditions and the heating needs for indoor spaces. Residential emissions are therefore distributed considering both population intensities and heating needs, with varying profiles from 1970 to 2022. EDGARv8.0 includes the latest population gridmaps developed by the Global Human Settlements GHS-POP R2023A (Schiavina et al., 2023b; Freire et al., 2016), which comprise residential population information for 12 epochs, ~~starting from~~over 1975~~- to~~ 2020 with five~~5~~-year~~s~~ time steps and projections to 2025 and 2030 obtained by distributing census data from CIESIN GPWv4.11 over global gridmaps~~ census data from CIESIN GPWv4.11~~. GHS-POP R2023A data at 30 arc-seconds (WGS84, EPSG:4326) (or about 1km) spatial resolution were used to develop the corresponding spatial proxies in EDGAR. Population density is then calculated for each gridcell and it is used as a proxy to allocate household emissions over populated areas. Small-scale combustion activities related with agriculture are distributed using rural population maps obtained from the GHS-SMOD R2023 product (including only low and very low density rural grid cells) (Schiavina et al., 2023a). For missing years, the closest population map to each epoch is taken (e.g. for the years 2001 and 2002 the population map ~~is the one of~~from 2000 is used, while for the years 2003 and 2004 ~~it is the one of~~the 2005 map is used).

To account for the effect of the weather (ambient temperature) on heating needs in the residential sector, heating degree days (HDD) have been computed using the 2 meters temperature data with hourly time resolution and 1 degree spatial resolution using the Copernicus ERA5 atmospheric reanalysis produced by ECMWF for the years 1970-2022 (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form). HDD is the cumulative number of degrees by which the mean daily temperature falls below a reference temperature (usually 18 °C or 19 °C which is adequate for human comfort). HDD were calculated following the methodology described by Spinoni et al. (2018) and assuming a reference temperature of 18°C. Cooling Degree Days (CDD) are not included in the development of the spatial proxies since they are mainly related with electricity consumption rather than to fuel combustion in the residential sector. An additional weight to the population distribution is therefore added by the HDD metric, thus increasing the emissions arising in colder regions subjected to more heating needs rather than in warm areas for the same amount of population.

Our approach does not aim to~~at~~ identify~~ing~~ and represent~~ing~~ the heating habits for all countries, while modulating within a single country the combustion of fuels for e.g. heating purposes due to the different temperatures across latitudes (climatic zones). Countries may in fact have ~~in fact~~ different habits in turning on and off their heating systems, thus requiring the use of different reference temperature values in the calculation of HDD (Atalla et al., 2018) which is not taken into account here. The process to build the residential proxy in EDGAR is shown in Fig. 9.

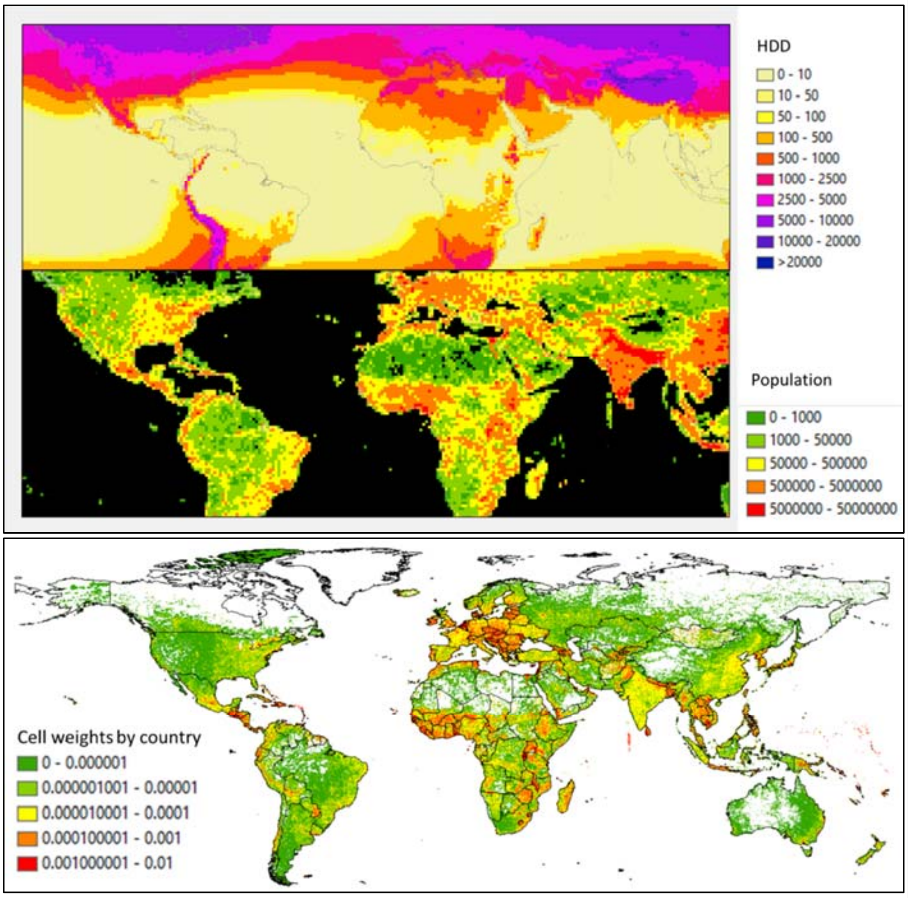

Figure 9 – Coupling heating degree days (a) and population density (b) as a proxy (c) to downscale residential emissions. Data refer to the year 2020.

6 Results

The purpose of this work is to describe the methodological improvements included in EDGARv8.0 linked to the update of the spatial data used to downscale country and sector specific emissions. In addition, a specific focus is dedicated to case studies showing the relevance of understanding the evolution of GHG emissions at sub-national level in order to support the development of regional climate mitigation and adaptation policies (Kuramochi et al., 2020). Therefore, the reader can refer to Crippa et al. (2023) for the description of country and sector-specific GHG emission trends at global level. In the following sections, insights on the global distribution of GHG emissions as well as their sub-national features are described.

6.1 Global GHG emissions in EDGARv8.0

Figure 10 shows global GHG emissions in 2022 as a result of the EDGARv8 gridding process, while Figure 11 reports the same emissions at the country and sub-national level. Complementary figures are also reported in the Supplement (Figs. S5-S8) showing the evolution of GHG, fossil CO_2, CH_4 and N_2O global emission maps from 1970 to 2022.

The main strength and novelty of EDGARv8.0 is related with the production of a global GHG emission database at different level of granularity in support of local, regional and global climate actions. The high spatial resolution global maps are available at 0.1°x0.1° WGS84 (EPSG4326), about 10km at the equator, both as emissions and emission fluxes (.txt and .NetCDF files, https://edgar.jrc.ec.europa.eu/dataset_ghg80) fulfilling the requirements of the global atmospheric modelling community but also bridging bottom-up and top-down (mostly satellite based) GHG emission estimates (see Fig. 10).

EDGARv8.0 allows full flexibility in the aggregation of emissions at the sub-national level, thus supporting the analysis of the spatio-temporal variability of the emissions not only at gridcell level but also over wider administrative domains, or areas of interest such as urban centres (Melchiorri, 2022). A second key product from EDGARv8.0 is represented by GHG emissions at sub-national level using the Global ADMinistrative layer version 4.1 (https://gadm.org/download_country.html) at level 1 and NUTS2 level for the EU extended geographical domain, as shown in Fig. 11. ~~In the next sections, case studies over the European, American and Asian domains are discussed more in detail.~~

Looking at province or city scale emissions requires not only associating e.g. point sources to NUTS3 level but also relying on a different approach from the downscaling of national totals, which may include the use of statistical information available over smaller territorial units. Therefore, considering the current purposes of EDGAR the NUTS2 level represents the right balance between accuracy of the final emissions and downscaling of national totals. The relevance of including not only country specific details, but also sub-regional information is essential when doing emission data extraction at sub-national level, thus avoiding border issues. Some inventory compilers (Kuenen et al., 2022), report point source information just as points without distributing them over a gridmap with a certain resolution. This approach is accurate since it provides the exact geographical coordinates of individual facilities; however, it does not reduce data extraction issues, since the allocation of a specific point to a certain grid cell may fall between the borders of e.g. two regions.

Another challenge that we address with this new gridding approach is related with the harmonization of national and sub-national data. Local and regional inventories are often developed independently, therefore, undermining the possibility to collate together sub-national emissions to retrieve the national values. The challenge of using different and not coherent databases is overtaken by the EDGAR database, being able to consistently work both at the national and regional level, thus offering the user the possibility to work across different geographical scales. This is achieved through the downscaling of national emissions to sub-national data making use of high-spatial resolution proxies, as discussed in this paper. In the next sections, case studies over the European, American and Asian domains are discussed more in detail.

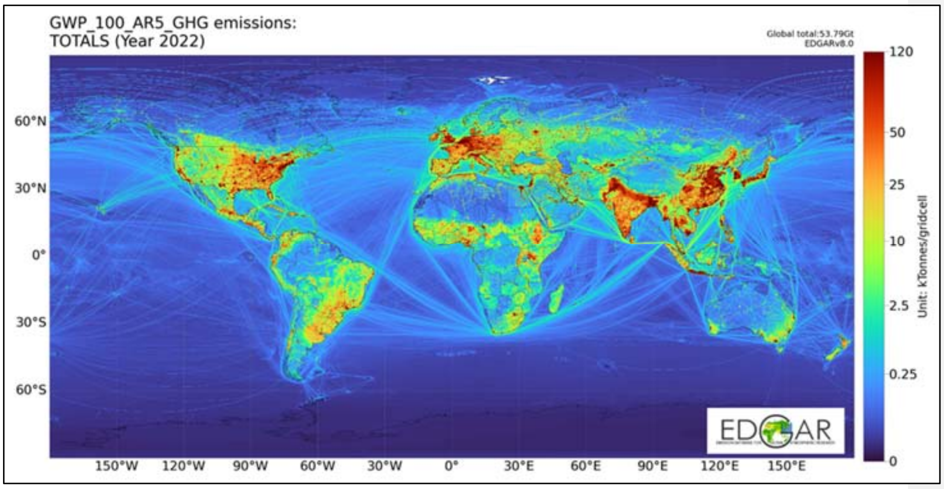

Figure 10 – Global GHG (expressed in CO2eq) emission map in 2022 from EDGARv8.0.

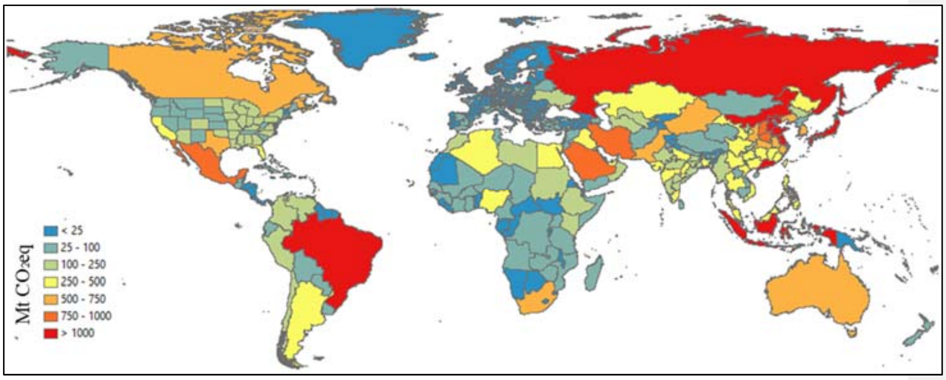

Figure 11 – Global GHG emissions by country and sub-national level in 2022 based on EDGARv8.0.

6.2 Sub-national emissions: the EU case

Climate and environmental territorial policies require robust and consistent knowledge of greenhouse gas (GHG) and air pollutant emissions at the sub-national level (e.g. NUTS2). No sub-national official reporting is available and the high spatial resolution data of EDGAR fill this knowledge gap. EDGAR sub-national GHG emissions are used as a reference by the European Commission in Cohesion Reports (European Commission, 2022), the EU semester process or Climate Action territorial analysis. Figure 12 shows how GHG emissions at NUTS2 level have changed from 1990 to 2021 both in absolute, per capita and per GDP terms. Out of 242 EU regions, 155 regions have shown a downward trend since 1990, while it is found for 206 and 204 regions since 2005 (on average -1.27% per year) and 2010 (on average -1.35% per year), respectively. However, in 2021, only 34 regions reached less than 5t CO2eq/person which corresponds to the average value needed to achieve the 2030 EU climate targets. The most contributing sectors to total EU GHG emissions in 2021 are ~~the~~ power generation (27%), industry (23%), transportation (20%), buildings (14%) and agriculture (11%), showing that the different regions in the EU have different transition challenges. For example, when looking at the NUTS2 level (see Fig. 12, middle bottom panel) the transport sector often represents the sector with the largest contribution at regional level, in particular in rural regions of Spain, France, Italy, or Germany. Figure 12 (bottom right panel) also shows the share of GHG emissions arising from small-scale combustion (buildings sector) at the NUTS2 level, highlighting several regions for which this sector contributes more than 15-20% to the regional total.

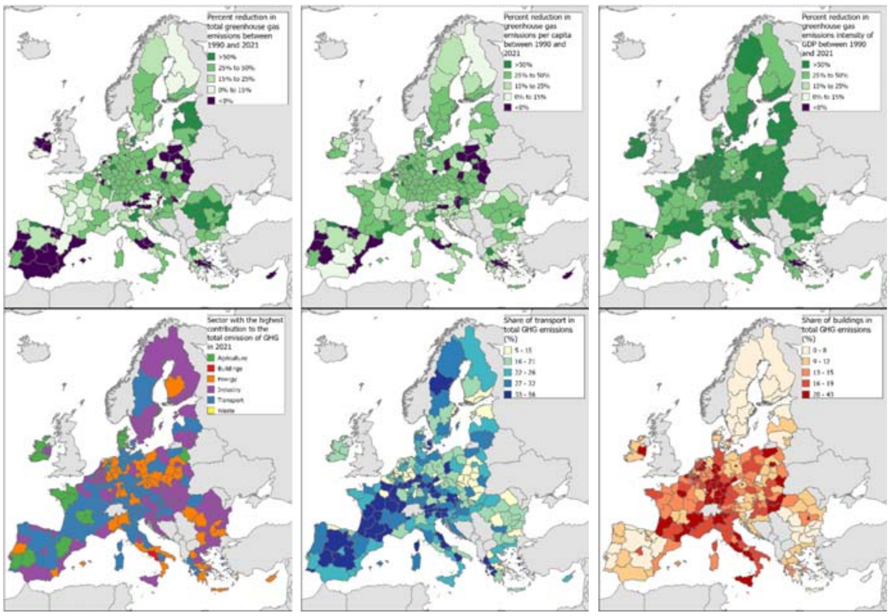

Figure 12 – Relative change of European GHG emissions by NUTS2 between 1990 and 2021 (top panels). Sector contribution of European GHG emissions by NUTS2 in 2021 (bottom panels). The sector with the highest contribution in 2021 for each NUTS2 is shown in the map on the left panel. The share of GHG emissions from transport (middle panel) and buildings (right panel) to total emissions in 2021 in Europe by NUTS2 is also shown.

6.3 Sub-national emissions in the United States, China and India

EDGARv8.0 includes GHG emission estimates at the sub-national level also for the United States (i.e. estimates for each US state, Fig. 13), for each Chinese province and each Indian state (Fig. 14). Based on our analysis, Texas emits 11.5% of the total US GHG emissions in

2022, followed by California with a contribution of 7.7% and Florida with a share of 4.6%. ~~Also i~~In 1990, Texas and California were the most emitting states, followed by Ohio, Pennsylvania and Illinois. Over the past three~~3~~ decades, the sector with the highest share of GHGs at state level over the US has changed, with a shift from power and industry towards transport (see Fig. 13).

In 2022, the five~~5~~ most emitting Chinese provinces contributed ~~for~~ to around 40% of the Chinese total GHG emissions. These were ~~and they were~~ Shandong (8.9% of the country total), Guangdong (8.4%), Jiangsu (7.4%), Hebei (6.6%) and Nei Mongol (6.5%), consistent~~ly~~ with ~~some~~ other literature studies addressing ~~province level~~provincial CO2 and GHG emissions in China (Jiang et al., 2019; Zhang et al., 2020). In 1990, the top five~~5~~ emitting provinces were Shandong (8.1%), Hebei (6.5%), Jiangsu (6.2%), Henan (5.9%) and Nei Mongol (5.8%) contributing ~~for~~ around 30% to the Chinese total GHG emissions.

~~-~~In 2022, five~~5~~ Indian states emitted around 50% of the country total GHG emissions, namely Maharashtra (11.8%), Tamil Nadu (11.7%), Uttar Pradesh (8.1%), Gujarat (8.0%) and~~,~~ Chhattisgarh (6.6%). In 1990, the most emitting Indian states were Tamil Nadu (18.4%), Maharashtra (9.5%), Uttar Pradesh (9.3%), West Bengal (6.6%) and~~,~~ Andhra Pradesh (6.0%). Compared to the US and European cases, a different picture is found over the Asian domain in terms of ~~most~~ top-emitting sectors at sub-national level (Fig. 14). The effect of the economic growth and the transition from an agricultural ~~based~~ towards a more industrialised economy can be seen in Fig. 14 (right panels). As a result, the sectors with the highest share changed from agriculture (in 1990) to energy and industry (in 2022) over China and India, with the exception of some few regions (e.g. Tamil Nadu, Assam, Jammu and Kashmir, Uttarakhand) which ~~kept~~ still had an agriculture~~-~~-based economy ~~also~~ in 2022. This type of information and analysis is instrumental for the definition of effective sector~~-~~-specific climate mitigation actions at the sub-national level.

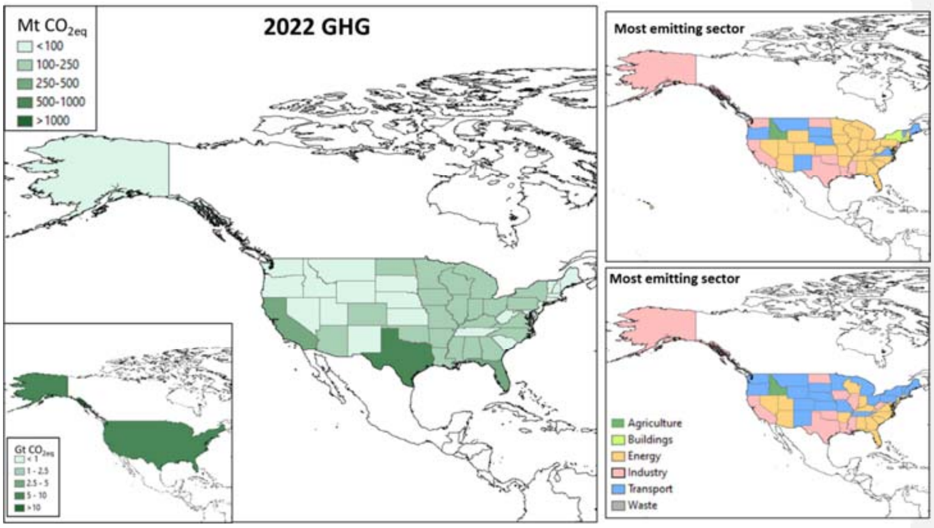

Figure 13 – 2022 GHG emissions at sub-national level in the United States are represented left panel and the sector with the highest contribution in 1990 and 2022 for each US state is shown in the maps on the right.

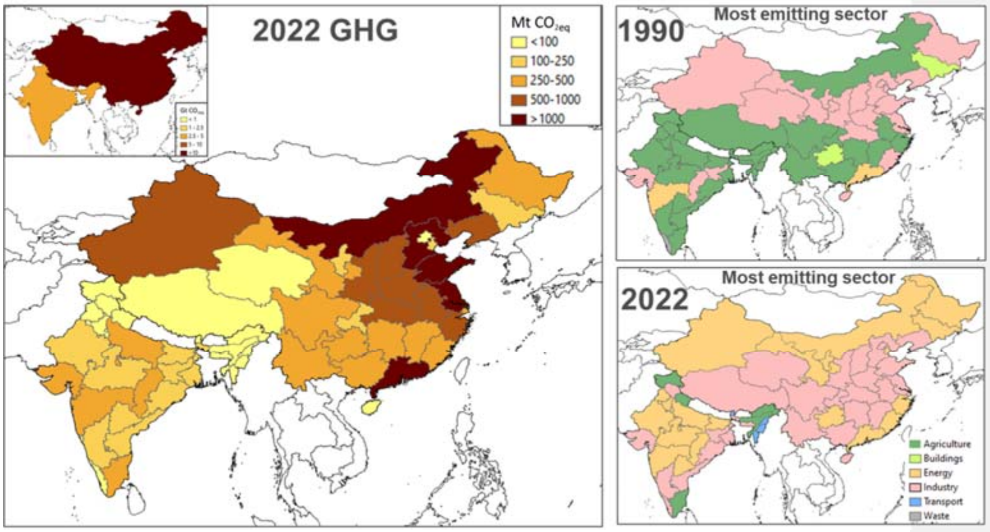

Figure 14 – 2022 GHG emissions at sub-national level over the Asian domain, with focus on China and India, (left panel) and the sector with the highest contribution in 1990 and 2022 for each Chinese and Indian province/state is shown in the maps on the right.

7 Data availability

The EDGARv8.0 GHG global emission maps can be freely accessed at https://doi.org/10.2905/b54d8149-2864-4fb9-96b9-5fd3a020c224 (Crippa, 2023a). The EDGARv8.0 subnational emissions can be accessed at doi:10.2905/D67EEDA8-C03E-4421-95D0-0ADC460B9658 (Crippa et al., 2023b). All data can also be accessed through the EDGAR website at https://edgar.jrc.ec.europa.eu/dataset_ghg80 and https://edgar.jrc.ec.europa.eu/dataset_ghg80_nuts2 (last access: November 2023).

Data are made available as emission gridmaps for each speciessubstance and for total GHGs as .txt and .nc files with emissions expressed in ton substance/0.1degree x 0.1degree/year. Emission fluxes are available as .nc files and they are expressed in kg substance/m2/s. Emission maps are available both as total and sector specific emissions.

8 Conclusions

Climate targets are often set at the global and national level, however their but implementation may occur at the subnational level, but also in many cases at national level.the implementation of mitigation actions occurs at local and regional level. It is therefore of the utmost relevance developing to develop sub-national GHG emission estimates for policy development and to monitor the progress towards climate targets or to evaluate their impacts. This work summarises the main updates developed within the Emissions Database for Global Atmospheric Research (EDGAR) for what concerns the use of high resolution and up to date spatial information to improve the global geospatial disaggregation of GHG emissions at sub-national level. Having accurate and up to date sector-specific GHG emission global maps at high spatial resolution (0.1x0.1 degrees) is instrumental for the design of effective climate mitigation options beyond (inter)national climate targets. EDGARv8.0 spatial proxies include globally consistent spatial data derived for example from the Global Energy Monitor, the Global Human Settlements Layer work, satellite based information to compute heating degree days or to identify hot-spots from agricultural activities, the STEAM model for ship track and many other global datasets. The use of satellite data to improve the EDGAR spatial proxies represents a successful cooperation between bottom-up inventory compilers and the Earth observation community, and the possibility to integrate relevant satellite based datasets and statistical information. In addition, EDGARv8.0 integrates spatial information from local databases (e.g. EPRTR for Europe, EIA data for the US) when including more detailed data ~~compared to~~than ~~what is~~that available in global databases.

Continuous updates and improvements of the spatial data used to downscale national emissions over the global grid are required to best represent the evolution of emission sources and their location. The strength and uniqueness of the EDGAR work are associated with its global coverage and consistency in computing and representing emissions for all countries, thus becoming a reference for many countries with limited capabilities for emissions estimation. However, several challenges are associated with the use of global databases of information, in particular dealing with the collection of point sources. Therefore, the use of local data, if available, is recommended when performing analysis at the highest spatial resolution (e.g. at city scale level, etc.).

A further improvement within EDGAR is related with the inclusion of sub-national information, representing a unique feature to address in a consistent way the evaluation of spatial patterns in the evolution of sub-national GHG emissions. Such spatial resolution and sub-national sector specific variability sets the ground for the production of city level emission data records, as used for example in the Urban Centre Database (https://ghsl.jrc.ec.europa.eu/ghs_stat_ucdb2015mt_r2019a.php). In this paper, few case studies are presented, with main focus on the European case where the EDGAR sub-national data are regularly used as input for the EU Semesters and contribute to climate action territorial and cohesion policies through the EU Cohesion Reports.

~~Continuous updates and improvements of the spatial data used to downscale national emissions over the global grid are required…limiations of the work…~~

9 Acknowledgements

~~We are grateful to the EDGAR team (M. Crippa, D. Guizzardi, E. Schaaf, M. Muntean, F. Pagani, M. Banja, W. Becker and F. Monforti-Ferrario) for the work needed to publish the EDGARv8.0 greenhouse gas emission datasets (https://edgar.jrc.ec.europa.eu/dataset_ghg80).~~

We are grateful to William Becker for the thorough review and proofreading of this manuscript. The views expressed in this publication are those of the author(s) and do not necessarily reflect the views or policies of the European Commission. All emissions, except for CO2 emissions from fuel combustion, are from the EDGAR (Emissions Database for Global Atmospheric Research) Community GHG database comprising IEA-EDGAR CO2, EDGAR CH4, EDGAR

N2O and EDGAR F-gases version 8.0 (2023). IASI-NH3 catalogue was updated in the framework of the ESA World Emission project (https://www.world-emission.com/). The ULB also gratefully acknowledges support from the TAPIR project (Air Liquide Foundation).

10 References

Ahsan, H., Wang, H., Wu, J., Wu, M., Smith, S. J., Bauer, S., Suchyta, H., Olivié, D., Myhre, G., Matsui, H., Bian, H., Lamarque, J. F., Carslaw, K., Horowitz, L., Regayre, L., Chin, M., Schulz, M., Skeie, R. B., Takemura, T., and Naik, V.: The Emissions Model Intercomparison Project (Emissions-MIP): quantifying model sensitivity to emission characteristics, Atmos. Chem. Phys., 23, 14779-14799, 10.5194/acp-23-14779-2023, 2023.

Alessandrini, A., Guizzardi, D., Janssens-Maenhout, G., Pisoni, E., Trombetti, M., and Vespe, M.: Estimation of shipping emissions using vessel Long Range Identification and Tracking data, Journal of Maps, 13, 946-954, 10.1080/17445647.2017.1411842, 2017.

Atalla, T., Gualdi, S., and Lanza, A.: A global degree days database for energy-related applications, Energy, 143, 1048-1055, https://doi.org/10.1016/j.energy.2017.10.134, 2018.

Bieser, J., Aulinger, A., Matthias, V., Quante, M., and Denier van der Gon, H. A. C.: Vertical emission profiles for Europe based on plume rise calculations, Environmental Pollution, 159, 2935-2946, https://doi.org/10.1016/j.envpol.2011.04.030, 2011.

CEIP: Inventory Review 2021 Review of emission data reported under the LRTAP Convention, https://www.ceip.at/fileadmin/inhalte/ceip/00_pdf_other/2021/inventoryreport_2021.pdf, Last Access: August 2023., 2021.

Clarisse, L., Van Damme, M., Clerbaux, C., and Coheur, P. F.: Tracking down global NH3 point sources with wind-adjusted superresolution, Atmos. Meas. Tech., 12, 5457-5473, 10.5194/amt-12-5457-2019, 2019.

Crippa, M., Guizzardi, D., Pagani, F., and Pisoni, E.: GHG Emissions at sub-national level, European Commission, Joint Research Centre (JRC) [Dataset] doi:10.2905/D67EEDA8-C03E-4421-95D0-0ADC460B9658 PID: http://data.europa.eu/89h/d67eeda8-c03e-4421-95d0-0adc460b9658, 2023b.

Crippa, M., Guizzardi, D., Pisoni, E., Solazzo, E., Guion, A., Muntean, M., Florczyk, A., Schiavina, M., Melchiorri, M., and Hutfilter, A. F.: Global anthropogenic emissions in urban areas: patterns, trends, and challenges, Environmental Research Letters, 16, 074033, 10.1088/1748-9326/ac00e2, 2021.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J. A., Monni, S., Doering, U., Olivier, J. G. J., Pagliari, V., and Janssens-Maenhout, G.: Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4.3.2, Earth Syst. Sci. Data, 10, 1987-2013, 10.5194/essd-10-1987-2018, 2018.

Crippa, M., Guizzardi, D., Pagani, F., Banja, M., Muntean, M., Schaaf, E., Becker, W., Monforti-Ferrario, F., Quadrelli, R., Risquez Martin, A., Taghavi-Moharamli, P., Köykkä, J., Grassi, G., Rossi, S., Brandao De Melo, J., Oom, D., Branco, A., San-Miguel, J., and Vignati, E.: GHG emissions of all world countries, Publications Office of the European Union, Luxembourg, doi:10.2760/953322, JRC134504, 2023.

Crippa, M., Guizzardi D., Pagani F., Banja M., Muntean M., Schaaf E., Becker, W., Monforti-Ferrario F., Quadrelli, R., Risquez Martin, A., Taghavi-Moharamli, P., Grassi, G., Rossi, S., Brandao De Melo, J., Oom, D., Branco, A., San-Miguel, J., Vignati, E.: EDGAR v8.0 Greenhouse Gas Emissions, European Commission, Joint Research Centre (JRC) [Dataset] doi: 10.2905/b54d8149-2864-4fb9-96b9-5fd3a020c224 PID: http://data.europa.eu/89h/b54d8149-2864-4fb9-96b9-5fd3a020c224, 2023a.

de Meij, A., Krol, M., Dentener, F., Vignati, E., Cuvelier, C., and Thunis, P.: The sensitivity of aerosol in Europe to two different emission inventories and temporal distribution of emissions, Atmos. Chem. Phys., 6, 4287-4309, 10.5194/acp-6-4287-2006, 2006.

Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C., and Ghosh, T.: Supporting international efforts for detecting illegal fishing and GAS flaring using viirs, 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 23-28 July 2017, 2802-2805, 10.1109/IGARSS.2017.8127580,

EPRTR: E-PRTR database v18, https://www.eea.europa.eu/data-and-maps/data/member-states-reporting-art-7-under-the-european-pollutant-release-and-transfer-register-e-prtr-regulation-23/european-pollutant-release-and-transfer-register-e-prtr-data-base/eprtr_v9_csv.zip, 2020.

European Commission: Cohesion in Europe towards 2050 - Eighth report on economic, social and territorial cohesion, doi: 10.2776/624081, 2022.

European Commission: GHSL Data Package 2023, Publications Office of the European Union, Luxembourg, JRC133256, doi:10.2760/098587, 2023.

European Union: European Commission, Joint Research Centre (JRC), EDGAR (Emissions Database for Global Atmopheric Research) Community GHG database, comprising IEA-EDGAR CO2, EDGAR CH4, EDGAR N2O and EDGAR F-gases version 8.0 (2023). Unless otherwise noted, all material owned by the European Union is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence. This means that reuse is allowed, provided that appropriate credit is given and any changes are indicated, 2023.

EUROSTAT: https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts, 2021.

Feng, L., Smith, S. J., Braun, C., Crippa, M., Gidden, M. J., Hoesly, R., Klimont, Z., van Marle, M., van den Berg, M., and van der Werf, G. R.: The generation of gridded emissions data for CMIP6, Geosci. Model Dev., 13, 461-482, 10.5194/gmd-13-461-2020, 2020.

Freire, S., MacManus, K., Pesaresi, M., Doxsey-Whitfield, E., and and Mills, J.: Development of new open and free multi-temporal global population grids at 250 m resolution, Geospatial Data in a Changing World, Association of Geographic Information Laboratories in Europe (AGILE), 2016.

Global Energy Monitor: Global Gas Plant Tracker, https://globalenergymonitor.org/projects/global-gas-plant-tracker/, 2022a.

Global Energy Monitor: Global steel plant tracker, https://globalenergymonitor.org/projects/global-steel-plant-tracker/, 2022b.

Global Energy Monitor: Global Coal Mine Tracker, https://globalenergymonitor.org/projects/global-coal-mine-tracker/, 2022c.

Global Energy Monitor: Global Coal Plant Tracker, https://globalenergymonitor.org/projects/global-coal-plant-tracker/, 2022d.

Guevara, M., Enciso, S., Tena, C., Jorba, O., Dellaert, S., Denier van der Gon, H., and Pérez García-Pando, C.: A global catalogue of CO2 emissions and co-emitted species from power plants at a very high spatial and temporal resolution, Earth Syst. Sci. Data Discuss., 2023, 1-41, 10.5194/essd-2023-95, 2023.

Guevara, M., Enciso, S., Tena, C., Jorba, O., Dellaert, S., Denier van der Gon, H., and Pérez García-Pando, C.: A global catalogue of CO2 emissions and co-emitted species from power plants, including high-resolution vertical and temporal profiles, Earth Syst. Sci. Data, 16, 337-373, 10.5194/essd-16-337-2024, 2024.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J. I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369-408, 10.5194/gmd-11-369-2018, 2018.

IEA-EDGAR CO2: A component of the EDGAR (Emissions Database for Global Atmospheric Research) Community GHG database version 8.0 (2023) including or based on data from IEA (2022) Greenhouse Gas Emissions from Energy, www.iea.org/data-and-statistics, as modified by the Joint Research Centre, 2023.

Jalkanen, J. P., Johansson, L., Kukkonen, J., Brink, A., Kalli, J., and Stipa, T.: Extension of an assessment model of ship traffic exhaust emissions for particulate matter and carbon monoxide, Atmos. Chem. Phys., 12, 2641-2659, 10.5194/acp-12-2641-2012, 2012.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., Petrescu, A. M. R., Solazzo, E., and Oreggioni, G. D.: EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012, Earth Syst. Sci. Data, 11, 959-1002, 10.5194/essd-11-959-2019, 2019.

Jiang, J., Ye, B., and Liu, J.: Peak of CO2 emissions in various sectors and provinces of China: Recent progress and avenues for further research, Renewable and Sustainable Energy Reviews, 112, 813-833, https://doi.org/10.1016/j.rser.2019.06.024, 2019.

Johansson, L., Jalkanen, J.-P., and Kukkonen, J.: Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution, Atmospheric Environment, 167, 403-415, https://doi.org/10.1016/j.atmosenv.2017.08.042, 2017.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J. P., Super, I., and Denier van der Gon, H.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling, Earth Syst. Sci. Data, 14, 491-515, 10.5194/essd-14-491-2022, 2022.

Kuramochi, T., Roelfsema, M., Hsu, A., Lui, S., Weinfurter, A., Chan, S., Hale, T., Clapper, A., Chang, A., and Höhne, N.: Beyond national climate action: the impact of region, city, and business commitments on global greenhouse gas emissions, Climate Policy, 20, 275-291, 10.1080/14693062.2020.1740150, 2020.

Melchiorri, M.: The global human settlement layer sets a new standard for global urban data reporting with the urban centre database, 10, 10.3389/fenvs.2022.1003862, 2022.

NOAA: Visible Infrared Imaging Radiometer Suite (VIIRS), https://www.ngdc.noaa.gov/eog/viirs.html, Latest Access: July 2023, 2017.

Pesaresi, M. and Politis, P.: GHS-BUILT-S R2023A - GHS built-up surface grid, derived from Sentinel2 composite and Landsat, multitemporal (1975-2030), European Commission, Joint Research Centre (JRC), http://data.europa.eu/89h/9f06f36f-4b11-47ec-abb0-4f8b7b1d72ea, doi:10.2905/9F06F36F-4B11-47EC-ABB0-4F8B7B1D72EA, 2023.

Schiavina, M., Melchiorri, M., and Pesaresi, M.: GHS-SMOD R2023A - GHS settlement layers, application of the Degree of Urbanisation methodology (stage I) to GHS-POP R2023A and GHS-BUILT-S R2023A, multitemporal (1975-2030), European Commission, Joint Research Centre (JRC), PID: http://data.europa.eu/89h/a0df7a6f-49de-46ea-9bde-563437a6e2ba, doi:10.2905/A0DF7A6F-49DE-46EA-9BDE-563437A6E2BA, 2023a.

Schiavina, M., Freire, S., Carioli, A., and MacManus, K.: GHS-POP R2023A - GHS population grid multitemporal (1975-2030). European Commission, Joint Research Centre (JRC), http://data.europa.eu/89h/2ff68a52-5b5b-4a22-8f40-c41da8332cfe, doi:10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE, 2023b.

Spinoni, J., Vogt, J. V., Barbosa, P., Dosio, A., McCormick, N., Bigano, A., and Füssel, H. M. J. I. J. o. C.: Changes of heating and cooling degree-days in Europe from 1981 to 2100, 38, e191-e208, https://doi.org/10.1002/joc.5362, 2018.

Thunis, P., Kuenen, J., Pisoni, E., Bessagnet, B., Banja, M., Gawuc, L., Szymankiewicz, K., Guizardi, D., Crippa, M., Lopez-Aparicio, S., Guevara, M., De Meij, A., Schindlbacher, S., and Clappier, A.: Emission ensemble approach to improve the development of multi-scale emission inventories, EGUsphere, 2023, 1-27, 10.5194/egusphere-2023-1257, 2023.

US EIA: US Coal mines, https://atlas.eia.gov/datasets/eia::coal-mines-1/explore, 2022a.

US EIA: US Energy Atlas, https://atlas.eia.gov/datasets/eia::power-plants/explore?location=41.629235%2C-118.496000%2C3.79, 2022b.

USGS: USGS Mineral Resources On-Line Spatial Data, http://mrdata.usgs.gov/, Last Access: January 2019, 2019.

Van Damme, M., Clarisse, L., Whitburn, S., Hadji-Lazaro, J., Hurtmans, D., Clerbaux, C., and Coheur, P.-F.: Industrial and agricultural ammonia point sources exposed, Nature, 564, 99-103, 10.1038/s41586-018-0747-1, 2018.

Wang, C., Corbett, J., and Firestone, J.: Improving Spatial Representation of Global Ship Emissions Inventories, Environmental science & technology, 42, 193-199, 10.1021/es0700799, 2008.

World Bank: Global Gas Flaring Tracker Report, https://www.worldbank.org/en/programs/gasflaringreduction/global-flaring-data, Last Access: August 2023, 2023.

World Resources Institute: Global Power Plant Database, Global Energy Observatory, Google, KTH Royal Institute of Technology in Stockholm, Enipedia, 2018.

WRI: Global Power Plant Database v1.3.0, https://datasets.wri.org/dataset/globalpowerplantdatabase, 2021.

Zhang, X., Geng, Y., Shao, S., Dong, H., Wu, R., Yao, T., and Song, J.: How to achieve China's CO2 emission reduction targets by provincial efforts? – An analysis based on generalized Divisia index and dynamic scenario simulation, Renewable and Sustainable Energy Reviews, 127, 109892, https://doi.org/10.1016/j.rser.2020.109892, 2020.

Table 1 – Overview of updated spatial proxies in EDGARv8.0, including data sources and methods.

Sector and spatial coverage	OLD EDGAR proxies	NEW EDGAR proxies	Details NEW EDGAR proxies	Time coverage	Data access
Power plants (global)	CARMAv3 (not anymore available): 2004, 2009, 2014, fuel type derived from plant capacity (assumption)	**Global coal/gas plant tracker (Global Energy Monitor)**	Coal, Gas	1970-2050	https://globalenergymonitor.org/projects/global-coal-plant-tracker/ and https://globalenergymonitor.org/projects/global-gas-plant-tracker/ (2022)
		Global Power Plant Database v1.3.0	Biomass, Other, Oil		https://datasets.wri.org/dataset/globalpowerplantdatabase
		US EIA	USA power plants, all fuels	All	https://atlas.eia.gov/datasets/eia::power-plants/explore?location=41.629235%2C-118.496000%2C3.79
		CARMAv3	Autoproducers, missing countries	2004, 2009, 2014	http://carma.org/
All other industries (Europe)	EPRTR v4*	**European Pollutant Release and Transfer Register (EPRTR), v18**	All industries and waste plants (with the exception of power plants, iron and steel and coal mines)	2007-2017	https://www.eea.europa.eu/data-and-maps/data/member-states-reporting-art-7-under-the-european-pollutant-release-and-transfer-register-e-prtr-regulation-23/european-pollutant-release-and-transfer-register-e-prtr-data-base/eprtr_v9_csv.zip
Iron and Steel (global)	In-house EDGAR	**Global steel plant tracker (Global Energy Monitor)**		1970-2050	https://globalenergymonitor.org/projects/global-steel-plant-tracker/

Coal mines (global)	USGS derived proxies, Global Energy Observatory (China)	**Global coal mine tracker (Global Energy Monitor)**	Brown and hard coal, surface and underground	1970-2050	https://globalenergymonitor.org/projects/global-coal-mine-tracker/
		Global Energy Monitor + EIA (Energy Information Administration)	USA all fuels, more precise open and close years	1970-2050	https://atlas.eia.gov/datasets/eia::coal-mines-1/explore
		EDGAR old proxy	For missing countries	Key years	
Flaring (global)	NOAA-NDGC (2015) VIIRS data https://www.ngdc.noaa.gov/eog/viirs.html	Global Gas Flaring Tracker Report (2023)	Used both for venting and flaring activities	2012-2022	https://www.worldbank.org/en/programs/gasflaringreduction/global-flaring-data
Small scale combustion (global)	Global Human Settlements Layer (1975, 1990, 2000, 2015)	Global Human Settlements Layer data Package 2023 + Heating Degree Days from ERA5	For all fuels	Population every 5 years from 1975 to 2030, HDD every year from 1970 to 2022	https://ghsl.jrc.ec.europa.eu/ghs_pop2023.php and https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form)
Small scale combustion in agriculture (global)-Rural population	Global Human Settlements Layer (1975, 1990, 2000, 2015)	Global Human Settlements Layer data Package 2023, including GHS-SMOD R2023A - GHS settlement layers + Heating Degree Days from ERA5	For small-scale combustion in agriculture which are mostly associated to rural areas.	Population every 5 years from 1975 to 2030, HDD every year from 1970 to 2022	https://ghsl.jrc.ec.europa.eu/ghs_pop2023.php, https://ghsl.jrc.ec.europa.eu/ghs_smod2023.php, and https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form)
Intensive livestock and fertiliser industries (global)	Livestock density maps	ESA World Emission project +intensive livestock point sources were taken from EPRTRv18 for Europe.	For intensive livestock and fertiliser industry+ gapfilling with livestock density map	2008-2022	https://www.world-emission.com/
Gap-filling of industrial activities (global)	Population based	Built-up for non-residential areas from Global Human	It is used entirely when no information is available or attributing a fraction of	every 5 years from 1975	https://ghsl.jrc.ec.europa.eu/ghs_buS2023.php

		Settlements data package 2023	emissions which was not allocated to point sources.	to 2030	
International shipping	In-house EDGAR proxy based on LRIT and Wang et al. (2007) and Trombetti et al. (2017)	STEAM (Ship Traffic Emission Assessment Model)	Based on CO2 emissions for multi vessels and multi-years.	2000-2018	Jalkanen et al., 2012; Johansson et al., 2017