CoCO2

Prototype system for a Copernicus CO_2 service

Functional Requirements Specification Document

Frédéric Chevallier (CEA), Nicolas Bousserez (ECMWF), Thomas Kaminski (iLab), Grégoire Broquet (CEA), Audrey Fortems-Cheiney (CEA), Marine Remaud (CEA), Diego Santaren (CEA), Marko Scholze (ULUND), Luca Cantarello (ECMWF), Wolfgang Knorr (iLab), Peter Rayner (iLab), Michael Voßbeck (iLab)

Co-ordinated by ECMWF

CoCO2

Prototype system for a Copernicus CO_2 service

D6.4Functional Requirements Specification Document

Dissemination Level:	Public
Author(s):	Frédéric Chevallier (CEA), Nicolas Bousserez (ECMWF), Thomas Kaminski (iLab), Grégoire Broquet (CEA), Audrey Fortems-Cheiney (CEA), Marine Remaud (CEA), Diego Santaren (CEA), Marko Scholze (ULUND), Luca Cantarello (ECMWF), Wolfgang Knorr (iLab), Peter Rayner (iLab), Michael Voßbeck (iLab)
Date:	15/02/2023
Version:	1.1
Contractual Delivery Date:	31/12/2022
Work Package/ Task:	WP6/ T6.3 and T6.4
Document Owner:	ECMWF
Contributors:	Konstantinos Politakos (FORTH), Stavros Stagakis (FORTH) and Jean-Christophe Calvet (MF)
Status:	Final

CoCO2: Prototype system for a Copernicus CO_2 service

Coordination and Support Action (CSA)
H2020-IBA-SPACE-CHE2-2019Copernicus evolution – Research activities in support of a Europeanoperational monitoring support capacity for fossil CO2 emissions

Project Coordinator: Dr Richard Engelen (ECMWF)
Project Start Date: 01/01/2021
Project Duration: 36 months

Published by the CoCO2 Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX,
richard.engelen@ecmwf.int

The CoCO2 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958927.

Table of Contents

1 Introduction

The European Commission is establishing an operational observation-based anthropogenic CO_2 emissions monitoring and verification support capacity (CO2MVS) as part of its Copernicus Earth Observation programme. Demonstrator systems for this CO2MVS are being developed in the Prototype System for a Copernicus CO_2 service (CoCO2) project.

In its information to the 1st Global Stocktake entitled *Data contribution of the European CoCO2 project to the first Global StockTake* (Deliverable D6.5), CoCO2 introduced data from five demonstrator systems, each exploring and prefiguring different aspects of the future CO2MVS. The first three are specifically about fossil fuel emissions:

- **local** large fossil fuel CO_2 emissions estimated from observed CO_2 plume cross-sections
- **regional** fossil fuel CO_2 emissions estimated by an atmospheric inversion assimilating satellite retrievals of co-emitted species
- **global** fossil fuel CO_2 emissions estimated by a Carbon Cycle Fossil Fuel Data Assimilation System

The fourth demonstrator data is about the Agriculture, Forestry and Other Land Use sector (AFOLU)

- **global** AFOLU emissions estimated by CO_2 atmospheric inversions

The fifth demonstrator addresses all global emissions and absorptions synergistically:

- **global** data from an extension of the ECMWF Numerical Weather Prediction system

The present Functional Requirements Specification Document describes the codes that have been used to generate Figures 2-7 of D6.5. It explains what these codes do and how they do it. It explains how the user requirements identified by CoCO2[1] have been addressed from a technical point of view by each demonstrator. The references for each demonstrator are given directly at the end of the section where it is described.

2 Estimates of large point source fossil fuel CO_2 emissions based on satellite observations

2.1 Introduction

The algorithm follows the method successively developed by Zheng et al (2020) and Chevallier et al (2020, 2022). It is a fully automated process without any manual selection, giving it the advantage of being objective and making it fit for an operational application. Its implementation to generate Figure 2 of the CoCO2 submission to the GST is described in the following, as an update of the text of Chevallier et al (2022). Computations were performed on the High Performance Computing resources of https://www-hpc.cea.fr/ under the allocation A0110102201 made by Grand Equipement National de CalculIntensif (GENCI).

2.2 CO_2column retrievals

The name *Orbiting Carbon Observatory-2* (OCO-2) designates both a specific spacecraft in sun-synchronous polar orbit and the single sounding instrument it carries. This instrument

[1] https://www.coco2-project.eu/sites/default/files/2022-03/CoCO2-D6-3-V1-0.pdf

incorporates three boresighted spectrometers. “OCO-3” designates a similar instrument to that of OCO-2, but is installed as an external payload on the International Space Station (ISS). OCO-3 is therefore not autonomous, and follows viewing and operation constraints from the ISS. This disadvantage is offset by a new pointing mirror assembly which allows for much more agile vision than OCO-2. The ISS has a 12% faster ground-track velocity than OCO-2 (7.667 km s^{-1} vs. 6.75 km s^{-1}; see Eldering et al., 2019, and Baker et al., 2022), resulting in slightly less spatially dense retrievals, except when a specific scanning mode is enabled to perform side-by-side scans and provide snapshot area maps (SAMs) of ~80 × 80 km^2. The soundings of the two OCOs are therefore distributed differently in space and time, but their footprint and precision are similar (~3 km^2 and better than 1 ppm, respectively).

For the two OCO instruments, we use all operational quality-controlled bias-corrected retrievals of the column-average CO_2 dry air-mole fraction (XCO_2) with associated uncertainty estimate over both land and ocean made by the Atmospheric CO_2 Observations from Space (ACOS) algorithm of NASA (O’Dell et al., 2018). For OCO-2, a distinction is made between the emissions estimated from version 10 of NASA’s official column-retrieval algorithm and version 11. Version 10 covers the full OCO-2 record from September 2014 until February 2022, while version 11 has taken over since then, but with only a portion of the past OCO-2 record having been reanalysed by NASA so far. Version 11 generates more data than version 10, and with enhanced quality, so that the two data streams are not fully comparable. For OCO-3, version 10 is the latest processing to date and covers the period from August 2019 until October 2022. The OCO-2 and OCO-3 retrievals were produced by the OCO-2 and OCO-3 project at the Jet Propulsion Laboratory, California Institute of Technology.

We neglect the impact of the deviation of the column averaging kernel from an ideal vector of ones. When computing the XCO_2 line densities, our emission retrieval also uses the a priori surface pressure and the a posteriori total column water vapour of the ACOS data.

2.3 Estimating the CO2 emissions from the burning of fossil fuels

We process data from both instruments in exactly the same way.

Each OCO orbit is analysed with a 200 km moving window successively centred on each of the validated retrieval. The width of the moving window defines the maximum size of enhancements allowed. If a retrieval value stands out of the variability of the retrievals in the window, the procedure attempts to fit a function that represents a bell curve on top of a linear background, on the soundings located over the dominant surface type in the window. For the SAM observation mode of OCO-3, we exclude from the fit the retrievals that are not in the same scan line, in order to stay within a cross-sectional vision. The soundings are represented by their XCO_2 retrievals and by their position along the satellite track: their across-track position is ignored, the satellite narrow swath being used only to damp retrieval noise. The restriction to the soundings of the dominant surface type (land or water) mainly skips the artificial discontinuities in the retrievals linked to the corresponding changes in the surface albedo. The adjustable parameters are the standard deviation and the height of the Gaussian, and the slope and offset of the line. The centre of the peak is positioned in the middle of the 200 km moving window and is not adjusted. The choice of the Gaussian form is linked to its neutrality (in the sense of the principle of maximum entropy) rather than to a hypothesis on the exact shape of the plume. In particular, a visual inspection of the orbits reveals that many XCO_2 enhancements take the form of Gaussian mixtures that can still be well adjusted by a single Gaussian: these ones are retained by the automatic process.

The quality of the fit is evaluated by the following criterions: (a) its coefficient of determination r^2 (it has to be larger than 0.7^2), (b) the density of the retrievals within 1σ of the Gaussian centre, and between 2σ and 3σ of it (in both cases it has to be at least of 1 sounding per km on average on each side of the peak), (c) the value of the Gaussian standard deviation (it must be between 2 and 200/6 km, except for the 80 × 80 km^2 SAM observation mode of OCO-3, for which the maximum acceptable σ is reduced to 20 km), (d) the value of the Gaussian height relative to the retrieval variability in the window (the former has to be larger than the latter). A

validated fit is interpreted as reflecting the potential crossing of a CO_2 plume from an emitting area or point source. If several validated fitting functions, centred on different retrievals, overlap, only the one with the largest r^2 is kept. Such an overlap can happen for instance between estimates made from different scan lines of a SAM.

By definition, the area under the Gaussian is the XCO_2 line density (in ppm m^{-1}, then converted to kg m^{-1}), which reflects the distribution of CO_2 abundance in the CO_2 plume. We multiply it with the wind speed in the direction normal to the OCO track to retrieve the corresponding emission under the assumption of a steady wind. The wind speed is taken from the fifth generation of ECMWF atmospheric reanalyses of the global climate (ERA5, Hersbach et al., 2020) in the model ninth level from the surface, that corresponds to a geometric altitude of about 250 m (after Brunner et al., 2019). The few cases with retrieved emissions less than 1 $ktCO_2$ h^{-1} or larger than 100 $ktCO_2$ h^{-1} are judged as unreliable and are left out.

This set of emission retrievals is then filtered to keep only the fresh plumes, defined here as plumes which are mostly less than 1.5-hour old, because older ones have a more complex transport history. Only OCO-observed plume transects that are either downwind of, or less than 30 km from, coal or steel plants whose cumulative emissions are larger than 1.0 $ktCO_2$ h^{-1} are kept. The location of the coal-fired power plants or steel plants is taken from the detailed time series reported and regularly updated by https://globalenergymonitor.org/. The versions of July and March 2022, for coal and steel, respectively, are used. For coal, this data includes an estimate of the annual CO_2 emissions for each plant that we correct for the capacity factor: indeed we assume that when an OCO sees a coal-fired power plant plume, this plant operates at the maximum of its capacity. For steel plants, we assume an actual production of 75% of the plant nominal capacity and a conversion factor or 1.85 tons of CO_2 per ton of steel produced. **Figure 1** shows a simple case where an XCO_2 enhancement is clearly visible in the OCO-2 data downwind of a coal-fired power plant, using the ERA5 wind estimate.

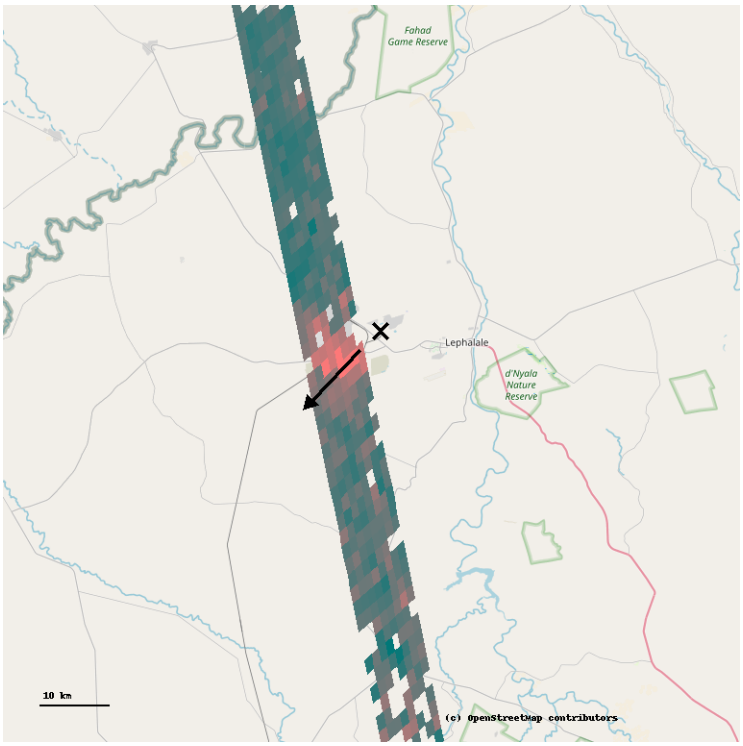

Figure 1 – plume identified on a portion of an OCO-2 orbit over South Africa centred on sounding 2014110711502571 (v10 retrievals, plotted over an OpenStreetmap background). The cross marks the Matimba coal power plant and the arrow shows the wind direction.

Note that we do not attempt to optimise the wind direction given the uncertainty in the origin of the plume, in particular when there are multiple emitters. Other sources of information would be needed to do it properly.

2.4 Definition of the footprint of the emission retrievals

The uncertainty in the wind field and the lack of information on the vertical distribution of the plume leaves a large margin for attribution of the plume to specific emitters, beyond the nearby ones that allow selecting the emission retrievals, as explained in Section 2.3. We define the plume footprint ad hoc as follows. First, we only consider coal-fired plants and steel plants from https://globalenergymonitor.org/ whose emission plumes could reach the orbit within less than 7 hr in a straight line, given the wind at the peak (as defined in Section 2.2). Older plumes are assumed to be too diffused to contribute to the observed enhancement (Broquet et al., 2018). Second, we exclude the plants whose emission would pass through the orbit in a straight line (given the wind at the peak) more than 5σ from the Gaussian centre. The loose 5σ criterion accounts for the uncertainty in the wind direction. For the same reason, we reintegrate emissions from the plants that are "downwind" of the orbit but less than 30 km from it. We keep the list of these plants together with the emission retrievals for evaluation, when we compare the latter with the cumulative emission of the former.

2.5 References

Baker, D. F., Bell, E., Davis, K. J., Campbell, J. F., Lin, B., and Dobler, J.: A new exponentially decaying error correlation model for assimilating OCO-2 column-average CO2 data using a length scale computed from airborne lidar measurements, Geosci. Model Dev., 15, 649–668, https://doi.org/10.5194/gmd-15-649-2022, 2022.

Broquet, G., Bréon, F.-M., Renault, E., Buchwitz, M., Reuter, M., Bovensmann, H., et al. (2018). The potential of satellite spectro-imagery for monitoring CO2 emissions from large cities. Atmospheric Measurement Techniques, 11, 681– 708. https://doi.org/10.5194/amt-11-681-2018

Brunner, D., Kuhlmann, G., Marshall, J., Clément, V., Fuhrer, O., Broquet, G., et al. (2019). Accounting for the vertical distribution of emissions in atmospheric CO2 simulations. Atmospheric Chemistry and Physics, 19, 4541– 4559. https://doi.org/10.5194/acp-19-4541-2019

Chevallier, F., Zheng, B., Broquet, G., Ciais, P., Liu, Z., Davis, S. J., et al. (2020). Local anomalies in the column-averaged dry air mole fractions of carbon dioxide across the globe during the first months of the coronavirus recession. Geophysical Research Letters, 47, e2020GL090244. https://doi.org/10.1029/2020GL090244

Chevallier, F., Broquet, G., Zheng, B., Ciais, P., & Eldering, A. (2022). Large CO_2 emitters as seen from satellite: Comparison to a gridded global emission inventory. *Geophysical Research Letters*, 49, e2021GL097540.https://doi.org/10.1029/2021GL097540

Eldering, A., Taylor, T. E., O'Dell, C. W., & Pavlick, R. (2019). The OCO-3 mission: Measurement objectives and expected performance based on 1 year of simulated data. *Atmospheric Measurement Techniques*, 12, 2341– 2370.https://doi.org/10.5194/amt-12-2341-2019

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 1, 51. https://doi.org/10.1002/qj.3803

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. R., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A. (2018), ImprovedRetrievals of Carbon Dioxidefrom the Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539-6576, doi:10.5194/amt-11-6539-2018.

Zheng, B., Chevallier, F., Ciais, P., Broquet, G., Wang, Y., Lian, J., & Zhao, Y. (2020). Observing carbon dioxide emissions over China's cities and industrial areas with the Orbiting Carbon Observatory-2. Atmospheric Chemistry and Physics, 20, 8501– 8510. https://doi.org/10.5194/acp-20-8501-2020

3 Estimates of fossil fuel emissions from regional inversions using co-emitted species

3.1 Overview

The sequence of computations implemented by Fortems-Cheiney and Broquet (2022) that is used for the estimate of the national-scale fossil fuel CO_2 ($FFCO_2$) emissions in Europe and for the generation of Figures 3 and 4 of the CoCO2 submission to the GST consists in:

1) The atmospheric inversions of maps of the NO_x or CO emissions over Europe during 2005-2021 at 1-day and 0.5° resolution. These inversions are based on the coupling between the variational mode of the Community Inversion Framework (CIF, Berchet et al., 2021), the CHIMERE regional atmospheric chemistry transport model (Menut et al., 2013, Mailler et al., 2017) and the adjoint code of this model (Fortems-Cheiney et al., 2021). The NO_x and CO inversions assimilate respectively atmospheric NO_2 and CO products from spaceborne instruments and apply corrections to the maps of anthropogenic emissions from an inventory of the European NO_x, CO and CO_2 emissions by the TNO organisation (Kuenen et al., 2014; Super et al., 2020) and to the maps of biogenic emissions of NO_x from the MEGAN model (Guenther et al., 2006),

2) The conversion of the daily maps of NO_x or CO anthropogenic emissions from these inversions into estimates of the fossil fuel CO_2 emissions at a national and monthly scale for five large groups of sectors of emitting activities. This conversion relies on the sectoral maps of emissions from the three species and, implicitly, on the emission ratios between the species for each sector, country and month from the inventory by TNO.

The following subsections detail the different components of this general scheme.

3.2 Satellite observations

3.2.1 OMI-QA4ECV-v1.1 NO_2

The Ozone Monitoring Instrument (OMI) is a ultraviolet-visible (UV-Vis) instrument launched in July 2004 onboard the Earth Observation System (EOS) Aura satellite, which flies on a 705 km sun-synchronous orbit that crosses the Equator at 13:40 LT. OMI provides measurements for the derivation of NO_2 Tropospheric Vertical Column Density (TVCD) observations. Compared with other UV-Vis instruments providing a long archive of NO_2 observations, it has the highest spatial resolution and least degradation (Schenkeveld et al., 2017; Levelt et al., 2018), allowing the study of long-term variations for various regions of the world. We use the OMI-QA4ECV-v1.1 NO_2 TVCD level-2 product (Boersma et al., 2017). The data selection follows the criteria of the data quality statement from Boersma et al. (2017): the processing error flag equals 0 for a pixel

- the solar zenith angle is lower than 80°,
- the snow ice flag is lower than 10 or equal to 255,
- the ratio of tropospheric air mass factor (AMF) over geometric AMF is higher than 0.2 to avoid situations in which the retrieval is based on very low (relative) tropospheric air mass factors,
- the cloud fraction is lower than 0.5.

3.2.2 TROPOMI-PAL NO_2

The Tropospheric Monitoring Instrument (TROPOMI, Veefkind et al., 2012) was launched onboard the Copernicus Sentinel-5 Precursor (S5P) satellite in October 2017. It flies on a 824 km altitude sun-synchronous orbit that crosses the Equator at 13:30 LT. This imaging spectrometer covers a UV-Vis band supporting the derivation of NO_2 TVCD observations. We use the Product Algorithm Laboratory (PAL) version 2.3 NO_2 TVCD level-2 product (Eskes et al., 2021).

Our TROPOMI data selection follows the criteria of VanGeffen et al (2018). We only select observations with a quality assurance (qa) value higher than 0.75 and a cloud radiance fraction lower than 0.5. As for OMI, we only select observations when the error associated with the retrieval is lower than 100%.

3.2.3 MOPITT CO

We use CO observations from the MOPITT instrument (Deeter et al., 2019). The MOPITT Infrared (IR) spectrometer has been flown onboard the NASA EOS-Terra satellite, on a low sun-synchronous orbit that crosses the Equator at 10:30 and 22:30 LST since 1999. The spatial resolution of its measurements is about 22 ×22 km^2 at nadir. It has been operated nearly continuously since March 2000. MOPITT CO products are available in three variants: thermal-infrared (TIR) only, near-infrared (NIR) only and the multispectral TIR-NIR product, all containing total columns and retrieved profiles (expressed on a 10-level grid from the surface to 100 hPa). We choose to use the MOPITT "surface" multispectral MOPITTv8-NIR-TIR product (also called MOPITT-v8J), as it provides the highest number of observations, with a good evaluation against in situ data from NOAA stations (Deeter et al., 2019).

3.3 Configuration of CHIMERE for the simulation of NO_2 and CO concentrations in Europe

CHIMERE is widely used to model air quality at regional scale (Cholakian et al., 2019; Ciarelli et al., 2019, Menut et al., 2020). It is one of the seven state-of-the-art chemistry transport models (CTMs) behind the operational ensemble simulations of the Copernicus Atmosphere Monitoring Service (CAMS) regional services.

Here, CHIMERE and its adjoint code are driven by the CIF to simulate NO_2 and CO atmospheric concentrations over Europe and their sensitivity to the surface emissions. It is run over a 0.5°x0.5° regular horizontal grid with 17 vertical layers, from the surface to 200 hPa, with 8 layers within the first two kilometers. The domain covers Europe (15.25°W-35.75°E; 31.75°N-74.25°N, see **Figure 2**) and therefore includes 101 (longitude) x 85 (latitude) x 17 (vertical levels) grid-cells.

The chemical scheme used in CHIMERE is MELCHIOR-2, with more than 100 reactions, including 24 for inorganic chemistry. CHIMERE is driven by the European Centre for Medium-Range Weather Forecasts (ECMWF) operational meteorological forecasts. The emissions maps and boundary conditions used for the simulations are discussed in the next sections.

Due to the need for a compromise between the robustness of the simulation of the chemistry in the model and the computational cost with a complex chemical scheme, the aerosols modules of CHIMERE have not been included in its adjoint code yet and are therefore not activated in the CHIMERE forward simulations.

3.4 Comparison between the satellite observations and CHIMERE

As the spatial resolution of the satellite data is finer than that of the CHIMERE grid, the selected OMI and TROPOMI NO_2 TVCDs, and MOPITT surface concentrations are aggregated into a single observation in each model grid-cell, called "super-observation". In order to associate the super-observations to a real averaging kernel, these super-observations are defined as the

observations (with their averaging kernel) corresponding to the median of the concentrations within the 0.5°×0.5° model grid-cell and within the CHIMERE physical time step of about 5-10 minutes.

The CHIMERE simulation of the NO_2 TVCDs and CO concentrations corresponding to these super-observations are derived from the NO_2 and CO simulated 3D-fields by applying vertical integrations with the averaging kernels associated to the super-observations (**Figure 2**).

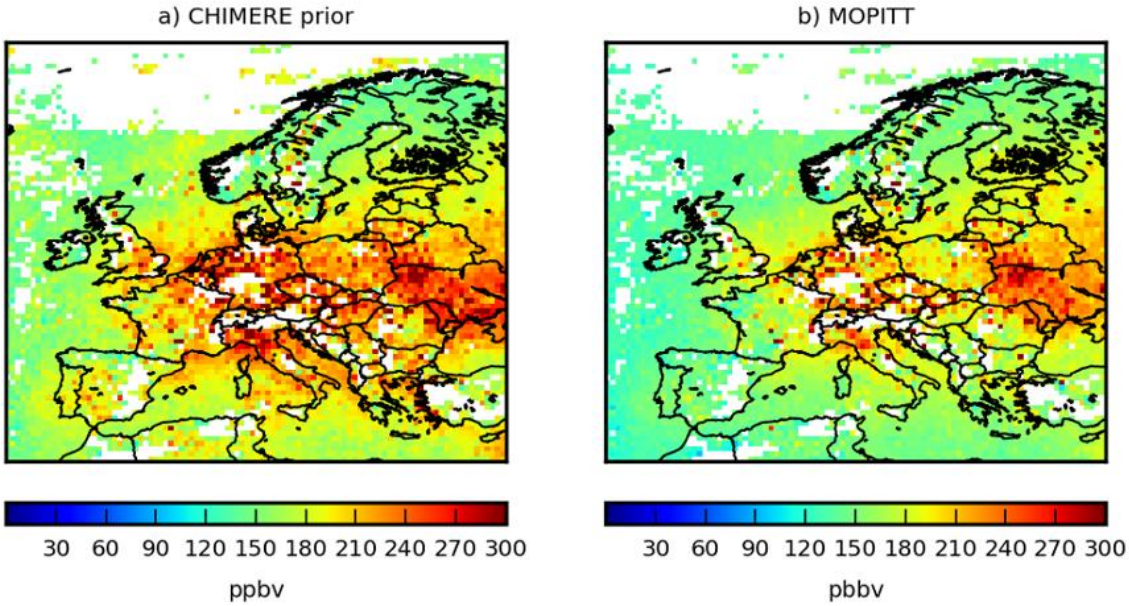

Figure 2 – Domain for the CHIMERE simulations and for the inversions with the averages of the CO concentrations a) simulated by CHIMERE using the prior TNO-GHGco-v3 anthropogenic emission estimate where and when MOPITT-v8J "surface" super observations are available, and b) observed by MOPITT-v8J, in ppbv, at the 0.5°x0.5° CHIMERE grid-cell resolution, in February 2015.

3.5 Prior estimate of the emissions, and of the boundary conditions

The Bayesian framework of the atmospheric inversion underlies the use of a prior estimate of the NOx and CO emissions, which is updated by the NO_x and CO emission inversions to better fit the observations. This prior estimate also feeds the initial simulation with CHIMERE in the iterative process of the variational inversions. The inversion system controls the initial and lateral and top boundary conditions for the concentrations in the model together with the surface emissions, and, as for the latter, it requires a prior estimate of these boundary conditions. Actually, here, the update of these prior estimates for the emissions and boundary conditions by the inversions consists in rescaling these prior estimates rather than in incremental corrections. Therefore, strictly speaking, the emission products detailed below belong to the observation operator of the inversions. Finally, the simulations of the atmospheric chemistry require estimates of the regional emissions of VOCs, which are not modified by the inversion.

The prior anthropogenic NO_x and CO emission estimates are taken from the TNO-GHGco-v3 gridded inventory. TNO-GHGco-v3 is an update of the TNO inventory described in Kuenen et al., (2014) and in Super et al., (2020). The dataset is based on the last European Monitoring and Evaluation Programme/Centre on Emission Inventories Projection (EMEP/CEIP) and UNFCCC official country reporting for air pollutants and greenhouse gases respectively. The emissions for the year 2019 have been extrapolated with an in-sample approach, and, here, the emissions for 2020 and 2021 are fixed to the values for 2019. In TNO-GHGco-v3 the NO_x, CO and $FFCO_2$ emissions are mapped at a 6x6 km^2 horizontal resolution. Annual and national budgets are disaggregated in space based on proxies of the different sectors (Kuenen et al., 2014). The temporal disaggregation is based on temporal profiles provided per GNFR sector code with typical month to month, weekday to week-end and diurnal variations. Following the

GENEMIS recommendations (Kurtenbach et al., 2001; Aumont et al., 2003), we speciated the TNO-GHGco-v3 NO_x emissions as 90% of NO, 9.2% of NO2, and 0.8% of nitrous acid (HONO) emissions. We have aggregated the TNO-GHGco-v3 maps at the 0.5°x0.5° horizontal resolution of the CHIMERE grid.

The prior biogenic NO_x emissions are derived from the MEGAN model (Guenther et al., 2006). CO biogenic emissions are assumed to be negligible and are not taken into account. The fixed anthropogenic emissions for VOCs are obtained from the EMEP inventory (EMEP/CEIP website). The fixed VOCs biogenic emission estimates come from the MEGAN model (Guenther et al., 2006).

Climatological values from the LMDZ-INCA global model (Szopa et al., 2008, Folberth et al., 2006) are used to prescribe concentrations at the lateral and top boundaries and the initial atmospheric composition in the domain of CHIMERE (Menut et al., 2013). The fields of the global model have a horizontal resolution of 3.8° in longitude and 2.5° in latitude. They are interpolated along the boundaries of our CHIMERE domain.

3.6 Variational inversion of the NO_x and CO emissions

The inversions of the NO_x or CO emissions follow a Bayesian theoretical framework. They consist in correcting the prior estimate of these emissions and of the model initial and boundary conditions to improve the fit between CHIMERE and the satellite NO_2 or CO data, respectively. Fortems-Cheiney et al. (2021) provides details on the principle and configuration for these variational inversions. The optimal (posterior) estimate of the emissions in a statistical sense is found by iteratively minimising the following cost function $J(\mathbf{x})$:

$$J(\mathbf{x})=(\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)+(H(\mathbf{x})-\mathbf{y})^T\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{y})$$

Where:

- $\mathbf{x}$is the control vector and $\mathbf{x}^b$ is its prior value, which are detailed in the following;
- the observation vector $\mathbf{y}$ gathers the super-observations from the OMI, TROPOMI or MOPITT datasets presented in sections 3.2 and 3.4;
- the observation operator H projecting the control vector to simulations of concentrations corresponding to these super-observation combines the emission and boundary conditions presented in section 3.5, the CHIMERE model presented in section 3.3 and the sampling and vertical interpolation corresponding to the observations as explained in section 3.4;
- the prior and observation error covariance matrices $\mathbf{B}$ and $\mathbf{R}$ characterise the uncertainties in $\mathbf{x}^b$ and in $\mathbf{y}$ and H respectively, under the assumption that these uncertainties are unbiased and follow a normal distribution.

Series of inversions over successive 1-day windows are conducted for NO_x and then combined to provide an estimate of the NO_x emissions over the whole period of analysis. The control vector $\mathbf{x}$ in the NOx inversions contains scaling factors to be optimised by the NO_x inversion:

- factors scaling the NO and NO_2 anthropogenic emissions at a 1-day temporal resolution, at a 0.5° × 0.5° (longitude, latitude) horizontal resolution and over the first 8 vertical levels of CHIMERE i.e. for each of the corresponding 101×85×8 grid cells,
- factors scaling the NO biogenic emissions at a 1-day temporal resolution, at a 0.5° × 0.5° (longitude, latitude) resolution and at the surface (over 1 vertical level only), i.e. for each of the corresponding 101×85×1 grid cells,
- factors scaling the daily NO and NO_2 3D initial conditions at 0:00 UTC, at a 0.5° × 0.5° (longitude, latitude) resolution and over the 17 vertical levels of CHIMERE.

For CO, series of inversions over successive 1-month windows are conducted. The control vector $\mathbf{x}$ for the CO inversions is similar to that for the NO_x inversions. However, due to the

relatively long lifetime of CO, we have to account for the uncertainties in the CO lateral boundary conditions at the borders of the CHIMERE domain. Furthermore, the chemical production of CO by VOCs is fixed and not controlled here by the inversion. Therefore, the control vector for the CO inversions contains the following scaling factors:

- factors scaling the CO anthropogenic emissions at a 1-day temporal resolution, at a 0.5° ×0.5° (longitude, latitude) horizontal resolution and over the first 8 vertical levels of CHIMERE, i.e. for each of the corresponding (30 or 31 days) x 101 × 85 × 8 grid cells,
- factors scaling the CO lateral boundary conditions at 1-day temporal resolution, at a 0.5° × 0.5° (longitude, latitude) horizontal resolution, i.e. for each of the corresponding (30 or 31 days) × 372 grid cells, over the 17 vertical levels of CHIMERE,
- factors scaling the CO 3D initial conditions at 0:00 UTC the first day of the month, at a 0.5° × 0.5° (longitude, latitude) horizontal resolution and over the 17 vertical levels of CHIMERE.

The prior estimate $\mathbf{x}^b$ of the control vectors for the inversions consist in sets of unitary scaling factors, which implicitly fixes the prior knowledge of the emissions and boundary conditions to the products detailed in section 3.5. The assumptions made to define the **B** and **R** matrices have been detailed in Fortems-Cheiney et al. (2021) and the main features for the parts of **B** corresponding to the surface emissions are recalled in Table 1.

Table 1 – Assumptions on the uncertainty in the NO_x and CO emission products detailed in section 3.5 that are used to derive the covariance matrix of the prior uncertainty B.

	Anthropogenic emissions	**Biogenic emissions**
NOx	50% at 0.5° × 0.5° and daily scale Spatial correlations of 50km No day to day temporal correlation	100% at 0.5° × 0.5° and daily scale Spatial correlations of 50km
CO	100% at 0.5° × 0.5° and daily scale Spatial correlations of 50km No day to day temporal correlation	

The inversions are conducted using the variational mode of the CIF with the M1QN3 algorithm for the minimization of the cost function *J* (Gilbert and Lemaréchal, 1989). At each iteration of this minimization, the CIF uses a CHIMERE simulation to compute *J* for a new estimate of **x** and the adjoint code of CHIMERE to compute for this new estimate of x (Fortems-Cheiney et al., 2021, Berchet et al., 2021).

Three inversions are conducted over multiple years:

- the inversion of NO_x emissions based on the OMI data over 2005-2020

- the inversion of NO_x emissions based on the TROPOMI data over 2019-2020

- the inversion of CO emissions based on the MOPITT data over 2011-2020

The posterior estimate of the anthropogenic emissions from these different NO_x or CO inversions are used for the subsequent analysis.

3.7 Derivation of the European $FFCO_2$ emissions

The derivations of the 1-month/national scale budgets of the European $FFCO_2$ emissions relies, first, on the conversion of the posterior estimate of the NO_x or CO total anthropogenic emissions at 0.5° resolution from the different inversions into estimates of sectoral NO_x or CO anthropogenic emissions at national scale. For this purpose, monthly averages of the maps of sectoral NO_x or CO emissions provided by the TNO-GHGco-v3 inventory restricted to each

European country are compared to the monthly averages of the maps of total anthropogenic emissions from the inversions at 0.5° resolution. The sectoral distribution is defined by five large groups of activity sectors: energy, industry, residential, road transport and the rest of the sectors, within which all emissions from the TNO-GHGco-v3 inventory with a more detailed sectoral distribution are aggregated. The spatial distribution at 0.5° resolution of the inventory for a given country, month and sector is assumed to be accurate and significantly different from that for the other sectors. Based on this assumption, the spatial patterns of the differences between the sectoral maps from the inventory and the maps from the inversion are interpreted in terms of differences between the sectoral budgets of emissions at the country and 1-month scale. Therefore, for each country and month, the five temporally average sectoral maps of NO_x or CO emissions from the TNO-GHGco-v3 inventory regridded at 0.5° resolution are rescaled (with a set of five scaling factors) so that their resulting sum fits the temporally average map of NO_x or CO total anthropogenic emission at 0.5° resolution from the inversion. The optimal fit is found by minimizing the root sum square misfits at 0.5° resolution between the sum of the rescaled sectoral maps from the TNO-GHGco-v3 inventory and the map from the inversion. In practice, a simple analytical inversion scheme is used to infer the optimal set of five scaling factors to be applied to the five average sectoral maps of NO_x or CO emissions from the TNO-GHGco-v3 inventory

Finally, for each country and month, these optimal scaling factors are applied to the sectoral budgets of the $FFCO_2$ emissions from the TNO-GHGco-v3 inventory. The resulting $FFCO_2$ emission budgets are the product which is documented in the CoCO2 submission to the GST. Implicitly, this last operation is equivalent to use the NO_x-to-$FFCO_2$ or CO-to-$FFCO_2$ anthropogenic emission ratios from the TNO-GHGco-v3 inventory for each country, month and sector to convert the sectorized NO_x or CO emission budgets into $FFCO_2$ emission budgets.

3.8 References

Aumont, B., Chervier, F., and Laval, S.: Contribution of HONO sources to the $NO_x/HO_x/O_3$ chemistry in the polluted boundary layer. Atmospheric Environment, 37(4):487 – 498, 2003. https://doi.org/10.1016/S1352-2310(02)00920-2

Berchet, A., Sollum, E., Thompson, R. L., Pison, I., Thanwerdas, J., Broquet, G., Chevallier, F., Aalto, T., Berchet, A., Bergamaschi, P., Brunner, D., Engelen, R., Fortems-Cheiney, A., Gerbig, C., Groot Zwaaftink, C. D., Haussaire, J.-M., Henne, S., Houweling, S., Karstens, U., Kutsch, W. L., Luijkx, I. T., Monteil, G., Palmer, P. I., van Peet, J. C. A., Peters, W., Peylin, P., Potier, E., Rödenbeck, C., Saunois, M., Scholze, M., Tsuruta, A., and Zhao, Y.: The Community Inversion Framework v1.0: a unified system for atmospheric inversion studies, Geosci. Model Dev., 14, 5331–5354, 2021. https://doi.org/10.5194/gmd-14-5331-2021

Boersma, K. F., Eskes, H., Richter, A., De Smedt, I., Lorente, A., Beirle, S., Van Geffen, J., Peters, E., Van Roozendael, M. and Wagner, T.: QA4ECV NO_2 tropospheric and stratospheric vertical column data from OMI (Version 1.1) [Data set], Royal Netherlands Meteorological Institute (KNMI), 2017. http://doi.org/10.21944/qa4ecv-no2-omi-v1.1

Cholakian, A., Beekmann, M., Coll, I., Ciarelli, G., and Colette, A.: Biogenic secondary organic aerosol sensitivity to organic aerosol simulation schemes in climate projections, Atmos. Chem. Phys., 19, 13209-13226, 2019. https://doi.org/10.5194/acp-19-13209-2019

Ciarelli, G., Theobald, M. R., Vivanco, M. G., Beekmann, M., Aas, W., Andersson, C., Bergstrom, R., Manders-Groot, A., Couvidat, F., Mircea, M., Tsyro, S., Fagerli, H., Mar, K., Raffort, V., Roustan, Y., Pay, M.-T., Schaap, M., Kranenburg, R., Adani, M., Briganti, G., Cappelletti, A., D'Isidoro, M., Cuvelier, C., Cholakian, A., Bessagnet, B., Wind, P., and Colette, A.: Trends of inorganic and organic aerosols and precursor gases in Europe: insights from the EURODELTA multi-model experiment over the 1990-2010 period, Geosci. Model Dev., 12, 4923-4954, 2019. https://doi.org/10.5194/gmd-12-4923-2019

Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Mao, D., Martínez-Alonso, S., Worden, H. M., Ziskin, D., and Andreae, M. O.: Radiance-based retrieval bias mitigation for the MOPITT instrument: the version 8 product, Atmos. Meas. Tech., 12, 4561–4580, 2019. https://doi.org/10.5194/amt-12-4561-2019

EMEP/CEIP, https://ceip.at/ms/ceip_home1/ceip_home/webdab_emepdatabase/emissions_emepmodels/

Eskes, H., van Geffen, G., Sneep, M., Veefkind, P., Niemeijer, S., &Zehner, C.: S5p nitrogen dioxide v02.03.01 intermediate reprocessing on the s5p-pal system: readme file (Tech. Rep.). Retrieved from https://data-portal.s5p-pal.com/product-docs/no2/, 2021.

Folberth, G. A., Hauglustaine, D. A., Lathière, J., and Brocheton, F.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: model description and impact analysis of biogenic hydrocarbons on tropospheric chemistry, Atmos. Chem. Phys., 6, 2273–2319, https://doi.org/10.5194/acp-6-2273-2006, 2006.

Fortems-Cheiney, A., Pison, I., Broquet, G., Dufour, G., Berchet, A., Potier, E., Coman, A., Siour, G., and Costantino, L.: Variational regional inverse modeling of reactive species emissions with PYVAR-CHIMERE-v2019, Geosci. Model Dev., 14, 2939–2957, 2021. https://doi.org/10.5194/gmd-14-2939-2021

Fortems-Cheiney, A., and Broquet, G.: Final re-analysis of the national scale CO2 anthropogenic emissions over 2055-2015, VERIFY public deliverable D2.12, https://verify.lsce.ipsl.fr/index.php/repository/public-deliverables/wp2-verification-methods-for-fossil-co2-emissions/d2-12-final-re-analysis-of-the-national-scale-co2-anthropogenic-emissions-over-2005-2015

Gilbert, J. C., and C. Lemaréchal: Some numerical experiments with variable-storage quasi-Newton algorithms, Math. Program., 45, 407–435, 1989. https://link.springer.com/article/10.1007/BF01589113

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, 2006. https://doi.org/10.5194/acp-6-3181-2006

Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC_II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, Atmos. Chem. Phys., 14, 10963-10976, 2014. https://doi.org/10.5194/acp-14-10963-2014

Kurtenbach, R., Becker, K., Gomes, J., Kleffmann, J., Lorzer, J., Spittler, M.,Wiesen, P., Ackermann, R., Geyer, A., and Platt, U.: Investigations of emissions and heterogeneousformation of HONO in a road traffic tunnel.Atmospheric Environment, 35(20):9506 – 9517. 3385Ð3394, 2001. https://doi.org/10.1016/S1352-2310(01)00138-8

Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., van der A, R., McLinden, C., Fioletov, V., Carn, S., de Laat, J., DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K., de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H., and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmos. Chem. Phys., 18, 5699–5745, 2018. https://doi.org/10.5194/acp-18-5699-2018

Mailler S., L. Menut, D. Khvorostyanov, M. Valari, F. Couvidat, G. Siour, S. Turquety, R. Briant, P. Tuccella, B. Bessagnet, A. Colette, L. Letinois, and F. Meleux, CHIMERE-2017: from urban

to hemispheric chemistry-transport modeling, Geosci. Model Dev., 10, 2397-2423, 2017. https://doi.org/10.5194/gmd-10-2397-2017

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G.,Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, Geosci. Model Dev., 6, 981–1028, 2013. https://doi.org/10.5194/gmd-6-981-2013

Menut, L., Bessanet, B., Siour, G., Mailler, S., Pennel, R. and Cholakian, A. : Impact of lockdown measures to combat Covid-19 on air quality over western Europe, Science of The Total Environment, 741, 2020. https://doi.org/10.1016/j.scitotenv.2020.140426

Super, I., Dellaert, S. N. C., Visschedijk, A. J. H., and Denier van der Gon, H. A. C.: Uncertainty analysis of a European high-resolution emission inventory of CO_2 and CO to support inverse modelling and network design, Atmos. Chem. Phys., 20, 1795–1816, 2020. https://doi.org/10.5194/acp-20-1795-2020

Szopa, S., Foret, G., Menut, L., and Cozic, A.: Impact of large scale circulation on European summer surface ozone: consequences for modeling, *Atmospheric Environment*, 43, 1189–1195,doi:10.1016/j.atmosenv.2008.10.039, 2008.

Veefkind, J., Aben, I., McMullan, K., Forster, H., de Vries, J., Otter, G., Levelt, P.: TROPOMI on the esa sentinel-5 precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote Sensing of Environment, 120, 70-83, 2012. https://doi.org/10.1016/j.rse.2011.09.027

4 Estimates from a Carbon Cycle Fossil Fuel DataAssimilation System

4.1 General presentation

The Carbon Cycle Fossil Fuel Data Assimilation System (CCFFDAS, Kaminski et al., 2022) pursues an innovative approach to the estimation of fossil fuel emissions in that it combines top-down (inverse modelling) and bottom-up (forward modelling) of sectoral fossil fuel emissions and of the terrestrial biosphere.

With respect to the user requirements identified by CoCO2[2], the CCFFDAS addresses by construction the user requirement of disentangling fossil fuel CO2 emissions and natural flux. Further, it is worth noting that the system can be operated in two modes. The *synergistic* mode includes (IEA or other) sectoral national emission totals as an observation used in the model parameter calibration. Alternatively, CCFFDAS can be operated in *verification* mode, i.e., without using the IEA sectoral national emission totals. This latter mode provides emission estimates that are largely independent of data used for the bottom-up estimations of emission, and thus address another user requirement.

In the following we describe the setup that has produced Figure 5 of D6.5 *Data contribution of the European CoCO2 project to the first Global StockTake*.

[2]https://www.coco2-project.eu/sites/default/files/2022-03/CoCO2-D6-3-V1-0.pdf

4.2 Processing steps

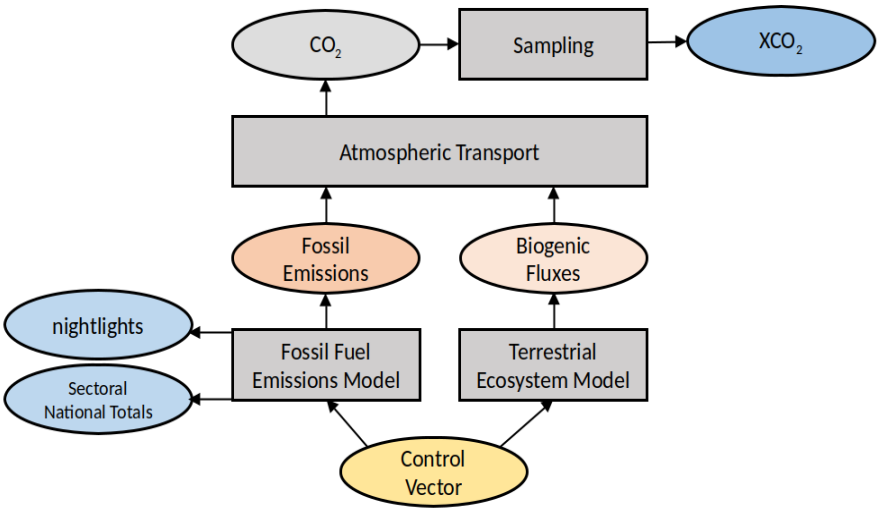

Figure 3 – Modelling framework with forward flow of information in CCFFDAS. Boxes represent calculation steps by models, blue ovals observables, and the yellow oval the control vector (model parameters and initial condition).

The CCFFDAS consists of a series of numerical, process-based models simulating the global atmospheric transport of CO_2(Heimann and Koerner, 2003), the sectoral emissions from fossil fuel usage and the exchange fluxes (photosynthetic uptake and respiratory release) from the terrestrial biosphere (Kaminski et al., 2017). The flow of information in the forward sense through the modelling chain is shown in**Figure 3**. The CCFFDAS version used here separates fossil fuel emissions into two sectors, an electricity generation sector and the complement, which is termed other sector (see Asefi-Najafabadi et al., 2014; Kaminski et al., 2022). The electricity generation is simulated per power plant using a prior emission data set per power plant (see section 4.3). Emissions from the other sector are simulated by the Kaya identity (Nakicenovic, 2004): the emissions F from the other sector in a 0.1 degree times 0.1 degree gridcell x in country c are given by

$$F(x,c)=pP(x)g(x)ef(c).$$

The variables on the right-hand side are a scalar multiplier, p, for the population density, P(x), a point-wise estimate of the per capita gross domestic product, g(x), a global constant for the energy intensity of the economy, e, and a country-wise estimate for the carbon intensity of energy production, f(c). Input data sets are listed in section 4.3. The output of these models, i.e. the CO_2 fossil fuel emissions (from both the electricity generation and the other sector) and the CO_2 land-atmosphere exchange fluxes, depend on a set of model parameters in the equations (the control vector) used for calculating the fluxes. In the Kaya identity these are denoted by lower case letters and for the electricity generation sector these are the emissions from each listed power station directly. The control vector components for the terrestrial ecosystem model are biome specific parameters controling the photosynthesis and heterotrophic respiration. The CCFFDAS operates in two steps:

1. Calibration against observations: The system is used to calibrate these parameters by assimilating the observational data streams listed in section 4.3. This calibration is based on a Bayesian formalism and uses a variational approach, i.e. a cost function consisting of two terms is minimised iteratively. The first term quantifies the misfit of simulated

counterparts to the observational data streams and the second term the deviation to prior estimates of the parameter vector. The minimisation algorithm relies on code for evaluation of the gradient of the cost function with respect to the parameters. This gradient is provided efficiently by the adjoint of the modelling chain. This adjoint is derived by the automatic differentiation tool TAPENADE (Hascoët and Pascual, 2003).

2. Target Simulation: The calibrated model is used to simulate relevant target quantities, i.e. sectoral fossil and biofuel emission fields on a 0.1 degree global grid that are consistent with the observations and the modelling chain.

The estimated sectoral emissions on the 0.1 degree grid resolution can then be aggregated to various territorial units such as national or regional totals (e.g. the NUTS levels in Europe).

4.3 Input datasets

The input data sets are differentiated into observational data sets that are used for calibration, a data set that provides part of the prior information and additional data sets used by the model.

The observational data sets for calibration are:

- 2018 sectoral national emission totals from the International Energy Agency (IEA)[3]
- Night-light intensities observed by satellites (Elvidge et al., 2021)
- Atmospheric CO_2 concentrations from measurements provided by the Greenhouse Gases Observing Satellites 1 and 2 and from the Orbiting Carbon Observatory 2 (Reuter et al., 2013)

The prior data set consists of:

- Locations of and emissions from individual power plants for the year 2018 derived within CoCO2[4]

Further data sets used by the modelling chain include:

- Fraction of Absorbed Photosynthetically Active Radiation by plants derived by the Joint Research Centre-Two-stream Inversion Package from satellites (Pinty et al., 2011)and made available by the Joint Research Centre of the European Commission via (https://fapar.jrc.ec.europa.eu).This product is selected for its outstanding performance in a series of validation exercises and cross ECV assessments, see, e.g., Mota et al. (2021) and references therein.
- A map of population density (Sims et al., 2022)
- Meteorological data from the fifth generation of ECMWF atmospheric reanalyses of the global climate (Hersbach et al.,2020)

[3] see https://www.iea.org/data-and-statistics/data-product/greenhouse-gas-emissions-from-energy
[4] see Hugo Denier van der Gon and CoCO2 WP2 team; D2.1 Prior Emission data 2018 documentation report, available athttps://coco2-project.eu/node/327

4.4 References

Asefi-Najafabady, S., Rayner, P. J., Gurney, K. R., McRobert, A., Song, Y., Coltin, K., Huang, J., Elvidge, C., and Baugh, K. (2014), A multiyear, global gridded fossil fuel CO_2 emission data product: Evaluation and analysis of results, *J. Geophys. Res. Atmos.*, 119, 10,213– 10,231, doi:10.1002/2013JD021296.

Elvidge CD, Zhizhin M, Ghosh T, Hsu F-C, Taneja J. Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019. *Remote Sensing*. 2021; 13(5):922. https://doi.org/10.3390/rs13050922

Hascoët, L.L., Pascual, V.V., 2013. The Tapenade Automatic Differentiation tool: principles, model, and specification. ACM Trans. Math. Softw. 39 (3).

Hersbach, H, Bell, B, Berrisford, P, et al. The ERA5 global reanalysis. *Q J R Meteorol Soc*. 2020; 146: 1999– 2049. https://doi.org/10.1002/qj.3803

Kaminski, T., Scholze, M., Voßbeck, M., Knorr, W., Buchwitz, M., and Reuter, M.: Constraining a terrestrial biosphere model with remotely sensed atmospheric carbon dioxide, Remote Sensing of Environment, pp. 109 – 124, https://doi.org/http://dx.doi.org/10.1016/j.rse.2017.08.017, URL http://www.sciencedirect.com/science/article/pii/S0034425717303838, 2017.

Kaminski, T., Scholze, M., Rayner, P., Voßbeck, M., Buchwitz, M., Reuter, M., Knorr, W., Chen, H., Agustí-Panareda, A., Löscher, A., and Meijer, Y.: Assimilation of atmospheric CO2 observations from space can support national CO2 emission inventories, Environmental Research Letters, 17, 014 015, https://doi.org/10.1088/1748-9326/ac3cea, URL https://doi.org/10.1088/1748-9326/ac3cea, 2022.

Mota B., Gobron N., Morgan O., Cappucci F., Lanconelli C., and Robustelli M. Cross-ECV consistency at global scale: LAI and FAPAR changes. Remote Sensing of Environment, 2021, 263. DOI: 10.1016/j.rse.2021.112561

Nakicenovic, N.: Socioeconomic driving forces of emissions scenarios, in: The global carbon cycle: Integrating humans, climate, and the natural world, edited by Field, C. B. and Raupach, M. R., pp. 225–239, Island Press, Washington, D.C., 2004.

Reuter, M., Bösch, H., Bovensmann, H., Bril, A., Buchwitz, M., Butz, A., Burrows, J. P., O'Dell, C. W., Guerlet, S., Hasekamp, O., Heymann, J., Kikuchi, N., Oshchepkov, S., Parker, R., Pfeifer, S., Schneising, O., Yokota, T., and Yoshida, Y.: A joint effort to deliver satellite retrieved atmospheric CO_2 concentrations for surface flux inversions: the ensemble median algorithm EMMA, Atmos. Chem. Phys., 13, 1771–1780, https://doi.org/10.5194/acp-13-1771-2013, 2013.

Pinty, B., Clerici, M., Andredakis, I., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., and Widlowski, J.-L.: Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 2. Fractions of transmitted and absorbed fluxes in the vegetation and soil layers, Journal of Geophysical Research – Atmospheres, 116, D09 106, https://doi.org/10.1029/2010JD015373, 2011.

Sims, K., Reith, A., Bright, E., McKee, J., & Rose, A. (2022). *LandScan Global 2021* [Data set]. Oak Ridge National Laboratory. https://doi.org/10.48690/1527702

5 Estimates of AFOLU emissions from CO_2 atmospheric inversions

5.1 Introduction

Figure 6 of the CoCO2 submission to the GST is an update of Figures 1 and 2 of Chevallier (2021). In the following we update its description. Computations were performed on the High Performance Computing resources of https://www-hpc.cea.fr/ under the allocation A0110102201 made by Grand Equipement National de CalculIntensif (GENCI).

5.2 Inversion Data

Twice per year since 2011, CAMS and its three precursor projects Monitoring Atmospheric Composition and Climate (MACC, https://www.ecmwf.int/en/research/projects/macc-ii) have been producing multi-decade analyses of CO_2 surface atmospheric measurements over the globe in the form of optimised gridded surface fluxes with associated 3D concentration fields (Chevallier et al., 2010, with successive updates described in https://atmosphere.copernicus.eu/supplementary-services). In 2019, the operational chain has been paralleled with four-monthly "Fast Track" experimental analyses of XCO_2 retrieved from OCO-2 (Chevallier et al., 2019, with successive updates described in https://atmosphere.copernicus.eu/supplementary-services). At the time of writing, the latest versions of these two products have been: (a) v21r1, that assimilated surface air-sample measurements of the CO_2 dry air mole fraction made at 146 sites over the globe, from January 1979 to May 2022, and (b) FT21r2 that assimilated NASA's ACOS bias-corrected land XCO_2 retrievals, version 10 from September 2014 until February 2020 (and another instrument before September 2014, but the corresponding data are of lower quality and we do not use it here). "FT" stands for Fast Track, since the satellite observations are available faster than most surface measurements, but the surface-driven inversion remains the CAMS reference product due to slightly lesser agreement with independent atmospheric measurements, as reported in the product documentation. The CAMS data are available from https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-greenhouse-gas-inversion.

The inversion system follows Bayesian principles formulated within a variational framework. The Bayesian approach compensates for the irreversibility of atmospheric transport and for the limited observational coverage by the use of prior information on surface fluxes provided by inventories (for fossil fuel use, cement production or biomass burning), a model (for terrestrial ecosystems) and the interpolation of measurements of surface ocean CO_2 fugacity combined with a simple gas transfer velocity model. For the interpretation of the following results, it is important to note that the annual average of the prior fluxes from terrestrial ecosystems have no interannual variability: All interannual variations in the optimised (or "posterior") vegetation and soil fluxes are therefore driven by the assimilated observations, modulated by the interannual variations in atmospheric transport represented in the underlying observation operator. The inversion does not solve for the emissions from fossil fuel use and cement production, the prior values of which come from the UNFCCC National Inventory Reports (NIRs) as much as possible (Jones et al., 2021). The variational framework makes it possible to estimate the other CO_2 surface fluxes at the weekly grid-point scale, with day and night separation, in a unique assimilation window, thus ensuring physical and statistical consistency of each inversion throughout the years and decades. The size of the grid points corresponds to that of the underlying global atmospheric transport model, currently 3.75° in longitude and 1.875° in latitude. The reader is referred to the above articles and technical reports for a detailed description of these products and of the processing chain that generated them.

5.3 Aggregation in space and time

The relatively high resolution of the inversion (73,700 variables controlled per month) compared to the density of the assimilated atmospheric data, at least for the surface-air-sample-driven inversion, is motivated by the risk of degradation of the inversion product by the aggregation

error studied by Kaminski et al. (2001) and Bocquet et al. (2011), when running at coarse resolution. Despite the availability of high-resolution outputs, the inversion product is typically post-processed to be presented at annual subcontinent scales (Chevallier et al., 2010) or even in broad latitudinal bands (e.g., Friedlingstein et al., 2022). Here, we examine both CAMS CO_2 inversion products v21r1 and FT21r2 at country-scale resolution. In most cases, the country-scale resolution is an intermediate between the usual subcontinent resolution and the grid point resolution. In time, we choose the usual annual scale of the national reporting to the UNFCCC.

We use the 0.08° × 0.08° country mask of Klein Goldewijk et al. (2017) to determine, for each UNFCCC party, the fraction that it occupies in each land portion of the 3.75° × 1.875° model grid points. We multiply this fraction by the fraction of non-wild area deduced from the 1 km^2 global terrestrial human footprint maps of Venter et al. (2016, 2018a, 2018b) that synthesise diverse information on human population pressure (population density), human land use and infrastructure (built-up areas, night-time lights, land use/land cover), and human access (coastlines, roads, railroads, navigable rivers) into a generic Global Human Influence Index with values between 0 and 50. In the absence of a clear delineation between managed and unmanaged plots in the UNFCCC reports, all index values larger than zero (i.e., all values that indicate some human influence) are considered to correspond to managed areas here. The maps are available for years 1993 and 2009: We linearly interpolate the fractions in-between those years and keep the 1993 and 2009 values for earlier and later years, respectively. This fractional coverage allows us to aggregate the grid-point land vegetation flux values (i.e., the total posterior fluxes minus the prior emissions from fossil fuel use and cement production), while conserving the mass of carbon fluxes at all spatial scales. Lastly, weekly diurnal and nocturnal fluxes have been aggregated on an annual scale.

5.4 Uncertainty statistics

The uncertainty of the posterior fluxes provided by the inversion system is directly deduced from the Bayesian framework which combines the uncertainty in the prior fluxes, the uncertainty of the observations as defined from the inversion observation operator (consisting mainly of an atmospheric transport model and, for the assimilation of satellite retrievals, of a vertical averaging kernel) and the observation operator itself, to estimate it. The statistical models of the CAMS inversion system represent error variances and correlations at various spatiotemporal scales. They have been the subject of dedicated studies at different scales which give us confidence in their broad realism (Chevallier et al., 2006, 2012; Chevallier and O'Dell, 2013). In practice, we use a robust Monte Carlo method, made of ensembles of two-year inversions (much easier to parallelize than the inversions made on the full inversion periods), to explore the uncertainty space in a manner consistent with the assigned prior and observation error statistics (Chevallier et al., 2007). Each member of the inversion ensembles was aggregated as the reference data to produce the uncertainty statistics at the annual scale and at the appropriate spatial scale. Note that the ensemble of inversions over two years barely describes the interannual correlation of the uncertainty. These correlations would prevent the uncertainty of multi-annual averages from decreasing with the square of the number of years.

5.5 Inventory data

We have used the latest AFOLU data reported under the UNFCCC at the time of writing and available on the UNFCCC portal (https://unfccc.int/ghg-inventories-annex-i-parties/2022). They cover the period between 1990 and 2020, at least partially, depending on the party. Their territorial scope differs from the data reported within the framework of the Kyoto Protocol and available on the same portal with the same format: Our land mask for each party accounts for these subtle differences that involve the status of overseas territories, although they play an insignificant role in the following results. For Annex-I parties (Canada, EU27+UK, Kazakhstan, Russia and USA), we simply use the sum of the net CO_2 fluxes from the agriculture sector and the land-use, land-use change and forestry (LULUCF) sector reported in their common reporting format tables (CRF) submitted in 2022. In contrast to our uniform definition of managed versus unmanaged grid points for the inversion data described above, parties to the UNFCCC mainly

determine which areas are managed themselves, as long as they apply their definition consistently over time and report it in a transparent manner (see Chapter 3 of IPCC, 2006).

For Brazil, we use Appendix I of its fourth biennial update report (BUR) that gives emission values for years 1994, 2000, 2010, 2012, 2015, 2015, but note that no values are reported for the agriculture sector. Methodological details were taken from Brazil's Third National Communication.

For China, we use Tables 2−7, 2−13, 2−14, 2−15 and 2−16 of its second BUR that give emission values for years 1994, 2005, 2010, 2012, and 2014.

DR Congo has not submitted any BUR to UNFCCC yet and we rely on its third national communication to UNFCCC that provides CO_2 net emissions values for the LULUCF sector for the years between 2000 and 2010, but none for agriculture.

For India, we use Tables 2.2 and 2.19 of the initial and second national communications to UNFCCC, respectively, and Tables 2, 2.2 and 2.4, respectively, of the three successive BURs.

For Mongolia, we use Tables 2−4 of its first BUR. That table combines all reported greenhouse gases for each year between 1990 and 2014, but, consistent with the remark in Section of that BUR, its LULUCF values are very close to those for the AFOLU numbers for the sole CO_2 fluxes given in the annex for years 1990 and 2014. We therefore use the full time series of Table 4. Note that Mongolia only reported source categories forest land remaining forest land and harvested wood products for the LULUCF sector.

The various documents or the Annex I NIRs report an uncertainty analysis of the numbers. The two-sigma uncertainty is usually of a few tens of percent for the fluxes from the LULUCF sector (32% for Brazil in 2016, ~40% for Canada, 21% for China, ~20% for EU27+UK, 36% for Forest land Remaining Forest land in Mongolia in 2014, ~20% for the USA, no estimate for DR Congo and India for that CO2 sector). NIRs of Kazakhstan and Russia are in Russian and have not been exploited in order to avoid any misinterpretation.

For a like-for-like comparison, some emission and removal terms not included in the NIRs need to be subtracted from the inversion estimates and the terrestrial biospheric carbon input to the inland water network needs to be added to the inversion system. These correction terms are estimated from the scientific literature and from living databases like the one maintained by the FAO (Ciais et al., 2022, Deng et al., 2022).

5.6 Party Selection

Our rationale for limiting our analysis to the 10 parties named earlier is as follows: We limit the analysis to the parties whose geographical area is much larger than the inversion grid resolution and where the inversion provides a significant improvement to the prior fluxes. The first criterion is represented by a threshold on the surface areas: Parties smaller than 1 million square kilometres (i.e., about 15 times the grid point size) are excluded. The second criterion is represented by a maximum value of two thirds imposed on the ratio of posterior over prior uncertainty standard deviations for both v21r1 and FT21r2 for the party annual flux in 2015.

This selection leads to the following 10 parties: Brazil, Canada, China, DR Congo, EU27+UK (called “EUA” in the submission to UNFCCC), India, Kazakhstan, Mongolia, Russia and USA.

5.7 References

Bocquet, M., Wu, L., & Chevallier, F. (2011). Bayesian design of control space for optimal assimilation of observations. I: Consistent multiscale formalism. Quarterly Journal of the Royal Meterological Society, 137, 1340– 1356. https://doi.org/10.1002/qj.837

Chevallier, F., Bréon, F.-M., &Rayner, P. J. (2007). Contribution of the orbiting carbon observatory to the estimation of CO2 sources and sinks: Theoretical study in a variational data

assimilation framework. Journal of Geophysical Research, 112, D09307. https://doi.org/10.1029/2006JD007375

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., et al. (2010). CO2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements. Journal of Geophysical Research, 115, D21307. https://doi.org/10.1029/2010JD013887

Chevallier, F., & O'Dell, C. W. (2013). Error statistics of Bayesian CO2 flux inversion schemes as seen from GOSAT. Geophysical Research Letters, 40, 1252– 1256. https://doi.org/10.1002/grl.50228

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., &Cozic, A. (2019). Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions. Atmospheric Chemistry and Physics, 19, 14233– 14251. https://doi.org/10.5194/acp-19-14233-2019

Chevallier, F., Viovy, N., Reichstein, M., &Ciais, P. (2006). On the assignment of prior errors in Bayesian inversions of CO2 surface fluxes. Geophysical Research Letters, 33, L13802. https://doi.org/10.1029/2006GL026496

Chevallier, F., Wang, T., Ciais, P., Maignan, F., Bocquet, M., AltafArain, M., et al. (2012). What eddy-covariance flux measurements tell us about prior errors in CO2-flux inversion schemes. Global Biogeochemical Cycles, 26, GB1021. https://doi.org/10.1029/2010GB003974

Chevallier, F. (2021). Fluxes of carbon dioxide from managed ecosystems estimated by national inventories compared to atmospheric inverse modeling. Geophysical Research Letters, 48, e2021GL093565. https://doi.org/10.1029/2021GL093565

Ciais, P., Bastos, A., Chevallier, F., Lauerwald, R., Poulter, B., Canadell, P., Hugelius, G., Jackson, R. B., Jain, A., Jones, M., Kondo, M., Luijkx, I. T., Patra, P. K., Peters, W., Pongratz, J., Petrescu, A. M. R., Piao, S., Qiu, C., Von Randow, C., Regnier, P., Saunois, M., Scholes, R., Shvidenko, A., Tian, H., Yang, H., Wang, X., and Zheng, B.: Definitions and methods to estimate regional land carbon fluxes for the second phase of the REgional Carbon Cycle Assessment and Processes Project (RECCAP-2), Geosci. Model Dev., 15, 1289–1316, https://doi.org/10.5194/gmd-15-1289-2022, 2022.

Deng, Z., Ciais, P., Tzompa-Sosa, Z. A., Saunois, M., Qiu, C., Tan, C., Sun, T., Ke, P., Cui, Y., Tanaka, K., Lin, X., Thompson, R. L., Tian, H., Yao, Y., Huang, Y., Lauerwald, R., Jain, A. K., Xu, X., Bastos, A., Sitch, S., Palmer, P. I., Lauvaux, T., d'Aspremont, A., Giron, C., Benoit, A., Poulter, B., Chang, J., Petrescu, A. M. R., Davis, S. J., Liu, Z., Grassi, G., Albergel, C., Tubiello, F. N., Perugini, L., Peters, W., and Chevallier, F.: Comparing national greenhouse gas budgets reported in UNFCCC inventories against atmospheric inversions, Earth Syst. Sci. Data, 14, 1639–1675, https://doi.org/10.5194/essd-14-1639-2022, 2022.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., McGrath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pocock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., WillstrandWranne, A., Wright, R., Yuan, W., Yue, C., Yue, X.,

Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2022, Earth Syst. Sci. Data, 14, 4811–4900, https://doi.org/10.5194/essd-14-4811-2022, 2022.

IPCC (2006). In H. S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe (Eds.), In IPCC guidelines for national greenhouse gas inventories, prepared by the national greenhouse gas inventories programme. IGES.

Jones, M. W., Andrew, R. M., Peters, G. P., Janssens-Maenhout, G., De-Gol, A. J., Ciais, P., et al. (2021). Gridded fossil CO2 emissions and related O2 combustion consistent with national inventories 1959–2018. Science Data, 8, 2. https://doi.org/10.1038/s41597-020-00779-6

Kaminski, T., Rayner, P. J., Heimann, M., &Enting, I. G. (2001). On aggregation errors in atmospheric transport inversions. Journal of Geophysical Research, 106, 4703– 4715. https://doi.org/10.1029/2000jd900581

Klein Goldewijk, K., Beusen, A., Doelman, J., &Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene – HYDE 3.2. Earth System Science Data, 9, 927– 953. https://doi.org/10.5194/essd-9-927-2017

Venter, O., Sanderson, E., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., et al. (2016). Global terrestrial human footprint maps for 1993 and 2009. *Science Data*, 3, 160067.https://doi.org/10.1038/sdata.2016.67

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., et al. (2018a). Last of the wild project, Version 3 (LWP-3): 2009 human footprint, 2018 release. NASA Socioeconomic Data and Applications Center (SEDAC).https://doi.org/10.7927/H46T0JQ4

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., et al. (2018b). Last of the wild project, Version 3 (LWP-3): 1993 human footprint, 2018 release. NASA Socioeconomic Data and Applications Center (SEDAC).https://doi.org/10.7927/H4H9938Z

6 Estimates of CH_4 emissions from the extended Copernicus Atmosphere Monitoring Service global monitoring system

6.1 Forward model

The ECMWF global Integrated Forecasting System (IFS), which provides the operational Copernicus Atmosphere Monitoring Service (CAMS, https://atmosphere.copernicus. eu/, last access: 15 January 2022) greenhouse gas (GHG) forecast (Agustí-Panareda et al., 2019), was used to generate the forward model integrations used in this study. These were performed from January to June of 2019 and 2020, with additional case study simulations performed for June 2018, November 2019 and July to September of 2020. Computational cost prevented simulating the full period (2018–2020). Simulations were performed using a horizontal cubic octahedral reduced Gaussian grid (TCo399: ∼ 25 km) and 137 vertical levels with coupled meteorology at operational forecast timesteps of 15 min and 3-hourly output. Monthly gridded prior estimates of anthropogenic emissions were taken from the CAMS global emissions product, CAMS-GLOB-ANT v4.2, (Granier et al., 2019), which combines existing products. The Global Fire Assimilation System (GFAS) provided daily biomass burning emissions (Kaiser et al., 2012). We used a monthly climatology of wetland emissions based on the LPJ-WHyMe model (Spahni et al., 2011). Remaining fluxes from oceans (Lambert and Schmidt, 1993; Houweling et al., 1999), termites (Sanderson, 1996) and wild animals (Houweling et al., 1999) were used at the highest available spatiotemporal resolution. The atmospheric CH4 sink comprised of a monthly mean climatological loss rate field (Bergamaschi et al., 2009), which represents loss reactions with hydroxyl, chlorine and atomic oxygen radicals. A gridded surface soil sink was also used (Ridgwell et al., 1999). Initial conditions for the 3D atmospheric state of CH4 were taken from the CAMS CH4 inversion product (Segers and Houweling, 2018).

6.2 4D-Variational inversion

We used the 4D-Var IFS system, cycle 47R1, employed operationally at ECMWF between June 2020 and May 2021. More detailed information on the IFS 4D-Var system can be found in Rabier et al. (2000). The incremental algorithm used consists of solving a series of quadratic minimisation problems (inner-loop) constructed by linearising the initial (non-linear) cost function around updated estimates of the state vector (outer-loop). To constrain surface emissions, the state vector is augmented by a parameter control vector that consists of a 2D scaling factor applied to a prior emission inventory. In our configuration, the posterior scaling factors are optimised on a regular 2D grid (~ 80 km) within a 24-h window and then applied to the prior emission inventory defined on a grid of ~ 10-km resolution (**Figure 4**). Prior emission errors are assumed to be independent between 24-h inversion cycles (i.e. each 24-h inversion uses the same uniform scaling factor of 1 and the same prior errors). This choice was driven by the lack of information about temporal error correlations in current prior inventories. Currently, the error covariance for the CH4 initial state vector is taken from a climatology and fixed in time. As a result, posterior errors in methane emissions and 3D state are not propagated forward across data assimilation cycles in this configuration, which is a technical limitation of our current system and will be addressed in subsequent versions. We use an online 4D-Var data assimilation system, where the meteorological fields are part of the control vector and optimised jointly with the emission scaling factors. As a result, the transport errors associated with uncertainties in the initial conditions of the meteorological variables are accounted for in our inversion. This is in contrast with widely used offline inversion systems, wherein transport errors are typically prescribed on an ad-hoc basis and fixed. Note that in our experiments the background errors for the meteorological variables at the initial time are constructed based on a climatology, and therefore are not flow-dependent.

The scaling factors derived from the inversion were applied to sector-specific prior maps for source attribution. A caveat to this approach is the assumption that co-located sectors have the same scaling factor applied, which can only be overcome with the use of co-emitted species observations such as ethane or isotopologues. However, this is unlikely to noticeably impact these results as at the relatively high increment resolution used (~ 80 km) CH4 sectors are rarely co-located. Missing sources in the prior are also not accounted for when using a posterior scaling factor.

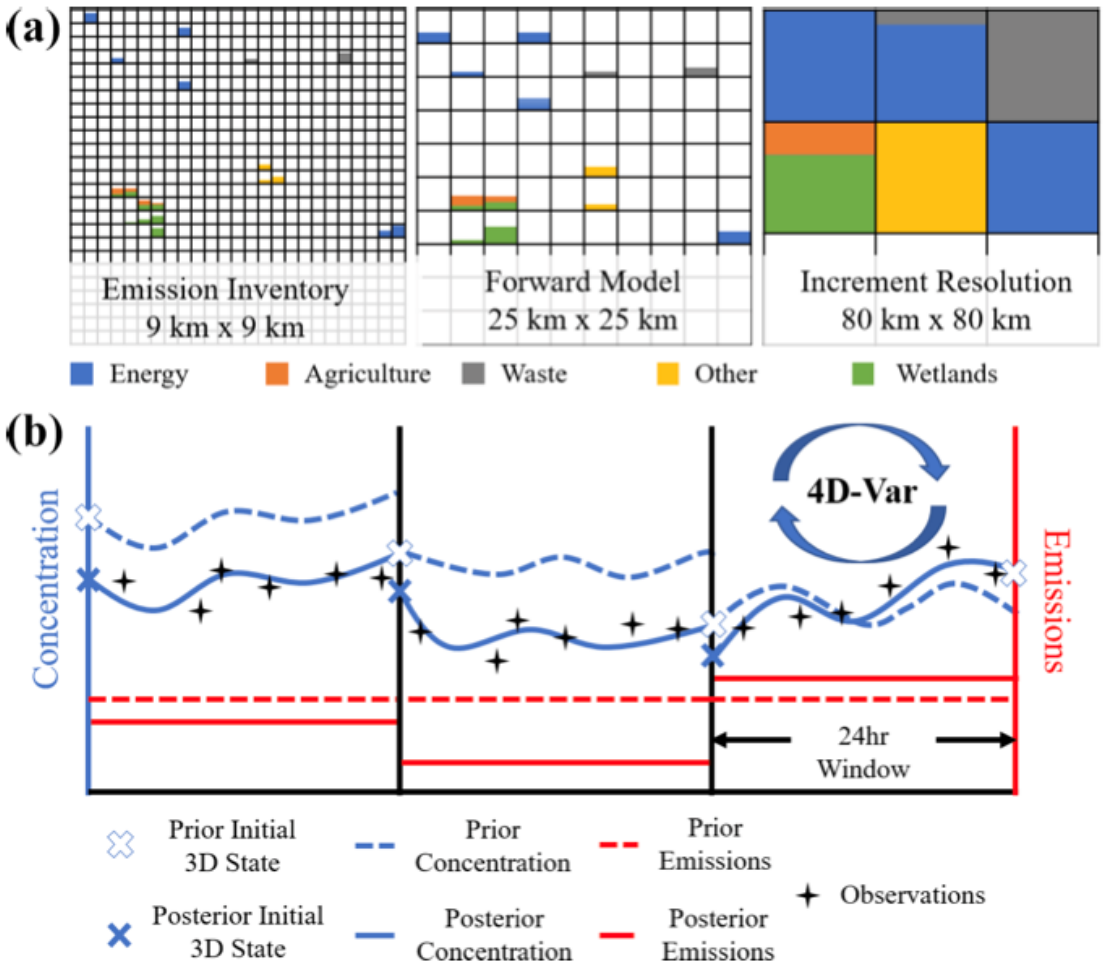

Figure 4 (a) Schematic of different resolutions used in the inversion shown by pseudo-data for five sectors. The magnitude of prior emissions at ~ 9 km (left panel) and those same emissions used as input to the forward model at ~ 25 km (middle panel). The inversion increment at ~ 80 km, resulting scaling factors are applied to all sectors within the grid cell, the boxes indicate relative contribution per sector (right panel). (b) Schematic of inversion setup using the 24-h window, correcting for the initial 3D state, emissions, and initial conditions in the prior of the subsequent window.

6.3 Prior information

Anthropogenic sector-specific grid cell uncertainties, taken from Maasakkers et al. (2016), provided the initial prior estimate for countries with well-developed statistical infrastructures or Annex I countries (IPCC, 2006). For Non-Annex I countries, the same sector-specific uncertainties were further increased by 50 %. Globally, constant wetland uncertainties were estimated at 58 %, taken as the standard deviation from the WetCHARTs ensemble (Bloom et al., 2017). We assume the standard deviation of the WetCHARTs ensemble to provide a reasonable uncertainty estimate of the LPJ-WHyMe emissions used here. Initially, all other biogenic uncertainties were estimated as 100 %. The atmospheric sink was not optimised by the inversion. Sensitivity experiments where prior errors were perturbed and validated against independent observations were used to evaluate prior uncertainty assumptions. Given that anthropogenic emissions are typically from point sources (e.g. fossil fuel extraction), we assumed no spatial prior error correlation given that the derived increments are at ~ 80 km. Wetland emissions would typically require defined spatial correlations; however, given the uncertainty of these structures, the focus of this study being anthropogenic emissions and limited occurrences of co-located emissions from wetland and anthropogenic sources we have chosen to omit these for simplicity. Total grid cell uncertainties, used in the control vector, were calculated using the error propagation method. All prior uncertainties are assumed to have a log-normal distribution to prevent negative emissions.

6.4 Observations

The observations used in the meteorological component of the IFS 4D-Var system include satellite radiances, conventional ground-based and radiosondes, and aircrafts and ships data, for which the coverage and quality is constantly monitored prior the assimilation. With specific focus on CH4, the TROPOMI instrument on-board the Sentinel-5 Precursor satellite provides near-global daily coverage of XCH4 with a nadir ground pixel size of 7 km × 7 km and near-surface sensitivity (Veefkind et al., 2012; Lorente et al., 2021). We used operational observations, which became available in April 2018 and were bias corrected, as in Barré et al. (2021). TROPOMI uncertainties (< 1 %) provided as part of the CH4 product were applied within the minimisation routine and averaging kernels were used (Hasekamp et al., 2019). Additional XCH4 observations from IASI and GOSAT, and their associated uncertainties of ~ 2 % and < 1 %, respectively, are assimilated into the system to provide additional constraints as described by Massart et al. (2014). Poor quality data are removed based on the provided quality flags.

6.5 References

Agustí-Panareda, A., Diamantakis, M., Massart, S., Chevallier, F., Muñoz-Sabater, J., Barré, J., Curcoll, R., Engelen, R., Langerock, B., Law, R. M., Loh, Z., Morguí, J. A., Parrington, M., Peuch, V.-H., Ramonet, M., Roehl, C., Vermeulen, A. T., Warneke, T., and Wunch, D.: Modelling CO2 weather – why horizontal resolution matters, Atmos. Chem. Phys., 19, 7347–7376, https://doi.org/10.5194/acp-19-7347-2019, 2019.

Barré, J., Aben, I., Agustí-Panareda, A., Balsamo, G., Bousserez, N., Dueben, P., Engelen, R., Inness, A., Lorente, A., McNorton, J., Peuch, V.-H., Radnoti, G., and Ribas, R.: Systematic detection of local CH4 anomalies by combining satellite measurements with high-resolution forecasts, Atmos. Chem. Phys., 21, 5117– 5136, https://doi.org/10.5194/acp-21-5117-2021, 2021.

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals, J. Geophys. Res., 114, D222301, https://doi.org/10.1029/2009jd012287, 2009.

Bloom, A. A., Bowman, K. W., Lee, M., Turner, A. J., Schroeder, R., Worden, J. R., Weidner, R., McDonald, K. C., and Ja- cob, D. J.: A global wetland methane emissions and uncertainty

dataset for atmospheric chemical transport models (WetCHARTs version 1.0), Geosci. Model Dev., 10, 2141–2156, https://doi.org/10.5194/gmd-10-2141-2017, 2017.

Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Liousse, C., Quack, B., Simpson, D., and Sindelarova, K.: The Copernicus Atmosphere Monitoring Service global and regional emissions (April 2019 version), Copernicus Atmosphere Monitoring Service, https://doi.org/10.24380/D0BN-KX16, 2019.

Hasekamp, O., Lorente, A., Hu, H., Butz, A., aan de Brugh, J., and Landgraf, J.: Algorithm Theoretical Baseline Document for Sentinel-5 Precursor Methane re- trieval, SRON-S5P-LEV2-RP-001, SRON, the Netherlands, https://sentinel.esa.int/documents/247904/2476257/ Sentinel-5P-TROPOMI-ATBD-Methane-retrieval (last access: 10 January 2021), 2019.

Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J., and Heimann, M.: Inverse modeling of methane sources and sinks using the adjoint of a global transport model, J. Geophys. Res., 104, 26137–26160, https://doi.org/10.1029/1999jd900428, 1999.H

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., 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, 527–554, https://doi.org/10.5194/bg-9-527-2012, 2012.

Lambert, G. and Schmidt, S.: Reevaluation of the oceanic flux of methane: Uncertainties and long term variations, Chemosphere, 26, 579–589, https://doi.org/10.1016/0045-6535(93)90443-9, 1993.

Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., aan de Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D., Pol- lard, D., Shiomi, K., Deutscher, N., Velazco, V., Roehl, C., Wennberg, P., Warneke, T., and Landgraf, J.: Methane retrieved from TROPOMI: improvement of the data product and valida- tion of the first 2 years of measurements, Atmos. Meas. Tech., 14, 665–684, https://doi.org/10.5194/amt-14-665-2021, 2021.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J.-X., Zhang, Y., Hersher, M., Bloom, A. A., Bowman, K. W., Worden, J. R., Janssens-Maenhout, G., and Parker, R. J.: Global distribution of methane emissions, emission trends, and OH concentrations and trends inferred from an inversion of GOSAT satellite data for 2010–2015, Atmos. Chem. Phys., 19, 7859–7881, https://doi.org/10.5194/acp-19-7859-2019, 2019.

Massart, S., Agusti-Panareda, A., Aben, I., Butz, A., Chevallier, F., Crevoisier, C., Engelen, R., Frankenberg, C., and Hasekamp, O.: Assimilation of atmospheric methane products into the MACC- II system: from SCIAMACHY to TANSO and IASI, Atmos. Chem. Phys., 14, 6139–6158, https://doi.org/10.5194/acp-14- 6139-2014, 2014.

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics, Q. J. Roy. Meteorol. Soc., 126, 1143–1170, https://doi.org/10.1002/qj.49712656415, 2000.

Ridgwell, A. J., Marshall, S. J., and Gregson, K.: Consumption of atmospheric methane by soils: A process-based model, Global Biogeochem. Cy., 13, 59-70, https://doi.org/10.1029/1998gb900004, 1999.

Segers, A. and Houweling, S.: Description of the CH4 Inversion Production Chain, CAMS (Copernicus Atmospheric Monitoring Service) Report, https://atmosphere.copernicus.eu/sites/default/ files/2018-11/CAMS73_2015SC3_D73.2.5.5-2018_201811_ production_chain_v1_0.pdf (last access: 8 November 2021), 2018.

Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., Frankenberg, C., Foster, P. N., Joos, F., Prentice, I. C., and van Velthoven, P.: Constraining global methane emissions and uptake by ecosystems, Biogeosciences, 8, 1643–1665, https://doi.org/10.5194/bg-8-1643-2011, 2011.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ., 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

Document History

Version	Author(s)	Date	Changes
1.0	All authors	25/01/2023	Initial version
1.1	All authors	15/02/2023	Address reviews

Internal Review History

Internal Reviewers	Date	Comments
Konstantinos Politakos	30/01/2023	Minor changes
Stavros Stagakis	02/02/2023	Minor changes