CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling

Jeroen Kuenen[1], Stijn Dellaert[1], Antoon Visschedijk[1], Jukka-Pekka Jalkanen[2], Ingrid Super[1], Hugo Denier van der Gon[1]

5   [1]TNO, Department of Climate, Air and Sustainability, Princetonlaan 6, 3584CB Utrecht, The Netherlands
[2]FMI, Department of Atmospheric Composition Research, P.O. Box 503, FI-00101 Helsinki, Finland

Correspondence to: Jeroen Kuenen (Jeroen.Kuenen@tno.nl)

Abstract

10   This paper presents a state-of-the-art anthropogenic emission inventory developed for the European domain for a 18-year time series (2000-2017) at a 0.1°x0.05° grid, specifically designed to support air quality modelling. The main air pollutants are included: NOx, SO_2, NMVOC, NH_3, CO, PM_{10} and $PM_{2.5}$ and also CH_4. To stay as close as possible to the emissions as officially reported and used in policy assessment, the inventory uses where possible the officially reported emission data by European countries to the UN Framework Convention on
15   Climate Change, the Convention on Long-Range Transboundary Air Pollution and the EU National Emission Ceilings Directive as the basis. Where deemed necessary because of errors, incompleteness of inconsistencies, these are replaced with or complemented by other emission data, most notably the estimates included in the Greenhouse gas Air pollution Interaction and Synergies (GAINS) model. Emissions are collected at the high sectoral level, distinguishing around 250 different sector-fuel combinations, whereafter a consistent spatial
20   distribution is applied for Europe. A specific proxy is selected for each of the sector-fuel combinations, pollutants and years. Point source emissions are largely based on reported facility level emissions, complemented by other sources of point source data for power plants. For specific sources, the resulting emission data were replaced with other datasets. Emissions from shipping (both inland and at sea) are based on the results from a separate shipping emission model where emissions are based on actual ship movement data, and agricultural waste burning
25   emissions are based on satellite observations. The resulting spatially distributed emissions are evaluated against earlier versions of the dataset as well as to alternative emission estimates, which reveals specific discrepancies in some cases. Along with the resulting annual emission maps, profiles for splitting PM and NMVOC into individual component are provided, as well as information on the height profile by sector and temporal disaggregation down to hourly level to support modelling activities. Annual grid maps are available in csv and
30   NetCDF format (Kuenen et al., 2021).

1 Introduction

Emission inventories are the key starting point for understanding the causes and possible mitigation of air pollution. They provide information about the sources of air pollution, which can be used in environmental assessment models and air quality models to obtain information on levels of air pollution. In particular, atmospheric dispersion models are using the information on the releases of air pollutants into the atmosphere to calculate levels of air pollution in various geographical domains to get a better understanding of the relation between emission sources and air pollutant concentrations (Belis et al., 2020; Liang et al., 2018). When inventories span a longer period the trend in air pollution and exposure can also be assessed, for example to identify the impact of changes in industrial processes, fuel mix or implementation of policies (Buonocore et al., 2021).

At the same time, emission inventories are the backbone of policies that control air pollution and climate change. In the climate change community, the United Nations Framework Convention on Climate Change (UNFCCC) relies on emission inventories to provide information on the reduction of emissions and progress towards future reduction commitments related to the Kyoto Protocol and the Paris Agreement. In the European air pollution community, the UNECE Convention on Long-Range Transboundary Air Pollution (CLRTAP) (UNECE, 2012) and the EU National Emission Ceilings Directive (NECD) (European Commission, 2016) set reduction commitments for air pollutant emissions. In this framework, emission inventories are used to identify mitigation options and to check progress towards objectives. In view of these policy needs, each country that is part of the Convention is required to develop an emission inventory that meets the requirements set for these inventories annually. These requirements have been standardized over the last decades by prescribing methodologies for both greenhouse gases and air pollutants. These methodologies, documented in the IPCC Guidelines for National Greenhouse Gas Inventories (Eggleston et al., 2006) for greenhouse gases and the EMEP/EEA Air Pollutant Emission Inventory Guidebook (EEA, 2019a) for air pollutants, provide a set of default methodologies that each country shall use to establish the emission inventory for each source. However, they explicitly say that if a country has better information than the default methodologies provided by the guidance, it should be used.

Apart from the national total emissions by sector, both CLRTAP and NECD require countries to submit gridded emission inventories to support air quality modelling and assessment at (sub-)national and European level. While until 2016, the requirement under CLRTAP was to report at a horizontal resolution of 50x50km2, the resolution for both CLRTAP and NECD has been increased to 0.1°x0.1°, which is equivalent to roughly 5-10km. However, most of the EU Member States do report gridded emission data that are good quality, some of them and most of the non-EU countries submit incomplete or erroneous data or no data at all (EMEP, 2017). Therefore, the merged submitted gridded inventory data does not yet provide a complete and reliable inventory for the entire European domain that is usable for air quality modelling. Using the country reporting of gridded data, CEIP annually compiles a European spatially distributed emission inventory for modellers. Given the shortcomings described in the reported gridded emissions, also this work involves significant gapfilling on both the sectoral total emissions per country as well as the spatial distribution component (Wankmüller, 2019). Where directly reported gridded data are not available, the spatial distribution pattern from other inventories, in particular CAMS-REG (this inventory) is used to disaggregate country and sector totals.

70 For CH$_4$, being a greenhouse gas, no requirements with regard to emission reporting exist in the framework of CLRTAP or NECD. Given that UNFCCC does not include any requirements for spatially distributed emissions, no spatially distributed CH$_4$ emissions are available from country reporting.

Next to the official inventories various initiatives have been providing gridded emission inventories as this information is a prerequisite for global, national and local air quality studies and policies.   Globally, the EDGAR
75 emission inventory was developed since the 1990s as a bottom-up emission inventory using activity data (e.g. energy statistics) and emission factors (Crippa et al., 2018). Other important global inventories include ECLIPSE (Klimont et al., 2017), CEDS (McDuffie et al., 2020) and CAMS-GLOB-ANT (Doumbia et al., 2021). At the European level, the TNO_MACC inventories have been developed in the MACC (Monitoring Atmospheric Composition and Climate) FP7 project since 2007 (Kuenen et al., 2014). The MACC project has evolved into the
80 Copernicus Atmosphere Monitoring Service (CAMS) under the umbrella of the EU Copernicus programme (https://www.copernicus.eu/en). CAMS identified a continuous need for up-to-date emission information that can be used in support of air quality production and forecasting systems at both the global and the European scale. To fulfill this need at the European scale the CAMS regional inventory (CAMS-REG) was developed. The CAMS-REG inventory covers both air pollutants and greenhouse gases, the latter to support upcoming initiatives for
85 GHG emission verification and inversions. This paper describes the methodology used to derive the CAMS-REG air pollutant (CAMS-REG-AP) inventory (hereafter referred to as CAMS-REG), in particular its version 4 which covers the years 2000-2017. Separately, CAMS-REG-GHG is available which includes emsisions of CO$_2$ and CH$_4$. Apart from the use in CAMS, the data are freely available for air quality modellers and other scientists who are in need of emission information at the European scale. To make the inventory fit for purpose CAMS-REG not
90 only provides the gridded emissions but also default profiles for typical emission height by source type, temporal profiles to distribute emissions over the year, and chemical speciation of PM and NMVOC emissions.

2   Methodology

The CAMS-REG emission inventory focuses on the main air pollutants (NO$_x$, SO$_2$, NMVOC, NH3, CO, PM$_{10}$ and PM$_{2.5}$). CH$_4$ is also included because of its role in atmospheric chemistry. For this inventory, the general
95 methodology is illustrated in Figure 1. The strategy was to use official reported emission data from national inventories for both the greenhouse gases and the air pollutants where possible, recognizing that these often contain specific national information which increases the accuracy of the data compared to estimates at continental or global scale (see Sect. 2.2.2 and 2.2.3). At the same time, it is clear that in specific cases these data have shortcomings e.g. in the form of missing or inaccurate data (EMEP, 2020). Therefore, an alternative
100 emission dataset available from the IIASA GAINS model (IIASA, 2018) was used to fill in the gaps or replace any data which are considered of insufficient quality (see Sect. 2.2.1 and 2.2.4). All of this was done at the level of annual sectoral emissions by country (not distributed in space). As a next step, the dataset holding emissions from these different sources was spatially distributed in a consistent manner using relevant proxies for each source (see Sect. 2.3). In addition, all shipping emissions were excluded and taken from a different data source
105 (see Sect. 2.4), to allow a consistent approach towards all shipping emissions given the mix of national and international shipping. For agricultural waste burning a similar approach was followed (see Sect. 2.5) given the limited reporting of emissions from this source.

This methodology was applied for UNECE-Europe, which refers to all European countries including Turkey (as a whole) and Russia (only the European part, until 60°E) at the eastern side. The domain stretches between 30°W and 60°E, and between 30°N and 72°N, including all the European countries fully within the domain. Emissions from countries outside of Europe but still part of the rectangular domain (most notably North Africa and the Middle-East) were taken from EDGAR-v4.3.2 (Crippa et al., 2018) to complete the overview of anthropogenic emissions for the entire domain. Finally, only those sources contributing to the national total emissions in the inventory reporting system were included, all (semi-)natural sources were excluded for as far as reported, which is in line with the sources included in official national total emissions used for compliance assessment.

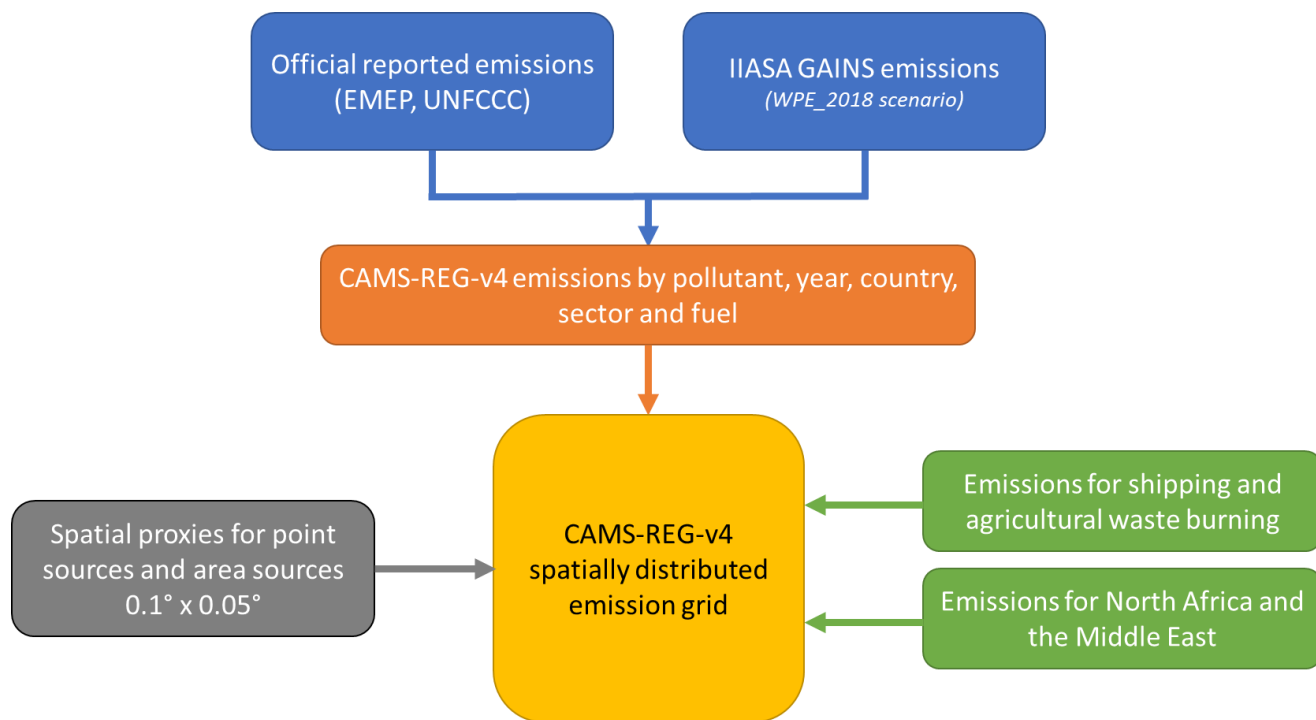



Figure 1: General methodology applied for the CAMS-REG emission inventory.

2.1 Source sector definition

The reported data from EMEP and UNFCCC (for CH_4) as well as the IIASA GAINS emissions were converted to a newly defined sector format, which combined the level of detail available in various datasets. A full overview of the sector definitions is provided in the supplementary information (SI, Table S1). This new sector and fuel definition was developed primarily since none of the existing classifications did contain all the necessary details. Also, it gave the opportunity to introduce a hierarchical and numerical structure which allows simple aggregation

and disaggregation of emission sectors where necessary. The sectoral structure defines 209 individual source categories at the highest level of detail, considering the highest detail in sectoral emissions for each of the data sources. Each sector can be aggregated up to a minimum of 7 main groups: energy industry, manufacturing industry and product use, road transport, non-road transport, small combustion activities, agriculture and waste. In practice, the emission data were processed at the highest sectoral detail possible, restricted by the level of detail in the different data sources (EMEP, UNFCCC and IIASA GAINS).

Apart from the detailed source sector definition, also an aggregated sector level was defined which is used as the default aggregation level in which the gridded emissions are provided. This aggregation is based on the GNFR sector level, which is an aggregation of the NFR (Nomenclature For Reporting) that is used as the basis for reporting spatially distributed emissions of air pollutants by European countries. A complete overview of the GNFR sectors in this inventory is given in the supplementary information (SI, Table S1).

2.2   Emission data collection and processing

2.2.1   GAINS emissions

The GAINS model is developed by IIASA to explore emission control strategies for air pollutants and greenhouse gases by modelling the impact of possible measures. Originally developed for Europe, it now covers most regions of the world. The model is used to underpin policies such as the UNECE Gothenburg Protocol (UNECE, 2012) and the EU National Emission Ceilings Directive (European Commission, 2016). For this inventory, the most recent emission data from the model was taken, as it was incorporated in the CEP_post2014_CLE scenario updated in December 2018 (IIASA, 2018). This scenario takes into account historical emission data up to 2015, and future emissions for 5-yearly intervals (2020, 2025, 2030) were modelled based on the latest available information on activity data and control measures available at the time. The emission data were obtained at the level of detailed source categories and fuels for each country for 5-yearly intervals (2005, 2010, 2015, 2020). Linear interpolation was used to estimate emissions for each of the years in between. To estimate emissions prior to 2005, an earlier GAINS dataset which was used in the TNO_MACC-II inventory (Kuenen et al., 2014) and does include the year 2000 was used to extrapolate the trend backwards in time until 2000.

The GAINS sector and fuel classifications were converted to our own sector and fuel definitions (see SI: Table S1 for definition, Table S4 for the links). For industrial combustion, an additional step was needed since the GAINS sector classification has most industrial combustion aggregated to one industrial combustion sector. To split these over various industrial sectors, a specific bottom-up emission inventory was set up for the industrial sectors. This bottom-up inventory uses energy consumption from the IEA energy statistics combined with default emission factors to calculate emissions per pollutant, country, year and industrial sector. The share of each sector in this bottom-up inventory was then used to disaggregate the GAINS industrial combustion emissions over the different industrial sectors as identified in this inventory (see SI, Table S1).

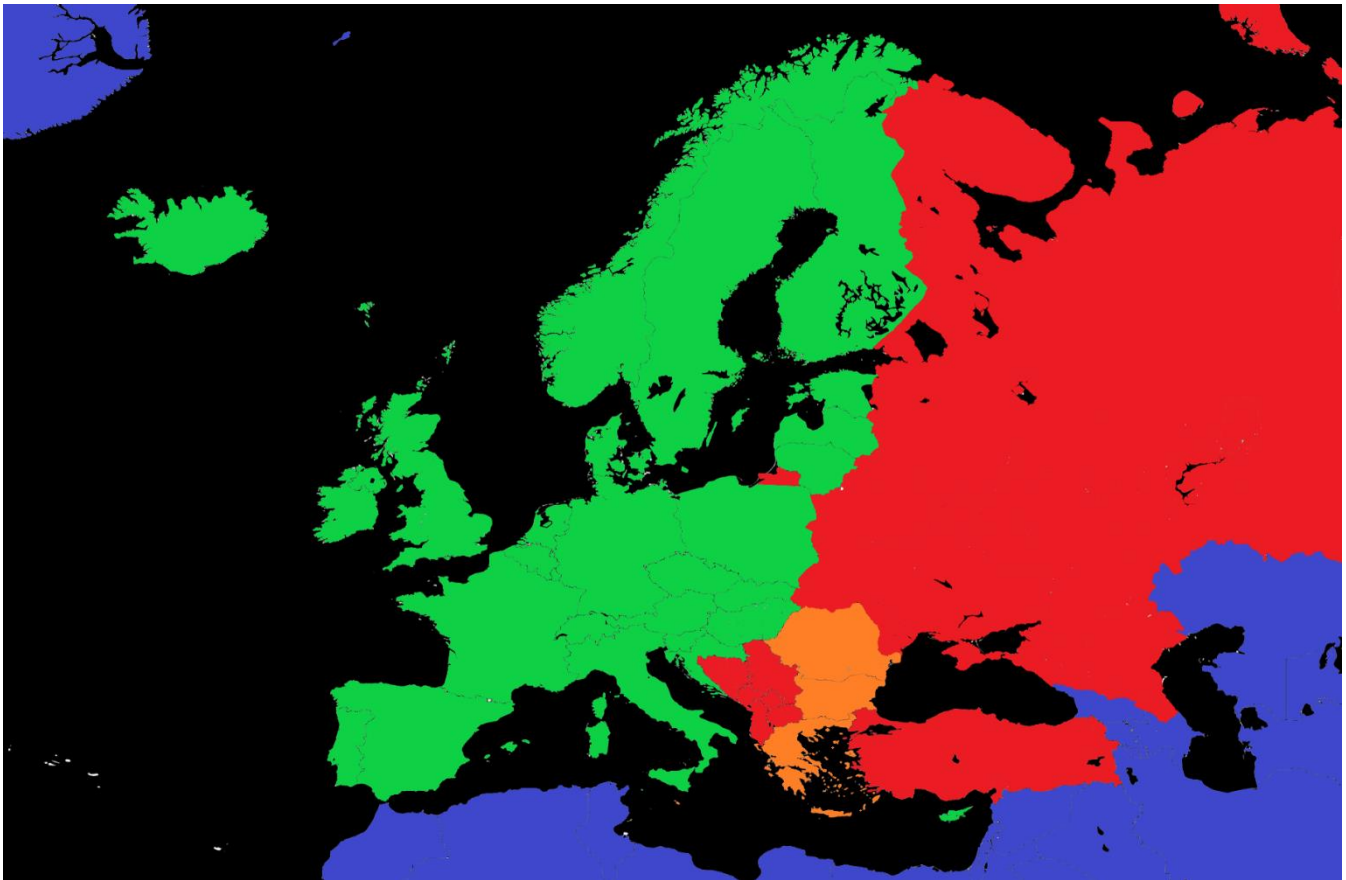
2.2.2 Reported emissions for greenhouse gases

CH_4 emissions for 2000-2017 (based on reporting year 2019) were obtained from the national inventory submissions to the UNFCCC (CRF tables) (UNFCCC, 2019). Emission data at CRF level were extracted from
165 the CRF tables and combined into a single database. For categories 1A1-1A4 which concern emissions from the combustion of fuels, emission data were collected for each individual fuel. Subsequently, the CRF sectors were converted to our own sector definitions (see SI, Table S3).

2.2.3 Reported emissions for air pollutants

Official reported emissions (reporting year 2019) were obtained from CEIP for the years 2000-2017 (CEIP,
170 2019), containing sectoral emissions reported under the EMEP reporting requirements, for all the countries for which data were available and for all air pollutants included in the scope of this inventory. Reporting follows the NFR (Nomenclature for Reporting) structure, which was converted to our own sector definitions as developed for this inventory (see SI, Table S2). Whereas the UNFCCC emissions from combustion activities and IIASA GAINS emissions are available per main fuel type, the EMEP emissions are not. Therefore, the relative
175 distribution of fuels for each pollutant, year, country and sector from the GAINS emissions dataset was used to add the fuel split to the dataset. Where for a specific combination, no emission was available from the GAINS dataset for the same pollutant, year, sector and country, an average fuel split was used which was calculated by taking the GAINS data for the sum of all years 2000-2017. Ultimately, if also this average split was not available, default fuel splits were calculated for each pollutant and sector based on the total emissions from GAINS for the
180 pollutant and sector (in all countries and all years).

2.2.4 Combination and processing

As a next step, the reported data were quality checked to decide in which cases these are fit for use. In the quality checks, it was taken into account that the greenhouse gas emissions from Annex I Parties have been reviewed on an annual basis for many years. For air pollutants, this annual review cycle is in place since 2017 as part of the
185 NEC Directive, thus covering only the EU Member States. While this annual review cycle has improved reporting by countries over time, still shortcomings are identified for EU Member States (IIASA, 2019). For non-EU countries, the completeness and quality of the inventory data differs significantly between countries, and for some countries no national emission inventory is available at all (EMEP, 2020). All in all, a thorough check on completeness and accuracy is key before using reported emission inventory data for air quality assessment.
190 Therefore, as a starting point the reported data from each EU Member State (including the United Kingdom and also including Iceland, Norway and Switzerland) were used, whereas reported emission data from other countries were not used. Hereafter these 31 countries are referred to as the EU+ countries. While for all of these other countries (consisting of Turkey, Balkan countries, and former Soviet Union countries) in some cases reported emission data are submitted, no consistent emission inventory for air pollutants is available on an annual basis.
195 Therefore GAINS emissions were used for these countries. And while for some of these countries GHG emissions are being reported to UNFCCC, for consistency reasons GAINS data were also used for the CH_4. Figure 2 shows the data sources for each European country.

200 **Figure 2 Emission domain and choice of data source for each country. Green: reported data, orange: reported data with significant corrections/gapfilling, red: GAINS emissions, blue: countries outside of UNECE-Europe, gridded emissions from EDGAR-v4.3.2 inserted for these locations**

Next, a quality check is performed for the 31 countries for which reported data were used. These checks focussed
205 on completeness (emissions reported for each GNFR sector where emissions are expected to occur), time series consistency (trend in reported emissions as expected, no missing years), and checks on the distribution of emissions over sectors to identify possible missing sectors. Compared to the assessment of reported data for the earlier TNO_MACC-II inventory (Kuenen et al., 2014), the number of gaps and inconsistencies was significantly smaller, and no major issues of such nature were identified. For 3 countries (Romania, Malta and Lithuania),
210 emissions prior to 2005 were found to be incomplete and/or inconsistent with later years. This is likely related to the fact that under the NEC Directive, 2005 is the base year and there is relatively little attention for earlier years. For these 3 countries, the 2000-2004 emissions were replaced with an extrapolation of the 2005 emissions based on the trend in GAINS emissions, per GNFR category.

At the same time, larger inconstancies were identified for agricultural emissions. For NMVOC from animal
215 husbandry and manure application, a methodology to estimate emissions was only recently included in the

EMEP/EEA Guidebook (EEA, 2019a), which led to inconsistent and incomplete reporting by countries. Given that these emissions are also not included in the GAINS model, it was decided to leave this source out of the CAMS-REG-v4 inventory. For NO_x from agriculture, also reporting is found to be inconsistent between countries. In addition, one of the main sources of agricultural NO_x emissions is soil NO_x, and many air quality models have separate modules to calculate soil NO_x internally. To avoid double counting it was therefore decided to exclude NO_x emissions from agriculture.

Minor issues that were found include mainly gaps or outliers in a specific year for a specific sectors. These were identified by looking at year-to-year changes for each country and pollutant at the level of GNFR categories (aggregated sectors) and take those cases where the average importance of the GNFR category in the national total was above 3%, and at the same time there was a change of more than 5% from one year to the next in the time series. Where this was based on reported data, the underlying detailed sector data were checked, and gaps or other errors were identified these were fixed using interpolation and/or extrapolation or by keeping emissions constant from the previous/next year (based on a manual case-by-case assessment).

Finally, some other modifications were made to the dataset to make it consistent and fit for purpose for the spatial distribution:

- Emissions of NMVOC from natural gas production and distribution systems (NFR category 1B2b) were split into production, high-pressure distribution and low-pressure distribution based on the relative contribution of these subsectors in GAINS to total emissions of CH_4.
- Emissions from combustion in energy industries (excl. power/heat plants and refineries) were split into fuel consumption in coal mines, oil extraction, gas extraction and coke ovens (see SI Chapter 1 for details).
- Emissions from road transport are available at different levels of aggregation (different vehicle type groups) which were harmonized. Also, a road type split between highway and non-highway (urban and rural) emissions was added based on information obtained from the COPERT model (Ntziachristos et al., 2009). More details are provided in the SI (Chapter 3).

2.3   Spatial distribution

Each combination of sector and fuel was assigned a specific proxy for the spatial distribution. The proxy is a variable which is available in gridded form (at the resolution of 0.1° x 0.05°) and can be used to mimic the spatial distribution of the emission source. The proxy is defined as the fraction of the national total to be allocated to this grid cell, and the sum of fractions for each country always equals one.

Some examples of proxies include the network of highways in each country with traffic intensities, which is used to distribute emission from road transport on highways, and a list of emissions from individual power plants and their emission strength for specific pollutants is used to distribute emission from power plants. The summary of proxies used for each sector is provided in Table 1, and a complete overview of the selected proxies for each source is provided in the Supplementary Information (SI, Table S5).

Table 1 Summary table with main proxies per GNFR source category

GNFR category	Main proxies used
A Power plants	- E-PRTR & LCP reporting combined with Platts-WEPP - CORINE land cover 2012 Industrial area
B Industrial sources	- E-PRTR - Own point source database - CORINE land cover 2012 Industrial area
C Other stationary combustion	- Population density - CORINE land cover 2012 Arable land (for stationary agricultural emissions) - Own wood consumption map
D Fugitives	- Own point source database - Population density
E Solvents	- CORINE land cover 2012 Industrial area - Population density
F Road Transport	- Road network - Population density
H Aviation	Airports
I Off-road	- Population density - CORINE land cover 2012 Industrial area, Arable land
J Waste	- E-PRTR - Own point source database - Waste water treatment plants - Population density
K Agriculture livestock	FAO gridded livestock
L Agriculture other	- CAPRI model distributions - CORINE land cover 2012 Arable land

255 For a number of the sectors, point source information was used in the spatial distribution. This concerns power plants, industrial sources, airports and waste water treatment plants, these are described in Sect. 2.3.1 to 2.3.3.

Non-point source distribution proxies are described in the other subsections. For power plants and industry, the E-PRTR emissions are not only used relatively (as a proxy) but their absolute value is used. The remainder is then distributed as an area source, as this is expected to represent those sources which are below the threshold for reporting in E-PRTR.

2.3.1 Public power and heat plants

For public power and heat plants, three different datasets were combined. E-PRTR (the European Pollutant Release and Transfer Register) collects facility level emission data for all EU Member States, with as main goal to inform the European citizens about pollutants transfers and releases in their area. The data are reported on an annual basis since 2007 (before that three-yearly since 2001) and publicly available (EEA, 2019b). In addition to E-PRTR, EU Member States are required to report information (fuel consumption and emissions) at the level of individual installations (stacks) of large combustion plants annually. This requirement follows from the LCP Directive (European Commission, 2001), superseded by the Industrial Emissions Directive (European Commission, 2010). Thirdly, a commercial dataset on power plants known as Platts-WEPP was used, which records characteristics of power plants worldwide, such as the technology type, fuel type, capacity, etc. (Platts, 2017).

In processing the E-PRTR and LCP reported emission data at facility level, a sector designator was used to select the facilities corresponding to the public power and heat sector. Matching and linking of identical facilities in the LCP and E-PRTR datasets was partly possible using a joint 'national ID' field. Linking was then manually completed based on similarities in facility name and location and each individual facility was assigned a new and unique identifier. Emissions of NO_x, SO_2 and dust reported to the LCP dataset, were supplemented with an estimate of CO_2 emissions based on reported fuel use and default CO_2 emission factors (Eggleston et al., 2006). While CO_2 is not explicitly part of this work, it has been included in the point source database since it is relatively well covered and has a relatively low uncertainty, and could therefore be used as an indicator for gapfilling.

Emissions reported to E-PRTR and LCP datasets were then combined by facility, pollutant and year using the new joint ID field. Since the scope of the E-PRTR is the most complete[1], where available the E-PRTR emissions were selected. Where E-PRTR emissions were missing, but LCP emissions had been reported, LCP emissions were used instead. Since the scope of LCP and E-PRTR reporting are not identical, the LCP emission values used for gapfilling were adjusted based on the average ratio between the E-PRTR and LCP emission values for the years where both were reported. A final step of emissions gapfilling was performed using the ratio between reported emission values for CO_2 and the other pollutants. When both an air pollutant and CO_2 are reported together for several years, the average ratio between the emission values was multiplied with the reported (or calculated) CO_2 value for years the pollutant had not been included. To avoid introducing outlier emission values, the E-PRTR reporting threshold were applied as a maximum emission value in gapfilling.

The facility level emissions were then split by fuel type. This was done by calculating a proxy emission value using the LCP reported fuel input by fuel type and country-, fuel- and pollutant-specific emission factors from the

---

[1] LCP reporting is only for facilities with a thermal capacity >50 MW thermal, and covers emissions of NOx, SO2 and dust

IIASA GAINS model at the country, fuel and pollutant specific level. The relative contribution of each fuel type was then applied to the reported emission value to distribute it to the various fuel types. Where no fuel input data was available, the unit fuel type from the Platts WEPP dataset was used instead to assign the emission value to one or multiple fuel types. Finally, for the remaining facilities the fuel type used was searched for online to fill the gaps.

Several checks were then performed to compare the total emissions by country with the reported UNFCCC and EMEP sector totals for public power and heat production. This check led to the identification of unrealistically high emission values for some facilities, where there had evidently been an error in reporting. In these cases, the erroneous value was removed (and then gapfilled following the routine described above) or lowered in case of a likely unit error (e.g. factor 10 too high).

The final step was the assignment of point source and area source emissions to GNFR A. In case the reported sector level emissions were higher than the processed facility level emissions, the remainder is assigned as area source emissions under GNFR A, representing the smaller facilities which are below the threshold for reporting. The processed facility level emissions were then assigned as point source emissions under GNFR A. For some countries and years, the total facility level emissions were higher than the reported sector total. In that case the emissions for all facilities were scaled down to arrive at the country sector total emission value. For the point source emissions, the facility coordinates included in the E-PRTR dataset were used for spatial distribution of the emissions. For some facilities, coordinates were added or corrected manually when they were found to be incorrect or missing. For the area source emissions, the CORINE dataset was used to spatially allocate emissions to areas with industrial activity (see Sect. 2.3.5).

2.3.2 E-PRTR for industrial sources

For industrial point sources, similar to the power plants the main sources of information for the EU+ was E-PRTR, combined with various online sources and one commercial industrial directory, and the TNO point source database for the other countries in Europe (see Sect. 2.3.3). For industrial emissions in particular the E-PRTR registry is the most complete and best available database for the EU(28)+ but it nonetheless frequently contains facilities that have an incorrect sector code or erroneous or missing emission data. Compared to power plants, industrial point sources in E-PRTR are much more numerous and diverse in type, and as a consequence a somewhat less detailed approach had to be followed to correct any pressing deficiencies in E-PRTR.

For oil refineries and integrated iron and steel plants, external lists of all existing plants in the EU+, including operational status, were consulted to extract all corresponding records from E-PRTR (years 2001, 2004 and annually from 2007 onwards) regardless of the E-PRTR sector code. Any missing facilities were added. Complete lists of existing/operational oil refineries in the EU+ were available from for instance the OGJ Worldwide Refining Survey, or online directories such as the Wiki site "A barrel full" (A barrel full, n.d.). For integrated iron and steel plants, the extensive commercial Plantfacts Capacity Database was consulted to extract any facilities operational after the year 2000 from E-PRTR and to complete any missing plants. Plantfacts data is an online source (World Steel Dynamics, n.d.) which was initially compiled and maintained by the German VDEH Steel Institute.

330 Next, based on E-PRTR sector classification, the following types industrial facilities were selected from E-PRTR (including EPER data for 2001 and 2004):

- Coal mines
- Coke ovens not belonging to iron and steel plants
- Secondary iron and steel smelters and foundries
335 - Chemical plants
- Non-ferrous metal plants (primary and secondary, incl. aluminium)
- Non-metallic mineral plants (e.g. cement, lime, glass etc.)
- Paper and pulp plants
- Waste incinerators without energy recovery
340 - Landfills without energy production
- Other waste disposal plants (such as composting plants)
- All other industrial facilities (except oil and gas production/transport/processing); grouped as "Other industry"

All industrial facilities extracted from E-PRTR have been subjected to a brief individual check for plant
345 characteristics, to ensure that the plants were assigned the right industrial activity. This elaborate process used E-PRTR plant names and locations to identify the true primary industrial activity of the E-PRTR facility through online search, as the true main activity sometimes proved different from what E-PRTR indicated.

The basic aim for industrial point sources was to compile a complete emission time-series for plants that appeared to have been operational in the period 2000-2017, but also to take plant closures, production stops and
350 emission data below the reporting thresholds into account. For oil refineries and integrated iron and steel plants any missing (but expected) CO_2, NO_x and SO_2 data were estimated, in addition to NMVOC emission data for refineries specifically and CO and PM data for iron and steel plants. For all other industrial activities the basic assumption used in gapfilling was that all operational plants must have emission data for each relevant pollutant, for each reporting year. So principally all emission data missing in this sense have been gapfilled for each
355 facility, by assuming the average of the emissions reported by that facility for other years, unless:

- emission data reported for other years was ever close to the threshold (missing emission data is below threshold);
- emission data are missing at the beginning or end of a time series (facility may report missing emission data under a different facility ID in earlier or later years, or the plant was modernized or closed);
360 - a facility did not report any emission data at all for a specific year (facility is assumed to be temporarily shut down).

2.3.3 Other point source proxies

An **own point source database** was developed for the spatial distribution of point source emissions in earlier European inventories (Denier Van Der Gon et al., 2010; Kuenen et al., 2014). Since this point source dataset is
365 only available for the year 2005, it is only used for specific sectors where no point source data could be extracted from E-PRTR.

Airports have been included as point sources in this dataset. The contribution of each airport to the country total for this sector was calculated based on flight statistics per airport (Eurostat, 2019). First, a split was made at country level between passenger and freight traffic using the number of flights in the country as a whole.
370 Thereafter, freight traffic was distributed to individual airports using the tonnage of freight, and passenger traffic was distributed using the number of passengers per airport. This was done on an annual basis, to allow for changes in time (e.g. opening of a new airport in a different location). Eurostat data was only available from 2003, for 2000-2002 the distribution is based on the Eurostat statistics as of 2003.

For **domestic waste water treatment**, an EEA dataset on urban waste water treatment plants was used (EEA,
375 2014), which includes plant coordinates. This dataset provides the flow and capacity or urban waste water treatment plants in all European countries, from which their share in total emissions was inferred.

2.3.4   Population density

The LandScan Global dataset (Oak Ridge National Laboratory, 2017) for population density was obtained for the years 2005, 2010 and 2015 at high spatial resolution (~ 1km). A country mask was combined with this dataset to
380 allocate each grid cell to a specific country. For cells that include a border between countries the relative area of each country was used to split up population for the two (or more) countries.

Based on the population of each grid cell, an additional qualification was made whether the cell was allocated as urban or rural. The definition of urban and rural areas however differs significantly between different regions of the world and between countries. To ensure consistency across the domain, a fixed value of 250 inhabitants per
385 km^2 was chosen. Above this value the cell classifies as urban, below it classifies as rural. This results in around 75% of the people living in urban areas, which corresponds well to the urban population percentage for the EU as published elsewhere (World Bank, 2018).

As a final step the data was converted to a resolution of 0.1° x 0.05° for the entire domain, separately for urban, rural and total population.

390 ### 2.3.5   Land cover

The CORINE Land Cover dataset (Copernicus Land Monitoring Service, 2016) was obtained from Copernicus Land Monitoring. This dataset has a resolution of approximately 100 metres and for each grid cell the main use type is given. These high resolution grid cells were aggregated to 0.1° x 0.05°, and for each of these larger grid cells the fraction of different main use types was calculated by adding up the amount of grid cells with this main
395 use type and normalising the totals per 0.1° x 0.05° grid cell.

The dataset identifies around 45 different use classes. For this work, 3 different proxies were extracted from this dataset:

- Industrial area (taken as the sum of "Industrial or commercial units", "Port areas" and "Construction sites")

400   -   Arable land (taken as the sum of "Non-irrigated arable land", "Permanently irrigated land", "Complex cultivation patterns" and "Land principally occupied by agriculture, with significant areas of natural vegetation")
       -    Rice fields

Similar to the population proxies, a country mask was added to the dataset to be able to calculate a distribution
405   map for each country separately.

2.3.6   Road network

For road transportation, shape files from Open Transport Map (OpenTransportMap, 2017) and Open Street Map (Open Street Map, 2017) were obtained for the entire European domain.

While Open Street Map contains the road network across Europe, OTM adds the traffic volumes to these and
410   divides the roads in different classes (main roads, and from first class to fifth class roads). The two datasets were merged, which resulted in a European-wide map of roads with each an associated intensity. This intensity is directly used as a proxy for the emissions, not taking into account more detailed parameters like vehicle speed, traffic jams, etc. However, especially for smaller roads in many cases the traffic volume was not available and traffic volume had to be estimated for these cases. To do this, first a relation between the traffic volume and the
415   population density in each grid cell was determined based on all the grid cells with a known traffic volume, per country and per road class. Then this relation was used to estimate the traffic volume where this was not available, based on the population. Since smaller roads are expected to contain mostly local traffic, this approach is expected to represent reality reasonably well.

The traffic intensity map was classified per country, per vehicle type and per road type. The latter refers to the
420   separation between highway (assumed equivalent to main roads) and non-highway emissions. Thereafter, the non-highway emissions were split between urban and rural by means of overlaying the traffic intensity map with a population map (which includes rural and urban shares, see 2.3.4). This way, the fact that on average the road in urban areas will have higher per vkm emissions is taken into account in a generic way. Using this approach, 18 different traffic intensity maps were created (6 vehicle types, 3 road types) for Europe as a whole. As a final step,
425   the resulting dataset was combined with a country mask to introduce the country codes similar to the population and land use distributions, and the map was subsequently normalized with the country total traffic intensities to create the final proxy map.

2.3.7   Other proxies

For agriculture, a number of specific agricultural proxies have been used:

430   -   **Gridded livestock** of the world (FAO, 2010), available from the UN Food and Agricultural Organisation (FAO) has been obtained per animal type and converted to a 0.1°x0.05° resolution.
       -    Distributions representative for **manure application and fertilizer application** have been extracted from the Common Agricultural Policy Regionalised Impact (CAPRI) modelling system (CAPRI, 2020). This model is a global partial equilibrium model for the agricultural sector with a focus on the European

435   Union, developed in a series of EU research projects and is now used by the European Commission to underpin agricultural policies.

Given the importance of **residential wood combustion** for some pollutants, a specific distribution proxy was developed to represent wood combustion emissions, taking into account population density and also proximity to wood. Based on population density and degree of urbanization, a hypothetical fuel wood demand function is

440   derived. The starting point is that the more densely populated and urbanized (ranging from free standing houses in rural areas to urban high rise apartment dwellings), the less wood combustion appliances will on average be present in households. Next, a fuel wood supply function is derived which, depending on land cover class, estimates how much fuel wood a certain type of landcover is able to produce sustainably. Then the wood demand function and the wood supply function are overlaid spatially, assuming that a local source of fuel wood will

445   primarily provide to the nearby residential areas. In addition, it is assumed that a residential area cannot consume more wood than what the surrounding area is able to produce sustainably. The result is a wood use proxy that principally follows population distribution but which is also strongly influenced by local wood availability and degree of urbanization. Using this proxy, most of the wood consumption is thus allocated to rural areas especially those near forested areas. Despite this modification to the distribution of residential wood combustion, an

450   overallocation of the emissions in urbanized centres may still be present in the spatial distribution (Timmermans et al., 2013). In practice it is observed that the distribution also depends on national circumstances, e.g. bans on using wood or coal in urban areas, making it difficult to derive a generic distribution at European scale.

Finally, for **high distribution gas networks** and for **rail transport**, specific maps are available to spatially distribute emissions. These are similar to those used in the earlier TNO_MACC inventory (Kuenen et al., 2014).

455

2.4   Shipping emissions

Emissions from shipping can be divided into sea shipping and inland shipping, but also into domestic and international. The latter is typically used in official reporting, where domestic refers to a ship leaving and arriving in the same country, irrespective of the route. International shipping however is not a primary part of national

460   inventory reporting, and therefore reporting is more incomplete and inconsistent. Given the importance of shipping for emissions and air quality at European scale (Jonson et al., 2020; Viana et al., 2014), shipping emissions from national reporting are replaced with an alternative based on a consistent modelling approach. The STEAM model (Jalkanen et al., 2012; Johansson et al., 2017) provides global emissions at high resolution based on AIS (Automatic Identification System) records that track the whereabouts of ships around the world. The

465   STEAM model then computes emissions from shipping for CO2 as well as major air pollutants based on the generation of shipping routes from the AIS signals, and emission characteristics based on the characteristics of each ship. For the CAMS-REG inventory, the STEAM model was run at a resolution of $0.1°\times0.05°$ for the European domain covered by this inventory, and for the relevant pollutants (NO_x, SO_2, CO, NMVOC, PM, CO2), where PM was speciated into EC, OC, SO_4 and ash. The main limitation for this inventory was that STEAM data

470   were only available from 2014 onwards, given the availability of global AIS datasets used in the ship emission modelling. Therefore, for years 2000-2013 the emission data have been extrapolated backward in time using a separate estimate of shipping emissions per year, pollutant and sea area using historic activity data and

information on fuel quality regulations and policies. This extrapolation only concerned the total emissions per sea region, whereas the spatial distribution before 2014 is assumed constant. Therefore, the shipping emissions for the years 2014-2017 are less uncertain than the pre-2014 data when AIS data were of much lower quality and/or not available at all.

2.5   Agricultural waste burning

Field burning of agricultural waste is a separate reporting category in national inventories. Formally agricultural waste burning (AWB) is forbidden in the EU. It may however still happen albeit illegally, by accident or possibly under certain exemptions. As a result the reporting of this category is highly variable between countries and inconsistent. In previous TNO-MACC (Kuenen et al., 2014) and CAMS-REG inventories this category was gap-filled using data from the IIASA-GAINS model (IIASA, 2018). With the on-going development of fire detection with global satellite products it was now possible to include an AWB emission estimate based on earth observation. For this we used the CAMS Global Fire Assimilation System (GFAS) which assimilates fire radiative power (FRP) observations from satellite-based sensors to produce daily estimates of wildfire and biomass burning emissions (Kaiser et al., 2012). The GFAS data output includes spatially gridded Fire Radiative Power (FRP), dry matter burnt and biomass burning emissions for a large set of chemical, greenhouse gas and aerosol species. Data are available globally on a regular latitude-longitude grid with horizontal resolution of 0.1 degrees from 2003 to present. By overlapping these data with land use maps the emission from biomass burning on agricultural fields was derived. To match with the annual total emission data and complete timeseries in CAMS-REG the available high resolution GFAS data over the years 2003-2018 were processed to derive at an annual average with a monthly distribution pattern and average spatial distribution map which can be applied to all years in the time series. It is important to note that in the case of Europe this approach only makes sense when data of a resolution of 0.1°x0.1° are available because in many European countries land use is mixed and forests and agricultural lands may occur in the same pixel if the resolution is not high enough. In fact even at the resolution of 0.1°x0.1° this still introduces uncertainty. Nevertheless, it is by far the best resource available for this emission source. For further details we refer to the Chapter 2 in the SI.

2.6   Non-European countries

Apart from the European emissions, also emissions from just outside the European borders may influence air quality in the study area. Therefore, the land-based emissions for those countries that are not included in Europe, but are part of the rectangular area of the domain have been added to the dataset. These "missing countries" include parts of North Africa, the Middle East and the Eastern European, Caucasus and Central Asia (EECCA) countries. Emissions for these regions were taken directly from the EDGAR v4.3.2 inventory (Crippa et al., 2018) for air pollutants (covering 1970-2012). From 2013 onwards emissions for these regions have been kept equal to 2012 levels. EDGAR source categories were converted to GNFR categories, and since the resolution in the EDGAR inventory (0.1° x 0.1°) is a factor 2 coarser compared to the resolution in this inventory, each

EDGAR grid cell was divided in half, where both halves were each assigned 50% of the emission from the
510   original EDGAR grid cell.

2.7   Speciation profiles, temporal profiles and emission height

For the pollutants which are essentially groups of pollutants (PM, NMVOC), a speciation into actual components has been calculated and provided along with this inventory. At the most detailed sector level, the PM and
515   NMVOC were split in various components for each source. PM emissions are divided into EC (elemental carbon), OC (expressed in full molecular mass), sulphate (SO_4), sodium and other minerals. NMVOC emissions have been split in 23 different hydrocarbon groups. This was done by combining information many literature sources over the last decades. For PM most of the information is derived from an earlier EC and OC inventory for Europe (Visschedijk et al., 2009), supplemented by the EMEP/EEA Guidebook (EEA, 2019a) and other source
520   specific literature. The sources for NMVOC speciation consist of many older source-specific reports from which a complete and consistent database was derived (Olivier et al., 1996). This consistent database has been the basis for many later work on NMVOC speciation (Theloke and Friedrich, 2007), and it has also been the basis for the latest speciated NMVOC inventory which was developed for the EDGAR global inventory (Huang et al., 2017). Eventually, both PM and NMVOC profiles are provided per country and per year, reflecting the different shares
525   of subsectors and fuels in each situation. In the case of PM, the profiles distinguish between fine (<2.5μm) and coarse mode (2.5-10μm) particles.

Temporal emission profiles can be used to break down annual emissions into hourly values by means of applying factors representing the month, the day in the week and the hour in the day. These are default profiles per GNFR sector, to be applied for all pollutants, countries and years, largely based on the earlier temporal profiles provided
530   with the TNO_MACC-II inventory (Kuenen et al., 2014). Recently a more detailed set of temporal profiles (CAMS-TEMPO) has been proposed (Guevara et al., 2021), for which the evaluation is currently ongoing. Based on the results of this evaluation, the default temporal profiles may be updated in the future.

Finally, for the emission height a default height profile per sector is included, which accounts for the average effective emission height (including plume rise), based on earlier work (Bieser et al., 2011). Especially in sectors
535   which include stacks, emissions are released into the atmosphere at higher altitude, which has important consequences for air quality especially close to these sources.

The PM and NMVOC speciation files, the default temporal emission profiles as well as the default emission height profiles are all available as separate files which are provided along with the gridded data files.

3   Results

3.1   Resulting emissions

Table 2 shows the total emissions for each pollutant for selected years, as well as the emission trend. It shows that for each air pollutant the emissions have decreased during the whole period. However the level of reduction for the period 2000-2017 as a whole differs significantly between pollutants, with the largest reductions for SO_2 emissions and lowest reductions for NH_3. In Table 2 the trend is calculated separately for the period 2000-2010 and for 2010-2017, which shows that for most pollutants the reductions in the 2000s have been significantly larger than in the 2010s. For NO_x and PM however, reductions are more stable, whereas for NH_3 a small increase in emissions is found between 2010 and 2017.

Table 2 Emissions for selected years for all pollutants (sum of all countries and sectors, given in kton) and the trend in emissions between 2000 and 2010 and between 2010 and 2017

	2000	2005	2010	2015	2017	Trend 2000-2010	Trend 2010-2017
CH_4	47 425	45 147	41 431	39 665	39 448	-13%	-5%
CO	58 235	49 792	43 261	35 885	35 299	-26%	-18%
NH_3	5 877	5 490	5 261	5 339	5 410	-10%	3%
NMVOC	15 805	13 532	11 377	9 867	9 757	-28%	-14%
NO_x	17 190	16 212	12 987	11 090	10 397	-24%	-20%
PM_{10}	5 465	5 281	4 854	4 483	4 436	-11%	-9%
$PM_{2.5}$	3 705	3 560	3 347	3 044	3 025	-10%	-10%
SO_2	16 110	12 665	8 748	7 482	6 230	-46%	-29%

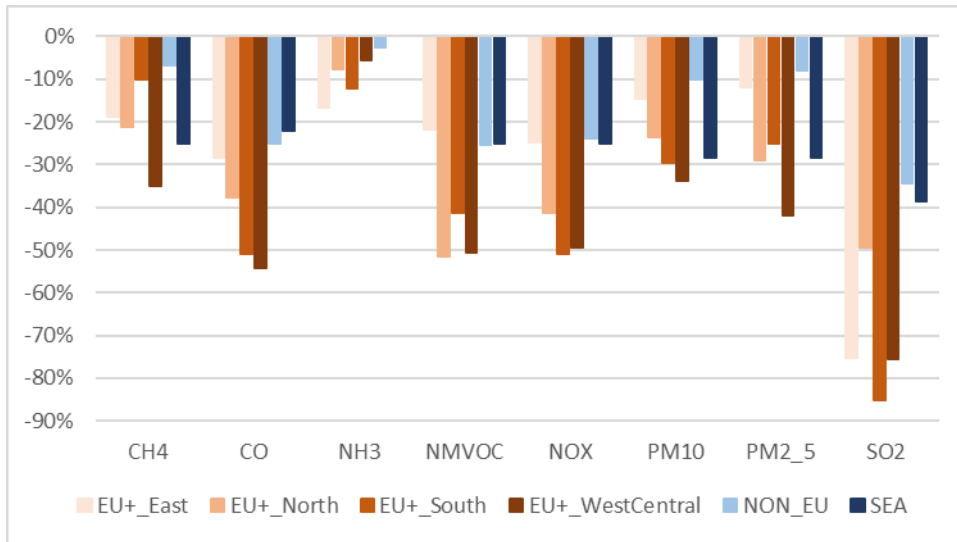
The emission reductions are not equally distributed over Europe. Figure 3 shows the change in emissions from 2000 to 2017 per country group. The country groups distinguish 4 different regions covering the EU+ countries, and separately non-EU countries and sea (international shipping). The exact definition of the country groups is provided in the SI (Table S6). It shows a mixed picture between different pollutants. Overall, largest emission reductions were achieved in the western, central and southern EU countries for most pollutants. Smaller reductions of emissions are seen for the non-EU countries and for shipping.

560 **Figure 3 Emission change between 2000 and 2017 per pollutant and per country group**

Figure 4 shows spatially distributed emissions for 2 selected pollutants (NMVOC and NO$_x$), for the sum of all sectors. In this plot, the resolution is aggregated here by a factor 2 to increase visibility of point sources in the map. Apart from these point sources (clearly visible as red dots in the maps, especially in areas where area sources are less important), other major sources (shipping, road transport) as well as urban areas are shown with

565 higher emissions. For NMVOC, emissions are more diffuse, with small combustion and solvent use as important contributors. But here also point sources are a significant contributor as shown in the map.

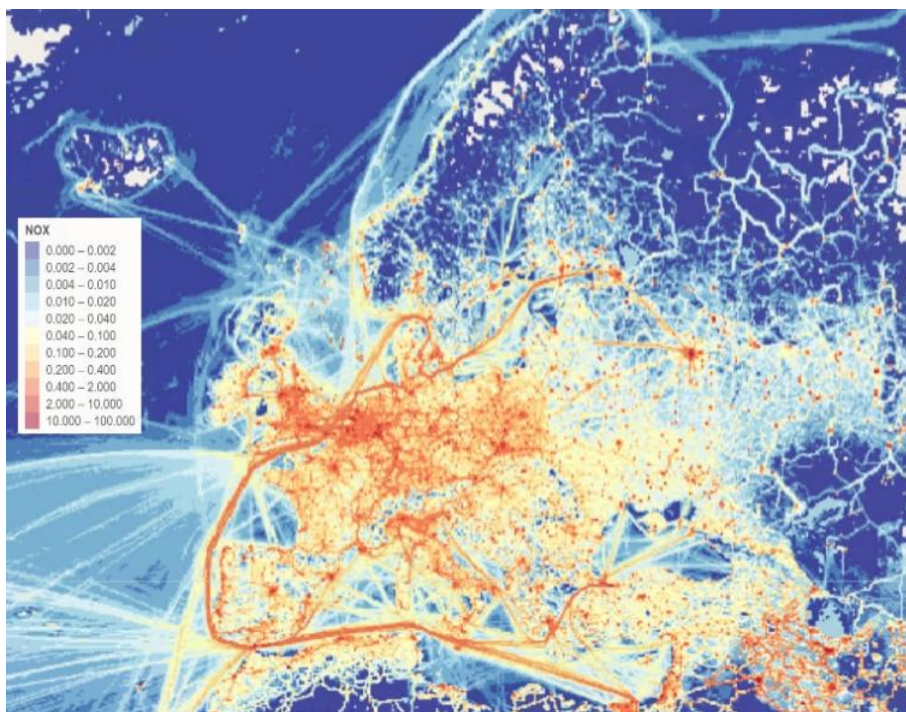
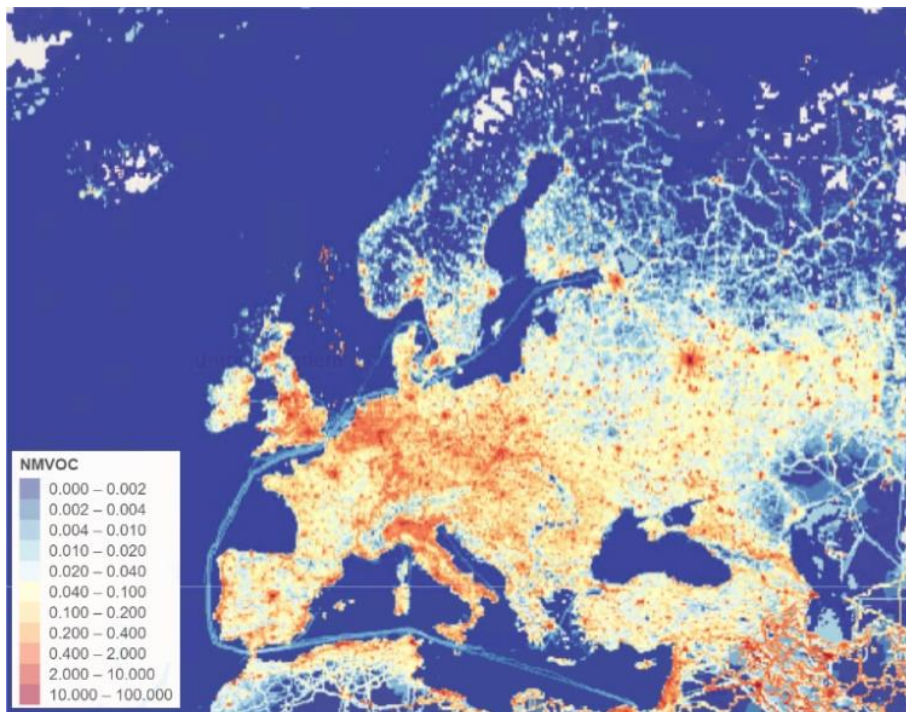



Figure 4 Spatially distributed emissions of NMVOC (upper panel) and NO$_x$ (lower panel) for 2017, in both cases the total for all sectors is shown

The difference between spatially distributed emissions in 2000 and 2017 is shown in Figure 5 for PM$_{2.5}$ emissions from small combustion (GNFR C) and NO$_x$ emissions from road transport – diesel exhaust (GNFR F2). Both examples show reductions in some countries and increases in other countries, but also within countries. The latter is due to differences in the relative contribution of underlying specific emission sources which are spatially distributed using different parameters. Figure 5 also indirectly illustrates that the EU has more coordinated policies on road transport engine technology and associated emissions than for residential combustion which shows a much more variable development. The reason why road transport NO$_x$ emission in eastern Europe do not follow the trend of the other EU regions is that the growth of road transport activity is larger than the emission reduction per vehicle.

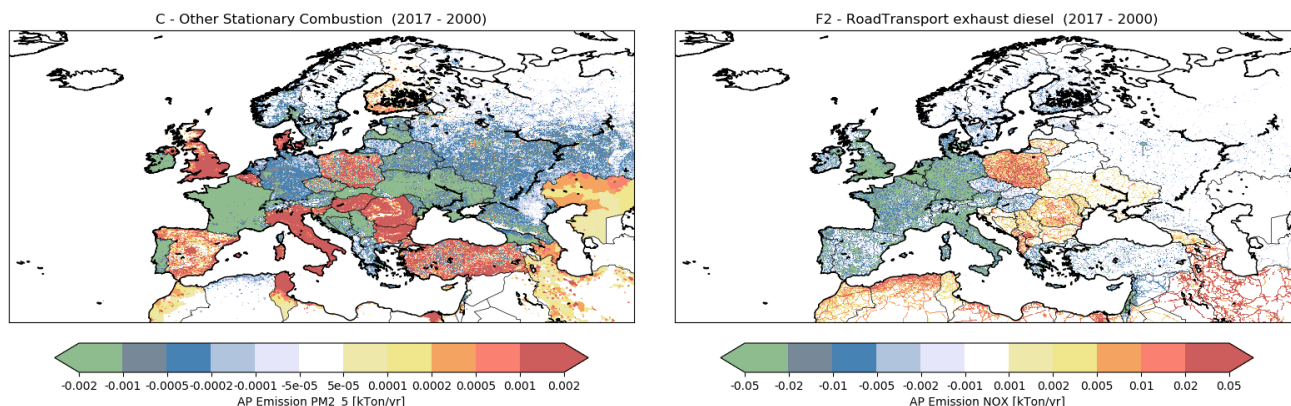



Figure 5 Difference between emissions in 2017 and 2000 for PM$_{2.5}$ from small combustion (left panel) and NO$_x$ from road transport, diesel exhaust (right panel) (figures are excluding international shipping)

3.1.1 Shipping

Shipping emissions include both shipping at sea and on rivers. For the main air pollutants, the contribution of inland shipping at the European scale is limited, typically 3-4% for PM and NO$_x$ and <0.5% for SO$_2$. Nevertheless, in major rivers and near large harbours, inland shipping emissions may be important for local air quality. As an example of the output from the STEAM model, Figure 6 shows the distribution of shipping emissions in the North Sea and the Benelux. This illustrates major shipping routes at sea, high emissions in and close to the major ports (Rotterdam, Amsterdam, Antwerp). Further transport along the main rivers is shown in the inland shipping map (center) which shows higher emissions along the inland waterways in the Netherlands, Belgium and Germany.

595 **Figure 6 Example emission distribution of CO_2 for sea shipping (left), inland shipping (middle) and sum of both (right) for a specific region covering parts of the North Sea, Baltic Sea as well as inland waterways (example for 2016, at high resolution of 1/60° x 1/120°) (figures created using Google Maps)**

As explained in Sect. 2.4, gridded emissions from the STEAM model were only available for 2014 onwards. For earlier years, scaling factors were developed for the shipping emissions to estimate emissions in the year 2000-
600 2013 by sea, taking into account environmental control measures such as the Sulphur Emission Control Areas (SECA). This is illustrated in Figure 7 where implementation of SECA on the North Sea in 2007 and consecutive sulphur reductions in 2010 and 2015 are clearly visible in the trend of SO_2 emissions.

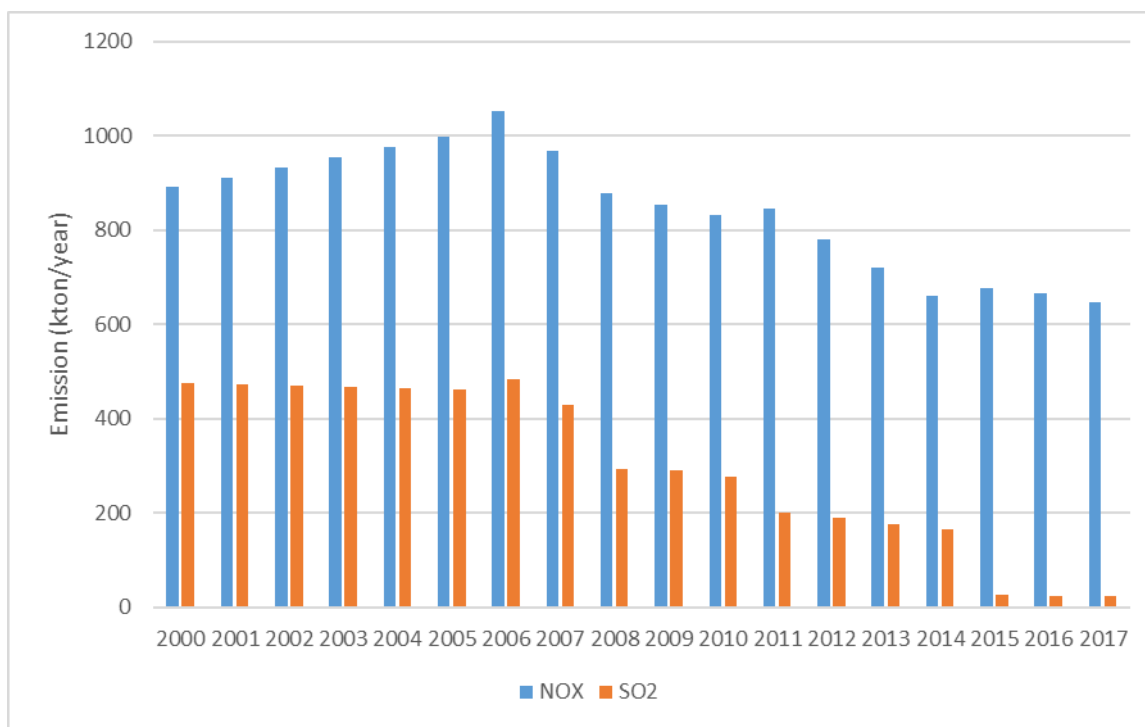



Figure 7 Emissions of NOₓ and SO₂ for the North Sea (NOS) including the English Channel (ENC) over 2000-2017.

3.2   Comparison to other inventories

3.2.1   Comparison to earlier versions

CAMS-REG-v4.2 does not only add new years to the inventory compared to earlier CAMS-REG versions, it also
provides updated emissions for the entire time series back to 2000. To illustrate the difference with earlier
versions, Table 3 shows the relative change between total emissions for the sum of all EU+ countries for the same
year with the two most widely used predecessors. TNO_MACC-III is an extension of the TNO_MACC-II
database (Kuenen et al., 2014), covering the years 2000-2011. CAMS-REG-v2 is an earlier version of the current
CAMS-REG-v4 dataset, covering the years 2000-2015. Differences between the same years in the different
inventories are partly related to methodological changes in the CAMS-REG inventory, but most can be explained
by recalculations of official reported emissions of air pollutants by each country. For TNO_MACC-III, emission
data from reporting year 2013 are used as the basis, for CAMS-REG-v2 this concerns emission data as reported
in 2017, while CAMS-REG-v4.2. builds on emission data as reported in 2019 The Table shows that PM and
NH₃ emissions are typically higher in CAMS-REG-v4.2 compared to TNO_MACC-III, while NOₓ, SO₂ and CO
emissions are typically lower.

One of the significant changes between this dataset and its predecessors is the approach to AWB as discussed in
Sect. 2.5. Table 3 shows that CAMS-REG-v4.2 for incomplete combustion related species like CO and PM$_{2.5}$ is 6

% lower that its immediate predecessor CAMS-REG-v2.2.1. While this is the net sum of various sources being adjusted downward and upward, an important contribution comes from the revised AWB estimate based on earth observation data. Over the entire European domain AWB now contributes 3.1% and 3.3% to total CO and $PM_{2.5}$ emission respectively. This used to be 8.2% and 11.2%, respectively in CAMS-REG-v2.2.1.

Table 3 Relative change between CAMS-REG-v4.2 and its predecessors (TNO_MACC-III and CAMS-REG-v2.2.1) for each pollutant for selected years, for the EU+ as a whole

	Difference against TNO_MACC-III			Difference against CAMS-REG-v2.2.1		
	2000	2005	2011	2005	2010	2015
CH_4	6%	6%	2%	-1%	-1%	0%
CO	-5%	-2%	-9%	1%	-3%	-6%
NH_3	9%	9%	6%	0%	-1%	-2%
NMVOC	-2%	-1%	-5%	2%	1%	-1%
NO_X	-4%	-4%	-7%	0%	0%	0%
PM_{10}	7%	14%	12%	2%	1%	-1%
$PM_{2.5}$	2%	6%	7%	-3%	-4%	-6%
SO_2	-5%	-6%	-14%	1%	-4%	1%

Figure 8 shows the trends in emissions for the EU+ countries from the most widely used earlier versions: TNO_MACC-III and CAMS-REG-v2. It illustrates the difference is not static in time and may change from year to year. Most of the differences can be explained by changes in reporting, which are in turn related to improved understanding of emissions and improved guidance for emission estimation.

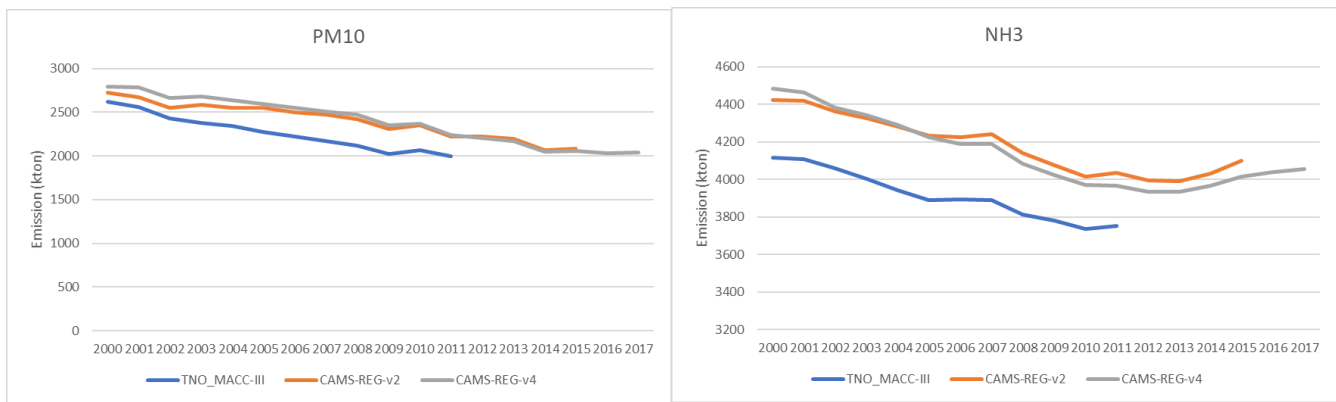



Figure 8 Comparison between CAMS-REG-v4.2 and its predecessors (TNO_MACC-III and CAMS-REG-v2.2.1) , for EU+ as a whole for PM$_{10}$ (left panel) and NH$_3$ (right panel)

Figure 10 shows a country specific comparison for PM$_{2.5}$ emissions from GNFR C (small combustion) in
640 TNO_MACC-III, CAMS-REG-v2 and CAMS-REG-v4, for the EU+ countries where reported data are used as the basis, therefore representative for the reporting of PM2.5 emissions from this source category in different years (2013, 2017 and 2019, respectively). The figure shows significant changes for some countries (e.g. EST, LTU, GBR, ESP, ITA, ROU) but only very small changes for others (e.g. DNK, NOR, DEU, FRA, SVK). These differences are significant and related to the inclusion of condensables in the emission inventories (which his
645 further discussed in Sect. 4).

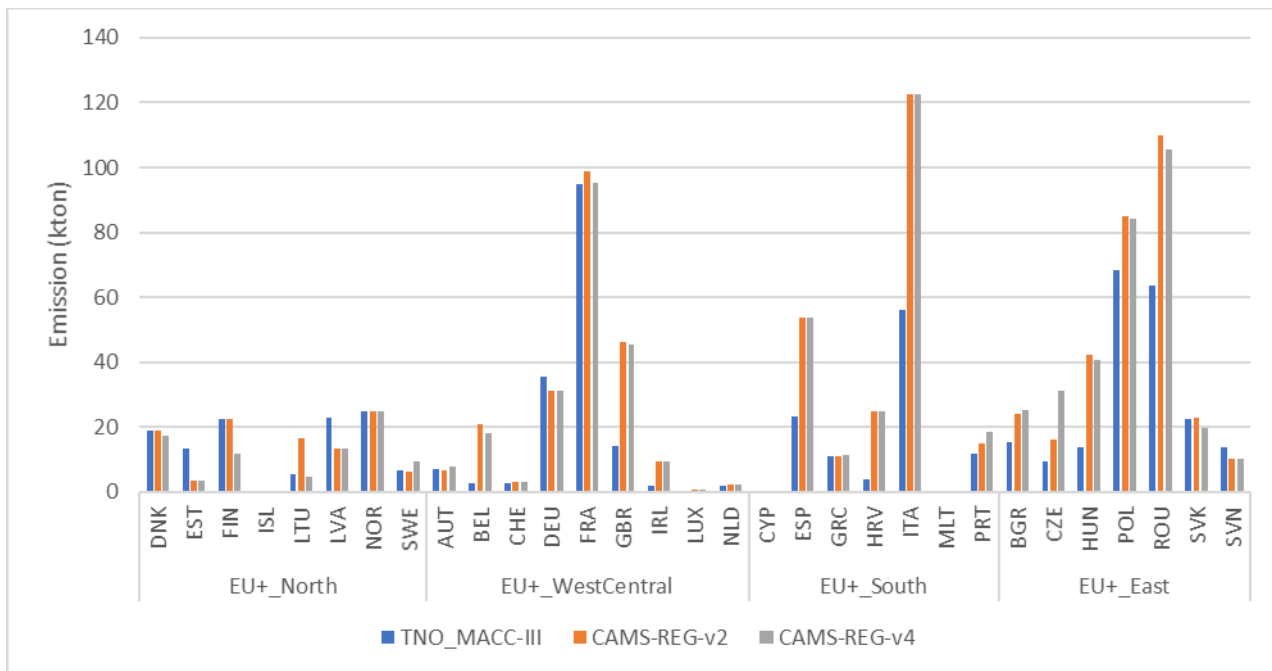



Figure 9 PM$_{2.5}$ emissions from small combustion (GNFR C) in TNO_MACC-III, CAMS-REG-v2 and CAMS-REG-v4 for the year 2010 and EU+ countries

3.2.2   Comparison to EDGAR

EDGAR (Emissions Database for Global Atmospheric Research) (Crippa et al., 2018, 2020) is a widely used global emission inventory, which uses a bottom-up approach for all sectors based on activity data (energy statistics, industrial production, etc.) combined with emission factors, developed independently from the national inventories from individual countries. Table 4 shows a comparison between the results from this inventory and EDGAR v5.0 (Crippa et al., 2020), for the year 2015. The comparison is made separately for the EU+ countries (where CAMS-REG is largely based on reported data from the countries) and non-EU countries (where CAMS-REG is largely based on the GAINS emissions). The Russian Federation is excluded from this analysis since EDGAR covers the total of the country, while CAMS-REG only includes the European part of Russia.

Total emissions from EDGAR and CAMS-REG at European scale differ considerably between pollutants, as illustrated in Table 4, especially for non-EU countries. However, also for the EU+ countries significant differences are seen for CH_4, NH_3, NMVOC and SO_2 in particular. In most cases EDGAR emissions are higher, except for PM in non-EU countries where CAMS-REG-v4.2 provides higher emissions.

	EU+ countries			Non-EU countries		
	CAMS-REG-v4.2	EDGAR v5.0	Difference	CAMS-REG-v4.2	EDGAR v5.0	Difference
CH_4	18 763	25 495	36%	5 967	9 976	67%
CO	20 250	22 025	9%	7 493	8 494	13%
NH_3	4 016	5 717	42%	892	1 798	102%
NMVOC	6 188	8 372	35%	1 450	2 668	84%
NO_x	7 218	7 676	6%	1 822	2 365	30%
PM_{10}	2 056	2 158	5%	1 332	957	-28%
$PM_{2.5}$	1 318	1 375	4%	934	642	-31%
SO_2	2 792	4 649	66%	3 311	3 789	14%

Table 4 Comparison between total annual emissions for 2015 from CAMS-REG-v4.2 and EDGAR-v5.0 for EU+ countries (left) and non-EU countries (right), excluding the Russian Federation (emissions in kton)

Figure 10 shows the difference between CAMS-REG-v4 and EDGAR-v5.0 on a GNFR sector level, for NO_x and NMVOC. To make this comparison, EDGAR emissions (which use the IPCC classification) were converted to GNFR sector classification. The match is not always perfect, which is illustrated by the figure for NMVOC, where the EDGAR-v5.0 emissions for GNFR category B (Industry) also include emissions from solvents (GNFR E), which explains the large discrepancy there. Also emissions from GNFR I (Off-road) are partly included in

670   other sectors, in particular in GNFR C. Key differences are seen for GNFR D (fugitives) and GNFR J (waste) where EDGAR includes significantly higher emissions, but also for agriculture livestock (GNFR K) there is a discrepancy since NMVOC emissions from this source are not included in CAMS-REG (see Sect. 2.2.4).

For NO_x, the comparison shows a large difference for shipping (GNFR G) where EDGAR emissions are significantly higher. This may be related to the allocation of shipping emissions between countries and sea
675   regions, as this comparison excludes international shipping. On the other hand, EDGAR emissions are significantly lower for GNFR I (off-road and other transportation) which is partly related to the sector allocation issue. Another difference is seen in GNFR L (other agriculture), which could be related to the inclusion of soil-NO_x in EDGAR which is not included in CAMS-REG-v4 (see Sect. 2.2.4).

680

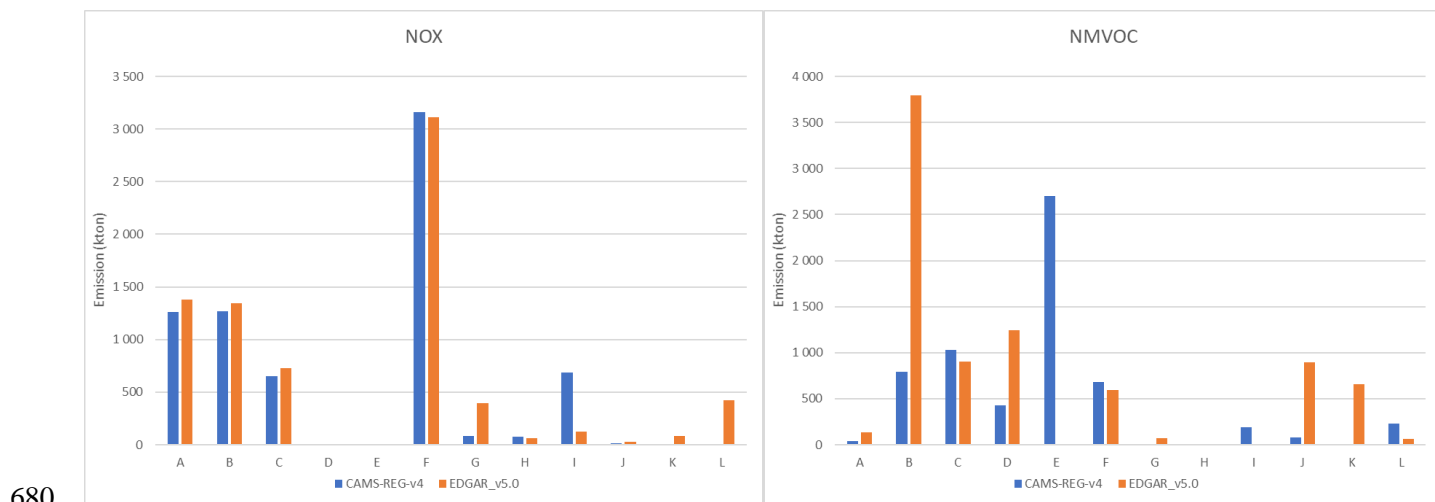



Figure 10 Comparison between CAMS-REG-v4.2 and EDGAR-v5.0 for NO_x and NMVOC for the year 2015 for EU+ countries (see Table 5 for GNFR sector explanations)

Figure 11 shows the emission trends between 2000 and 2015 both in CAMS-REG-v4.2 and EDGAR-v5.0 for 5 selected pollutants. It is shown that generally the trends are comparable in both datasets, but the downward trend
685   in CAMS-REG is stronger for each of these pollutants compared to EDGAR. This difference in trend is most notable for NMVOC, SO_2 and PM.

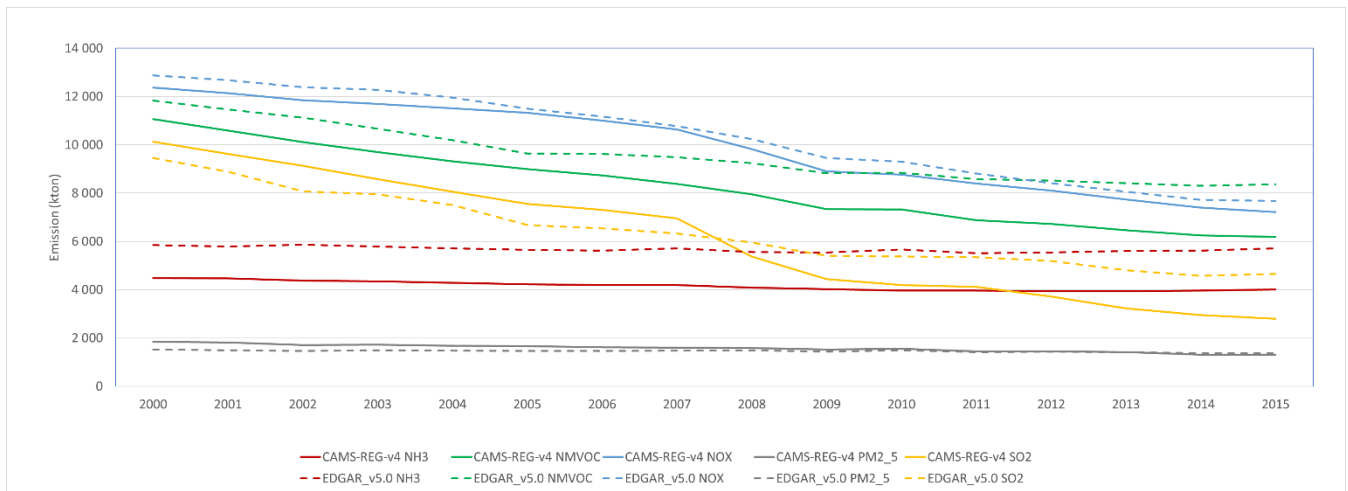



Figure 11 Trends in CAMS-REG-v4.2 (solid lines) and EDGAR-v5.0 (dashed lines) for 5 key pollutants between 2000-2015 (Russian Federation is excluded from the comparison)

690   For a comparison of previous versions of the CAMS-REG and TNO-MACC data to the inventories made by the US-EPA we refer to (Pouliot et al., 2015).

4   Discussion

The inventory described in this paper is an updated and improved inventory of the earlier described TNO_MACC inventories (Kuenen et al., 2014). Compared to this older inventory, the main changes are the further increased
695   resolution (0.1°x0.05°, was 0.125°x0.0625°) and the sectoral classification from SNAP to GNFR. The main reason for doing so was to allow easier (inter)comparison with national gridded data reported to EMEP and international datasets like EDGAR, both at 0.1° x 0.1°. The GNFR sectoral classification is used for official inventory reporting, and harmonization is beneficial for comparison and data exchange. At the same time, the inventory has been fully updated, especially with regard to the spatial distribution. This improves the
700   representation of emissions, especially when looking at larger time scales. The representation of point sources is significantly improved by incorporation of the E-PRTR (and associated datasets) which provide a robust representation of point source emissions over time, taking into account changes such as opening or closure of specific facilities as this may have significant impact on emissions in specific areas. For the distributing of area sources, a new population map was introduced, which better represents the actual population density over
705   Europe. By using this population map for 3 different years (2005, 2010, 2015) also the changes of the population distribution over time can be represented in the emission distribution, since urbanisation plays an important role in some countries over the almost 20 year period. Still however, the use of population density for distributing emissions is a simplification, which may be especially relevant for the residential sector, as different countries and urban areas may have cultural differences with regard to heating practices as well as different regulations in
710   this respect. Also the use of traffic intensity derived from a combination of Open Street Map and Open Transport Map as a proxy for road transport is a simplification, since this does not take into account that the dependency of emissions on many other parameters including vehicle speed, traffic flow and traffic jams. However, since the

goal of this inventory is to support air quality assessments at European scale, the approach is considered fit-for-purpose. This means however that when zooming in e.g. by looking at individual cities or urban areas, the
715 limitations of this inventory should be kept in mind.

The CAMS-REG-v4.2 emission inventory was constructed by combining different available datasets, similar to its previous versions. Wherever possible the choice of which dataset to use and in which situation, is based on objective criteria. For instance the use of reported emission data for EU+ countries but not using reported data for other countries. This assessment is based largely on the experience of working with these datasets in earlier years,
720 where for the non-EU+ countries the country reported data were not considered to be fit for purpose (Denier Van Der Gon et al., 2010; Kuenen et al., 2014). It was also found that even for the countries where reported emissions are expected to be good quality, errors and inconsistencies may exist. Key examples include NMVOC from agricultural husbandry and NO_x from all agriculture, which were excluded both, for the latter also to avoid double counting since there are air quality models that calculate NO_x from agricultural soils themselves. A thorough
725 assessment of the reported data for each EU+ country was performed. Since mistakes or inconsistencies in country reporting of emissions may concern various aspects (e.g. missing years, missing sectors, unrealistic distribution of emissions over sectors, etc.), it is difficult to apply a set of fixed rules to filter out errors, as these may miss specific errors as well as trigger false positives. For some further examples on inconsistency in reporting we refer to Kuenen et al. (2014). This makes that the use of expert judgement to make choices on what
730 (not) to use is a necessity.

The combination of different datasets and frequent use of expert judgement make the assessment of uncertainties more difficult. In the national emission inventories provided by individual countries, uncertainty assessment is one of the elements to be taken into consideration. The main goal of assessing uncertainties in national emission inventories is to help prioritize inventory improvements at national level by improving first those sectors with
735 relatively high uncertainty, thus efficiently reducing the uncertainty of the inventory as a whole. The EMEP/EEA Guidebook also includes a description of the methodology to be followed for such an uncertainty assessment (EEA, 2019a), which also provides indicate uncertainty ranges for activity data and emission factors depending on the source of the data. The EMEP/EEA Guidebook also provides default activity ranges, based on the assumption that all sources are calculated using activity data and emission factors. Direct measurements of
740 emissions in individual large installations would reduce uncertainty, therefore these values could be seen as an upper limit. Table 5 provides these default ranges. It should be noted that these do not take into account the uncertainty in spatial distribution of emissions.

	SO$_2$	NO$_X$	NMVOC	CO	NH$_3$	PM
A_PublicPower	10-30%	20-60%	50-200%	20-60%	Order of magnitude	50-200%
B_Industry	10-60%	20-200%	20-200%	20-200%	Order of magnitude	50-200%

C_OtherStationaryComb	10-30%	50-200%	50-200%	50-200%	Order of magnitude	100-300%
D_Fugitives	50-200%	50-200%	50-200%	50-200%		100-300%
E_Solvents			20-60%			
F_RoadTransport	10-30%	50-200%	50-200%	50-200%	Order of magnitude	50-200%
G_Shipping, H_Aviation, I_Offroad	20-60%	100-300%	100-300%	100-300%	Order of magnitude	100-300%
J_Waste	20-60%	20-60%	20-60%	50-200%	Order of magnitude	50-200%
K_AgriLivestock, L_AgriOther		100-300%	100-300%	100-300%	100-300%	Order of magnitude

Table 5 GNFR categories and their estimated uncertainty based on the default approach to emission inventories using activity data and emission factors, and the typical source of the data. Adapted from (EEA, 2019a).

However, despite requirements to do so, not all countries perform such an uncertainty assessment, which is also concluded in a recent report (Schindlbacher et al., 2021). The uncertainty values reported by different countries are shown in Figure 12. This illustrates a wide range of reported uncertainties in total emissions, ranging between <10% and >50% for NO_x and between 5% and nearly 40% for SO_2. The countries not shown here did not report quantitative information on uncertainties.

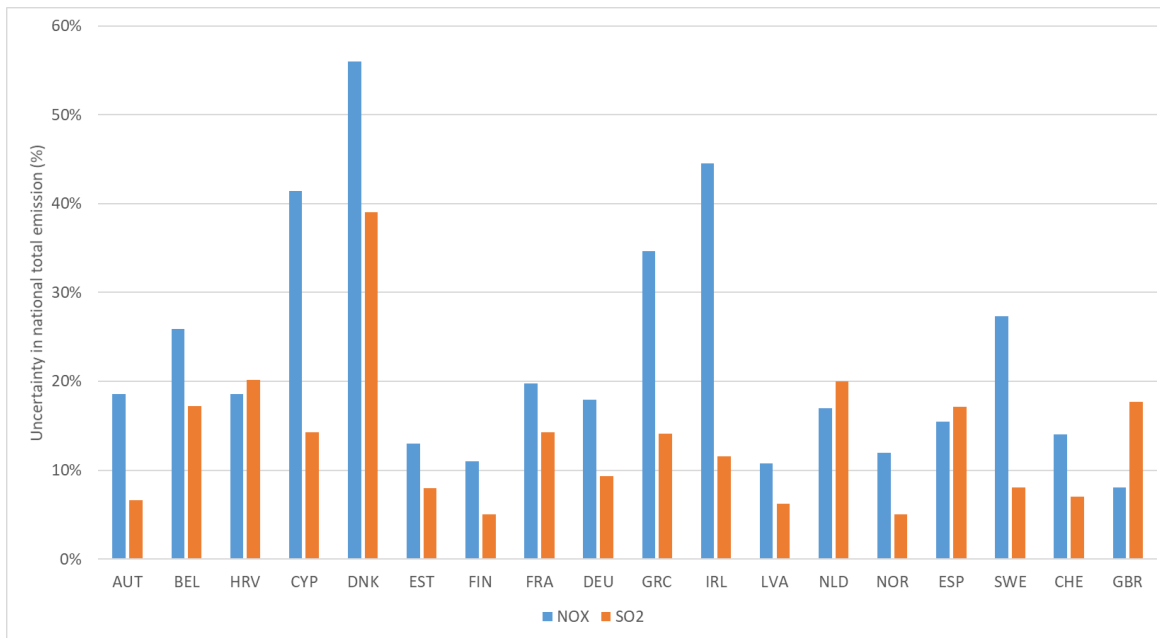



Figure 12 Uncertainties in total emissions of NOₓ and SO₂ as reported in the Informative Inventory Reports submitted by the countries in 2020 (data taken from (Schindlbacher et al., 2021))

The uncertainties shown in Figure 12 are generally lower than those reported in Table 5, since the emission inventories for these countries are well established, in many cases using directly measured emission data for important sources, which reduces the overall uncertainty. A large variation is shown between uncertainties for different countries, even for the same pollutant. This relates to the methodology and level of detail in which these uncertainty assessments are done and reported on differ per country. Also, the estimation of uncertainties for individual parameters such as activity data or emission factors often requires a significant amount of expert judgement in the absence of real data on uncertainties. Since countries perform these uncertainty assessment individually, resulting overall uncertainty estimates differ significantly per country, and applying the available uncertainty estimates directly to a European scale inventory will introduce inconsistencies in these uncertainties between countries. This poses significant shortcomings in the direct uptake of uncertainty data from country emission inventories in the European-wide uncertainty assessment.

An alternative to using country data on uncertainties would be to perform a complete uncertainty assessment. This requires uncertainties to be estimated for all parameters involved in the emission estimation, including activity data, emission factors, and something that accounts for the uncertainty in spatial distribution. A first attempt to quantify uncertainties in emissions in such a way was recently made by Super et al. (2020) for CO_2 and CO. Whereas the annual country-level CO_2 emissions have a relatively low uncertainty, for CO the emission factors had the largest contribution to the uncertainties. Also the uncertainty in the spatial proxies was assessed, showing a significant contribution to the uncertainty in the gridded inventory, especially at higher resolution. Another example of estimating uncertainties for greenhouse gases at global scale was made recently for the EDGAR inventory (Solazzo et al., 2021), but this study did not consider the uncertainty in the spatial distribution component. Given the variety of datasets used in CAMS-REG, the focus on air pollutants which have more

uncertain EFs compared to greenhouse gases and the almost 20-year time span, more work is needed to independently estimate emission inventory uncertainties, including spatial error correlations. This could be considered a future priority to develop.

Given the difficulty in deriving direct uncertainties of the emissions at grid level, comparisons to other
780   independent emission datasets are a useful way to identify key differences and apparent uncertainties. In this paper it is shown that EDGAR-v5.0 emissions differ significantly from CAMS-REG emissions, especially for non-EU countries (Table 4). Such differences could be regarded as indicative of the uncertainties associated with present-day anthropogenic emissions but this is a topic that deserves more attention and effort in the future, as mentioned earlier.

785   Also a comparison to satellite-based emission estimates can be helpful. In recent years many studies have attempted to estimate anthropogenic emissions or verify emissions inventories based on a combination of satellite data interpretation and (inverse) modelling approach, e.g. (Dammers et al., 2019; Goldberg et al., 2019; Lorente et al., 2019; Szymankiewicz et al., 2021). Most satellite studies focus on applications outside Europe given the limited availability of emission inventories in these regions, and the focus is mostly on those pollutants where
790   satellite retrieval products are relatively well established and where clear gradients in the emissions are expected from point sources (e.g. NO_x, SO_2). There are clear advantages when using space-based information, which mainly lie in the consistency (no methodological differences between countries or regions) and the high spatial and temporal resolution compared to the annual emission inventories. However, the satellite-based assessment has its own uncertainties because of the data interpretation and modelling involved. Also, the satellite only sees
795   the total for a pollutant (no sector breakdown of emissions).

For the comparison to the earlier versions of this dataset it is found that most of the differences in the emission estimates can be traced back to differences in national emission reporting. European countries are obliged to report their emissions in the national inventories for each year in the time series annually (back to 1990), and at the same time every year improvements and updates are made to the inventories, incorporating new information
800   on activity data or emissions factors. This means that every year the entire time series is revised, which may incur significant changes to the overall emissions in each country. Since each country has its own inventory team with its own challenges and data sources, such revisions do not always go in the same direction.

One example of large changes to earlier reported emissions over time can be found in the PM emissions from small residential combustion. The main reason for adjustments over time is the increasing insight and awareness
805   of the role that condensable, mostly organic compounds play in total PM emission from this source sector. These compounds are emitted in gaseous phase, but immediately after leaving the stack or chimney may condense to form particles. Depending on the measurement device used, these particles may or may not be captured in the PM emission measurement (Denier van der Gon et al., 2015). In the EMEP/EEA Guidebook (EEA, 2019a), the reference document for compiling air pollutant inventories in Europe, condensables were consistently introduced
810   for small combustion of biomass as part of the 2016 update, which gradually encouraged more and more countries to report on this basis. The main motivation for doing so is the better understanding of ambient PM as supported by better prediction skills of air quality models and better agreement with observations (Bergström et al., 2012; Simpson et al., 2020). Figure 9 shows $PM_{2.5}$ emissions reported in this inventory and 2 of its predecessors, where the emission data were based on reported data as reported in 2013, 2017 and 2019,

815   respectively. The differences between the datasets are representative for the differences in reporting from the inventories, which in turn are to a large extent related to the inclusion (or not) of condensables. Amongst others, it can be observed that between 2013 and 2017 the reported emissions in e.g. Belgium, the United Kingdom, Spain, Italy and Romania increased significantly, which was confirmed to be caused by the inclusion of condensables in these cases.

820   Given that the CAMS-REG inventories are derived to support air quality assessment at European scale, not all the details for specific national circumstances may be included. While specific national circumstances are incorporated for as far as sector totals are concerned (through the direct use of country reported data for EU+ countries), the spatial distribution methodology is uniform over Europe whereas individual countries may use detailed distribution proxies specific to their country. This means that when zooming in, differences with these
825   national distributions and also with local bottom-up estimates of emissions are typically found. These are to a large extent related to the spatial distribution component (Trombetti et al., 2018).

Due to the consistency in which country reported data are being processed the CAMS-REG datasets can also be used to investigate trends and derive information on specific sources. An example is the analysis presented by (Denier van der Gon et al., 2018) on the increasing importance of non-exhaust PM emissions from road transport.
830   This is a separate source category in CAMS-REG-v4.2 (GNFR sector F4). Stringent EU policies over time succeeded in reducing the road transport exhaust PM emissions and by now "non-exhaust" emissions from brake wear, tyre wear and road abrasion have started to dominate the PM_{10} emission from road transport. As the share of coarse PM is relatively high in wear emission, the $PM_{2.5}$ emission from road transport may still be dominated by exhaust emissions. An important aspect of having a separate category in the emission data for wear emissions
835   is that different chemical composition profiles can be applied, for instance to estimate various heavy metal emissions from non-exhaust road transport emissions (Denier van der Gon et al., 2018).

The CAMS-REG inventory is widely used in modelling activities worldwide because of its consistent approach and longer time series it is especially useful to support large intercomparison activities between models such as the Air Quality Modelling International Initiative (AQMEII) (Im et al., 2015). Another example where CAMS-
840   REG is used is the HTAP (Hemispheric Transport of Air Pollution) inventory (Janssens-Maenhout et al., 2015). Here, a global emission inventory (EDGAR) was updated by nesting specific regional inventories in specific regions of the world (e.g. CAMS-REG) to improve our understanding hemispheric transport of air pollution. Finally, with the increasing resolution of emission inventories to support modelling exercises at local to regional scale, the temporal distribution of emissions within the year becomes increasingly important. A specific set of
845   temporal profiles was derived specifically to be applied with the CAMS emission inventories including CAMS-REG. These profiles take into account more detailed temporal variations compared to the default temporal profiles provided along with this dataset, such as the variation of emissions with meteorology (e.g. temperature dependency of residential heating) (Guevara et al., 2021).

5   Data availability

850   Gridded emission maps with all pollutants are available for each year. The files are provided as NetCDF (Network Common Data Format) at a resolution of 0.05°x0.1° (latitude-longitude) for the European domain

(30°N-72°N, 30°W-60°E) (Kuenen et al., 2021). The emission data in the grids represent annual data per grid cell, all emissions are given in kg. Access is provided through the Emissions of atmospheric Compounds of Ancillary Data (ECCAD) system, which will be complemented with access through the ECMWF Atmosphere
855 Data Store (ADS) as soon as this is technically feasible. Since the ECCAD system requires a registration and login, for the purpose of the review process of this manuscript a sample of the emission files has been made available for download directly. This sample includes data for the year 2017 and is available through https://eccad.aeris-data.fr/essd-surf-emis-cams-reg/.

6   Conclusions and outlook

860 The current CAMS-REG-v4.2 emission inventory is developed in support of air quality modelling activities at European scale. It incorporates official national emission estimates from countries to the extent possible to facilitate the use of this inventory for policy applications and uses a uniform spatial distribution methodology across Europe to ensure a comparable and consistent emission grid across Europe. CAMS-REG-v4.2 is the latest version of the a series of emission inventories that was developed in support of modelling. Since in addition to
865 the gridded emissions, speciation profiles for PM and NMVOC as well as default information on temporal and height distribution are provided along with the dataset, it provides an excellent starting point for air quality modelling at European scale. On the other hand, the use of country reported data also implies that for some sectors there are limitations when the consistency in reporting is limited. Examples include PM emissions from residential combustion, but also agricultural NMVOC emissions (which are currently excluded from the CAMS-
870 REG inventory). These specific sources should be looked at in the future to work towards a consistent representation of these sources in the CAMS-REG inventory. Also non-reported sources such as PM from resuspension could be considered. In addition to that, developing uncertainties for the emissions in this dataset is a difficult task given the combination of different data sources and the limited availability of uncertainty estimates for these datasets. However, it will be important for the years to come to assess the uncertainties of
875 modelled concentrations and also to compare emission estimates from inventories to those derived from observations.

Supplement

The supplement related to this article is available online at (to be added)

880

Author contributions

JK coordinated and processed all the different emission data sources, with support from SD and IS. Methodological decisions and choices were made through discussions with AV and HvdG. The spatial distribution proxies were mostly prepared by SD with the help of AV, IS and JK. SD and AV developed the point
885 source databases used in this study. HvdG gave feedback on the whole inventory development and steered the directions. JPJ provided the shipping emissions. JK prepared the paper, with specific input sections from SD, AV and HvdG. IS reviewed the paper as a whole prior to submission.

Competing interests

890 The authors declare that they have no conflict of interest.

Special issue statement

(to be added)

895 **Acknowledgements**

The research leading to these results has received funding from:

- the Copernicus Atmosphere Monitoring Service (CAMS), which is implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Commission.
- The European Union's Horizon 2020 research and innovation programme under grant agreement No. 900   776186 (CHE project, coordinated by ECMWF).
- The European Union's Horizon 2020 research and innovation programme under grant agreement No. 776810 (VERIFY project, coordinated by CEA/LSCE).

The authors would like to give special thanks to the GFAS team and Johannes Kaiser for providing the spatially distributed emissions from agricultural waste burning, and the discussion regarding their use and uptake in 905 emission inventories.

The authors would like to express their thanks to CEIP for making available reported data of air pollutant emissions by all European countries, and to IIASA for making available GAINS emissions through its online tool. The EU Joint Research Centre (in particular Adrian Leip) is thanked for providing spatially distributed proxy maps that are used for agricultural emissions.

910

7   References

A barrel full: European refineries, [online] Available from: http://abarrelfull.wikidot.com/european-refineries, n.d.

Belis, C. A., Pernigotti, D., Pirovano, G., Favez, O., Jaffrezo, J. L., Kuenen, J., Denier van Der Gon, H., Reizer, 915   M., Riffault, V., Alleman, L. Y., Almeida, M., Amato, F., Angyal, A., Argyropoulos, G., Bande, S., Beslic, I., Besombes, J.-L., Bove, M. C., Brotto, P., Calori, G., Cesari, D., Colombi, C., Contini, D., De Gennaro, G., Di Gilio, A., Diapouli, E., El Haddad, I., Elbern, H., Eleftheriadis, K., Ferreira, J., Vivanco, M. G., Gilardoni, S., Golly, B., Hellebust, S., Hopke, P. K., Izadmanesh, Y., Jorquera, H., Krajsek, K., Kranenburg, R., Lazzeri, P., Lenartz, F., Lucarelli, F., Maciejewska, K., Manders, A., Manousakas, M., Masiol, M., Mircea, M., Mooibroek,

920 D., Nava, S., Oliveira, D., Paglione, M., Pandolfi, M., Perrone, M., Petralia, E., Pietrodangelo, A., Pillon, S., Pokorna, P., Prati, P., Salameh, D., Samara, C., Samek, L., Saraga, D., Sauvage, S., Schaap, M., Scotto, F., Sega, K., Siour, G., Tauler, R., Valli, G., Vecchi, R., Venturini, E., Vestenius, M., Waked, A. and Yubero, E.: Evaluation of receptor and chemical transport models for PM10 source apportionment, Atmos. Environ. X, 5, 100053, doi:https://doi.org/10.1016/j.aeaoa.2019.100053, 2020.

925 Bergström, R., van der Gon, H. A. C., Prévôt, A. S. H., Yttri, K. E. and Simpson, D.: Modelling of organic aerosols over Europe (2002–2007) using a volatility basis set (VBS) framework: application of different assumptions regarding the formation of secondary organic aerosol, Atmos. Chem. Phys., 12(18), 8499–8527, doi:10.5194/acp-12-8499-2012, 2012.

Bieser, J., Aulinger, A., Matthias, V., Quante, M. and Denier van der Gon, H. A. C.: Vertical emission profiles 930 for Europe based on plume rise calculations, Environ. Pollut., 159(10), 2935–2946, doi:https://doi.org/10.1016/j.envpol.2011.04.030, 2011.

Buonocore, J. J., Salimifard, P., Michanowicz, D. R. and Allen, J. G.: A decade of the U.S. energy mix transitioning away from coal: historical reconstruction of the reductions in the public health burden of energy, Environ. Res. Lett., 16(5), 54030, doi:10.1088/1748-9326/abe74c, 2021.

935 CAPRI: CAPRI Modelling System, [online] Available from: https://www.capri-model.org/dokuwiki/doku.php?, 2020.

CEIP: Officially reported activity data, [online] Available from: http://www.ceip.at/ms/ceip_home1/ceip_home/webdab_emepdatabase/reported_activitydata/, 2019.

Copernicus Land Monitoring Service: CORINE Land Cover 2012, [online] Available from: 940 https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012, 2016.

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(4), 1987–2013, doi:10.5194/essd-10-1987-2018, 2018.

Crippa, M., Guizzardi, D., Oreggioni, G., Muntean, M., Schaaf, E., Thunis, P., Cuvelier, C., de Meij, A. and 945 Pisoni, E.: EDGAR v5.0, [online] Available from: https://data.europa.eu/doi/10.2904/JRC_DATASET_EDGAR, 2020.

Dammers, E., McLinden, C. A., Griffin, D., Shephard, M. W., Van Der Graaf, S., Lutsch, E., Schaap, M., Gainairu-Matz, Y., Fioletov, V., Van Damme, M., Whitburn, S., Clarisse, L., Cady-Pereira, K., Clerbaux, C., Coheur, P. F. and Erisman, J. W.: $\chem{NH_3}$ emissions from large point sources derived from CrIS and IASI 950 satellite observations, Atmos. Chem. Phys., 19(19), 12261–12293, doi:10.5194/acp-19-12261-2019, 2019.

Denier van der Gon, H., Hulskotte, J., Jozwicka, M., Kranenburg, R., Kuenen, J. and Visschedijk, A.: Chapter 5 - European Emission Inventories and Projections for Road Transport Non-Exhaust Emissions: Analysis of Consistency and Gaps in Emission Inventories From EU Member States, in Non-Exhaust Emissions, edited by F. Amato, pp. 101–121, Academic Press., 2018.

955 Denier van der Gon, H. A. C., Bergström, R., Fountoukis, C., Johansson, C., Pandis, S. N., Simpson, D. and Visschedijk, A. J. H.: Particulate emissions from residential wood combustion in Europe – revised estimates and an evaluation, Atmos. Chem. Phys., 15(11), 6503–6519, doi:10.5194/acp-15-6503-2015, 2015.

Denier Van Der Gon, H. A. C., Visschedijk, A. J. H., Van den Brugh, H. and Dröge, R.: A high resolution European emission database for the year 2005. [online] Available from:

960 https://www.umweltbundesamt.de/sites/default/files/medien/461/publikationen/texte_41_2013_appelhans_e03_komplett_0.pdf, 2010.

Doumbia, T., Granier, C., Elguindi, N., Bouarar, I., Darras, S., Brasseur, G., Gaubert, B., Liu, Y., Shi, X., Stavrakou, T., Tilmes, S., Lacey, F., Deroubaix, A. and Wang, T.: Changes in global air pollutant emissions during the COVID-19 pandemic: a dataset for atmospheric chemistry modeling, Earth Syst. Sci. Data Discuss.,

965 2021, 1–26, doi:10.5194/essd-2020-348, 2021.

EEA: Urban Waste Water Treatment Directive (UWWTD) - reported data, [online] Available from: https://www.eea.europa.eu/data-and-maps/data/waterbase-uwwtd-urban-waste-water-treatment-directive-4, 2014.

EEA: EMEP/EEA Air Pollutant Emission Inventory Guidebook. [online] Available from: https://www.eea.europa.eu/publications/emep-eea-guidebook-2019, 2019a.

970 EEA: European Pollutant Release and Transfer Register, [online] Available from: https://industry.eea.europa.eu/, 2019b.

Eggleston, S., Buendia, L., Miwa, K., Ngara, T. and Tanabe, K.: 2006 IPCC guidelines for national greenhouse gas inventories, Institute for Global Environmental Strategies Hayama, Japan., 2006.

EMEP: Status report 1/2017: Transboundary particulate matter, photo-oxidants, acidifying and eutrophying

975 components. [online] Available from: https://emep.int/publ/reports/2017/EMEP_Status_Report_1_2017.pdf, 2017.

EMEP: Status report 1/2020: Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components. [online] Available from: https://emep.int/publ/reports/2020/EMEP_Status_Report_1_2020.pdf, 2020.

980 European Commission: Directive 2001/80/EC of the European Parliament and of the Council of 23 October 2001 on the limitation of emissions of certain pollutants into the air from large combustion plants, [online] Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32001L0080, 2001.

European Commission: Directive 2010/75/EU of the European Parliament and of the Council of 24 November 2010 on industrial emissions (integrated pollution prevention and control), [online] Available from: https://eur-

985 lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32010L0075, 2010.

European Commission: Directive (EU) 2016/2284 on the reduction of national emissions of certain atmospheric pollutants, [online] Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.344.01.0001.01.ENG&toc=OJ:L:2016:344:TOC, 2016.

Eurostat: Aircraft traffic data by main airport [avia_tf_aca], [online] Available from:

990 http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=avia_tf_aca&lang=en, 2019.

FAO: Gridded Livestock of the World (GLW3), [online] Available from: http://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1236449/, 2010.

Goldberg, D. L., Lu, Z., Streets, D. G., de Foy, B., Griffin, D., McLinden, C. A., Lamsal, L. N., Krotkov, N. A. and Eskes, H.: Enhanced Capabilities of TROPOMI NO2: Estimating NOX from North American Cities and

995   Power Plants, Environ. Sci. Technol., 53(21), 12594–12601, doi:10.1021/acs.est.9b04488, 2019.

Guevara, M., Jorba, O., Tena, C., van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C. and Pérez Garc\'\ia-Pando, C.: Copernicus Atmosphere Monitoring Service TEMPOral profiles (CAMS-TEMPO): global and European emission temporal profile maps for atmospheric chemistry modelling, Earth Syst. Sci. Data, 13(2), 367–404, doi:10.5194/essd-13-367-2021, 2021.

1000   Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., Guizzardi, D., Muntean, M., Schaaf, E. and Friedrich, R.: Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970--2012, Atmos. Chem. Phys., 17(12), 7683–7701, doi:10.5194/acp-17-7683-2017, 2017.

IIASA: Progress towards the achievement of the EU's Air Quality and Emission Objectives. [online] Available
1005   from: https://ec.europa.eu/environment/air/pdf/clean_air_outlook_overview_report.pdf, 2018.

IIASA: Horizontal report for the third phase of the review of national air pollution inventory data. [online] Available from: https://ec.europa.eu/environment/air/documents/NECReview2019horizontal-report_FINAL.doc, 2019.

Im, U., Bianconi, R., Solazzo, E., Kioutsioukis, I., Badia, A., Balzarini, A., Baró, R., Bellasio, R., Brunner, D.,
1010   Chemel, C., Curci, G., Flemming, J., Forkel, R., Giordano, L., Jiménez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L., Jorba, O., Knote, C., Kuenen, J. J. P., Makar, P. A., Manders-Groot, A., Neal, L., Pérez, J. L., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Tuccella, P., Werhahn, J., Wolke, R., Yahya, K., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C., Galmarini, S., Denier van der Gon, H., Flemming, J., Forkel, R., Giordano, L., Jiménez-Guerrero, P., Hirtl, M., Hodzic, A., Honzak, L.,
1015   Jorba, O., Knote, C., Makar, P. A., Manders-Groot, A., Neal, L., Pérez, J. L., Pirovano, G., Pouliot, G., San Jose, R., Savage, N., Schroder, W., Sokhi, R. S., Syrakov, D., Torian, A., Tuccella, P., Wang, K., Werhahn, J., Wolke, R., Zabkar, R., Zhang, Y., Zhang, J., Hogrefe, C. and Galmarini, S.: Evaluation of operational online-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part II: Particulate matter, Atmos. Environ., 115, 421–441, doi:https://doi.org/10.1016/j.atmosenv.2014.08.072, 2015.

1020   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(5), 2641–2659, doi:10.5194/acp-12-2641-2012, 2012.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Dentener, F., Muntean, M., Pouliot, G., Keating, T., Zhang, Q., Kurokawa, J., Wankmüller, R., Denier Van Der Gon, H., Kuenen, J. J. P., Klimont, Z., Frost, G., Darras, S.,
1025   Koffi, B. and Li, M.: HTAP-v2.2: A mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution, Atmos. Chem. Phys., 15(19), 11411–11432, doi:10.5194/acp-15-11411-2015, 2015.

Johansson, L., Jalkanen, J.-P. and Kukkonen, J.: Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution, Atmos. Environ., 167, 403–415,
1030   doi:https://doi.org/10.1016/j.atmosenv.2017.08.042, 2017.

Jonson, J. E., Gauss, M., Schulz, M., Jalkanen, J.-P. and Fagerli, H.: Effects of global ship emissions on European air pollution levels, Atmos. Chem. Phys., 20(19), 11399–11422, doi:10.5194/acp-20-11399-2020, 2020.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M.,
1035 Schultz, M. G., Suttie, M. and van der Werf, G. R.: Biomass burning emissions estimated with a global fire
assimilation system based on observed fire radiative power, Biogeosciences, 9(1), 527–554, doi:10.5194/bg-9-
527-2012, 2012.

Klimont, Z., Kupiainen, K., Heyes, C., Purohit, P., Cofala, J., Rafaj, P., Borken-Kleefeld, J. and Schöpp, W.:
Global anthropogenic emissions of particulate matter including black carbon, Atmos. Chem. Phys., 17(14), 8681–
1040 8723, doi:10.5194/acp-17-8681-2017, 2017.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I. and Denier van der Gon, H.: Copernicus
Atmosphere Monitoring Service regional emissions version 4.2 (CAMS-REG-v4.2), , doi:10.24380/0vzb-a387,
2021.

Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M. and Denier Van Der Gon, H. A. C.: TNO-MACC-II
1045 emission inventory; A multi-year (2003-2009) consistent high-resolution European emission inventory for air
quality modelling, Atmos. Chem. Phys., 14(20), 10963–10976, doi:10.5194/acp-14-10963-2014, 2014.

Liang, C.-K., West, J. J., Silva, R. A., Bian, H., Chin, M., Davila, Y., Dentener, F. J., Emmons, L., Flemming, J.,
Folberth, G., Henze, D., Im, U., Jonson, J. E., Keating, T. J., Kucsera, T., Lenzen, A., Lin, M., Lund, M. T., Pan,
X., Park, R. J., Pierce, R. B., Sekiya, T., Sudo, K. and Takemura, T.: HTAP2 multi-model estimates of premature
1050 human mortality due to intercontinental transport of air pollution and emission sectors, Atmos. Chem. Phys.,
18(14), 10497–10520, doi:10.5194/acp-18-10497-2018, 2018.

Lorente, A., Boersma, K. F., Eskes, H. J., Veefkind, J. P., van Geffen, J. H. G. M., de Zeeuw, M. B., Denier van
der Gon, H. A. C., Beirle, S. and Krol, M. C.: Quantification of nitrogen oxides emissions from build-up of
pollution over Paris with TROPOMI, Sci. Rep., 9(1), 20033, doi:10.1038/s41598-019-56428-5, 2019.

1055 McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa,
M., Brauer, M. and Martin, R. V: A global anthropogenic emission inventory of atmospheric pollutants from
sector- and fuel-specific sources (1970--2017): an application of the Community Emissions Data System
(CEDS), Earth Syst. Sci. Data, 12(4), 3413–3442, doi:10.5194/essd-12-3413-2020, 2020.

Ntziachristos, L., Gkatzoflias, D., Kouridis, C. and Samaras, Z.: COPERT: A European Road Transport Emission
1060 Inventory Model, in Information Technologies in Environmental Engineering, edited by I. N. Athanasiadis, A. E.
Rizzoli, P. A. Mitkas, and J. M. Gómez, pp. 491–504, Springer Berlin Heidelberg, Berlin, Heidelberg., 2009.

Oak Ridge National Laboratory: LandScan, [online] Available from: https://landscan.ornl.gov/, 2017.

Olivier, J. G. J., Bouwman, A. F., Van der Maas, C. W. M., Berdowski, J. J. M., Veldt, C., Bloos, J. P. J.,
Visschedijk, A. J. H., Zandveld, P. Y. J. and Haverlag, J. L.: Description of EDGAR version 2.0. [online]
1065 Available from: https://www.rivm.nl/bibliotheek/rapporten/771060002.pdf, 1996.

Open Street Map: Open Street Map, [online] Available from: https://www.openstreetmap.org/, 2017.

OpenTransportMap: Open Transport Map, [online] Available from: http://opentransportmap.info/, 2017.

Platts: World Electric Power Plants database, [online] Available from:
https://www.spglobal.com/platts/en/commodities/electric-power, 2017.

1070 Pouliot, G., van der Gon, H. A. C., Kuenen, J., Zhang, J., Moran, M. D. and Makar, P. A.: Analysis of the

emission inventories and model-ready emission datasets of Europe and North America for phase 2 of the AQMEII project, Atmos. Environ., 115, 345–360, doi:10.1016/j.atmosenv.2014.10.061, 2015.

Schindlbacher, S., Matthews, B. and Ullrich, B.: Uncertainties and recalculations of emission inventories submitted under CLRTAP. [online] Available from:
1075   https://www.ceip.at/fileadmin/inhalte/ceip/00_pdf_other/2021/uncertainties_and_recalculations_of_emission_inv entories_submitted_under_clrtap.pdf, 2021.

Simpson, D., Fagerli, H., Colette, A., Denier van der Gon, H., Dore, C., Hallquist, M., Hansson, H. C., Maas, R., Rouil, L., Allemand, N., Bergström, R., Bessagnet, B., Couvidat, F., El Haddad, I., Genberg, J., Goile, F., Grieshop, A., Fraboulet, I., Hallquist, A., Hamilton, J., Juhrich, K., Klimont, Z., Kregar, Z., Mawdsely, I.,
1080   Megaritis, A., Ntziachristos, L., Pandis, S., Prevot, A., Schindlbacher, S., Seljeskog, M., Sirina-Leboine, N., Sommers, J. and Astrom, S.: How should condensables be included in PM emission inventories reported to EMEP/CLRTAP? [online] Available from:
https://emep.int/publ/reports/2020/emep_mscw_technical_report_4_2020.pdf, 2020.

Solazzo, E., Crippa, M., Guizzardi, D., Muntean, M., Choulga, M. and Janssens-Maenhout, G.: Uncertainties in
1085   the Emissions Database for Global Atmospheric Research (EDGAR) emission inventory of greenhouse gases, Atmos. Chem. Phys., 21(7), 5655–5683, doi:10.5194/acp-21-5655-2021, 2021.

Szymankiewicz, K., Kaminski, J. W. and Struzewska, J.: Application of Satellite Observations and Air Quality Modelling to Validation of NOx Anthropogenic EMEP Emissions Inventory over Central Europe, Atmosphere (Basel)., 12(11), doi:10.3390/atmos12111465, 2021.

1090   Theloke, J. and Friedrich, R.: Compilation of a database on the composition of anthropogenic VOC emissions for atmospheric modeling in Europe, Atmos. Environ., 41(19), 4148–4160, doi:https://doi.org/10.1016/j.atmosenv.2006.12.026, 2007.

Timmermans, R. M. A., van der Gon, H. A. C., Kuenen, J. J. P., Segers, A. J., Honoré, C., Perrussel, O., Builtjes, P. J. H. and Schaap, M.: Quantification of the urban air pollution increment and its dependency on the use of
1095   down-scaled and bottom-up city emission inventories, Urban Clim., 6, 44–62, doi:10.1016/j.uclim.2013.10.004, 2013.

Trombetti, M., Thunis, P., Bessagnet, B., Clappier, A., Couvidat, F., Guevara, M., Kuenen, J. and López-Aparicio, S.: Spatial inter-comparison of Top-down emission inventories in European urban areas, Atmos. Environ., 173, 142–156, doi:10.1016/j.atmosenv.2017.10.032, 2018.

1100   UNECE: 1999 Protocol to Abate Acidification, Eutrophication and Ground-level Ozone to the Convention on Long-range Transboundary Air Pollution, as amended on 4 May 2012, [online] Available from: https://www.unece.org/fileadmin/DAM/env/documents/2013/air/eb/ECE.EB.AIR.114_ENG.pdf, 2012.

UNFCCC: National Inventory Submissions 2019, [online] Available from: https://unfccc.int/process-and-meetings/transparency-and-reporting/reporting-and-review-under-the-convention/greenhouse-gas-inventories-
1105   annex-i-parties/national-inventory-submissions-2019, 2019.

Viana, M., Hammingh, P., Colette, A., Querol, X., Degraeuwe, B., Vlieger, I. de and van Aardenne, J.: Impact of maritime transport emissions on coastal air quality in Europe, Atmos. Environ., 90, 96–105, doi:https://doi.org/10.1016/j.atmosenv.2014.03.046, 2014.

Visschedijk, A. J. H., Denier van der Gon, H. A. C., Dröge, R. and Van der Brugh, H.: A European high
1110 resolution and size-differentiated emission inventory for elemental and organic carbon for the year 2005., 2009.

Wankmüller, R.: Updated documentation of the EMEP gridding system, Technical Report CEIP 6/2019. [online] Available from:
https://www.ceip.at/fileadmin/inhalte/ceip/00_pdf_other/2019/emep_gridding_system_documentation_20191125
.pdf, 2019.

1115 World Bank: Urban population, [online] Available from:
https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=EU (Accessed 24 June 2021), 2018.

World Steel Dynamics: Plantfacts, Global Steel Information System, [online] Available from:
https://gsis.worldsteeldynamics.com/, n.d.

1120