CoCO2

Prototype system for a Copernicus CO_2 service

Intercomparison of national-scale inversions

Julia Marshall and Friedemann Reum (DLR), Gregoire Broquet, Elise Potier, Antoine Berchet and Audrey Fortems-Cheiney (CEA), Erik Koene and Dominik Brunner (EMPA), Sander Houweling and Eleftherios Ioannidis (VUA), Arjo Segers (TNO), Andrea Kaiser-Weiss, Anusha Sunkisala, Valentin Bruch and Thomas Rösch (DWD), Michał Gałkowski and Piotr Sekuła (AGH), Guillaume Monteil and Hans Chen (LUND), Paul Palmer and Tia Scarpelli (UEDIN), Anteneh Getachew Mengistu and Tuula Aalto (FMI)

Co-ordinated by ECMWF

CoCO2

Prototype system for a Copernicus CO_2 service

D4.6 INTER-COMPARISON OF NATIONAL-SCALE INVERSIONS

Dissemination Level: Public

Author(s):
Julia Marshall and Friedemann Reum (DLR), Gregoire Broquet, Elise Potier, Antoine Berchet and Audrey Fortems-Cheiney (CEA; EP now at LISA, AFC now at Science Partners), Erik Koene and Dominik Brunner (EMPA), Sander Houweling and Eleftherios Ioannidis (VUA), Arjo Segers (TNO), Andrea Kaiser-Weiss, Anusha Sunkisala, Valentin Bruch and Thomas Rösch (DWD), Michał Gałkowski and Piotr Sekuła (AGH), Guillaume Monteil and Hans Chen (LUND), Paul Palmer and Tia Scarpelli (UEDIN), Anteneh Getachew Mengistu and Tuula Aalto (FMI)

Date: 15/12/2023
Version: 1
Contractual Delivery Date: 31/10/2023
Work Package/ Task: WP4/ T4.4
Document Owner: DLR
Contributors: CEA, EMPA, VUA, TNO, DWD, AGH, LUND, UEDIN, FMI
Status: Final


ECMWF
Empa
Materials Science and Technology
ICOS
cea
Max Planck Institute
for Biogeochemistry
TNO innovation
for life
Lund University
VU
Vrije Universiteit Amsterdam
Wageningen
University & Research
BSC
Barcelona Supercomputing Center
CICERO
cmcc
cnrs
Deutscher Wetterdienst
Wetter und Klima aus einer Hand
FMI
FORTH
iLab
Mercator Ocean International
Météo France
FCiências.ID
The Cyprus Institute
DLR
Deutsches Zentrum
für Luft- und Raumfahrt

CoCO2: Prototype system for a Copernicus CO_2 service

Coordination and Support Action (CSA)
H2020-IBA-SPACE-CHE2-2019 Copernicus evolution – Research activities in support of a European operational monitoring support capacity for fossil CO_2 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

Figures

Tables

1 Executive Summary

This deliverable documents the development and application of 11 national-scale CO_2- and/or CH_4-focused inversion systems targeting the emissions of European countries and the USA, carried out within Task 4.4 of the CoCO2 project. The aim of this inverse modelling effort was to document the current strategies and capabilities to separately control the anthropogenic emissions and biospheric fluxes of CO_2 (and, to a lesser extent, CH_4) with relatively high-resolution inversion systems. It also aimed to provide support and guidance for the design of national operational systems and the regional-scale setup of the multi-scale inversion system for the CO2MVS. The work in this task led to improvements of current inversion systems and to the development of new regional inversion capabilities. In particular, the focus on the separate control of the CO_2 anthropogenic emissions was a major innovation, which is important to advance the potential of the regional inverse modelling of CO_2 in Europe.

The participating systems varied in terms of maturity, with some models having been applied to CO_2 inverse modelling for the first time. The inversions operated at different resolutions (between about 0.5° and 10 km) and covered different regions and countries, with the general ambition of having at least two models covering the same country.

The results demonstrate that major challenges are associated with the lack of maturity of some of the most recent inverse modelling components: the separate control of the CO_2 anthropogenic emissions, the co-assimilation of co-emitted species, the co-assimilation of surface and satellite CO_2 observations, and the underlying characterization of the fine-scale uncertainties in the inventories of the anthropogenic emissions of CO_2 and co-emitted species used as prior estimates for the inversions. The set of inversions reveal a lack of constraint on the CO_2 anthropogenic emissions at the monthly to annual and national scale when using the existing in-situ and satellite observations. The co-assimilation of CO and NO_2 data does not significantly increase this constraint. In parallel, there is a large spread in the estimates of biogenic CO_2 fluxes (and of the anthropogenic and natural CH_4 fluxes) across the different systems, or when assimilating surface versus satellite observations. The analysis of the results suggests that the use of spatial resolutions finer than 10 km for the transport modelling and control of the fluxes is advisable if targeting CO_2 anthropogenic emissions. Furthermore, a systematic analysis of the inverse modelling components that are responsible for the spread of the results across the inversions is recommended.

2 Introduction

2.1 Background

The parties of the Paris Agreement are individual countries, which need to report their emission reduction ambitions and their actual emissions on a regular basis. Supporting the countries with top-down estimates at the national scale will therefore be a critical component of the future CO2MVS.

Task 4.4 of WP4 supported the development, optimization and testing of national-scale inversion systems complementing the global system of ECMWF. The analysis with these systems should also provide insights for the configuration of the CO2MVS multi-scale inversion system, which will have to tackle the corresponding scales. The advantage of these regional systems is that they can run at higher resolution and incorporate more detailed, country-specific information. The main goal was thus to evaluate the potential of limited-area models to monitor national to regional (in the administrative sense) budgets of anthropogenic CO_2 and/or CH_4 emissions.

A total of 11 modelling systems were set up to study the emissions of single countries or groups of countries within the limited domain, with the ambition of having at least two systems covering the same country. Whenever possible, the simulations followed a common protocol specifying the years to be covered (2018 and 2021, and if focusing on 1-month inversions: February and July as reference months), the input data to be used (anthropogenic and natural fluxes, lateral boundary conditions) and the output to be generated.

The systems differed in terms of underlying transport model, resolution, coverage, maturity (some systems were developed during the project), the type of observations assimilated (in-situ and/or satellite), and also in the type of inversion approach. Some systems only focused on CO_2, others on CH_4, and some included both. The wide range of systems and approaches provides a comprehensive overview of the state-of-the-art in regional inverse emission modelling of CO_2 (and CH_4) in Europe.

In addition to evaluating the potential of the regional systems, the following goals were targeted by a subset of models:

- analyse the sensitivity of inversions to the type of observation network and the usage of satellite observations in addition to measurements from ground-based networks
- analyse the use of fossil-fuel-like co-emitted species like NO_x and CO to better differentiate between anthropogenic and biogenic CO_2 fluxes
- investigate the usage of future CO2M observations in synthetic experiments
- contribute to the global stocktake (GST) exercise in WP6

The national-scale inversion systems will be able to contribute to the future CO2MVS either by providing independent estimates for selected countries or by integrating the inversion strategies or results into the global system of ECMWF. A method for integrating local and national estimates into the global system in the form of the assimilation of ensembles of inversion results was developed in WP6 (see Deliverable 4.7).

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

The goal of this deliverable is to document the different regional model systems and the results obtained with a specific focus on how well national scale anthropogenic fluxes can be estimated. The deliverable summarises the current state-of-the-art, identifies strengths and limitations, and provides recommendations towards the future operational system.

2.2.2 Work performed in this deliverable

Eleven regional inversion systems were set up to perform inverse model simulations with a primary focus on six countries in Europe (Germany, France, Netherlands, Poland, Finland and UK, even though several systems cover the whole Europe) and for the US following a common modelling protocol. For several modelling systems, substantial further developments were required before they could be applied to the tasks outlined in the protocol.

2.2.3 Deviations and counter measures

One planned set of experiments was to perform OSSEs with synthetically generated CO2M observations to assess the potential of this future mission in comparison to the existing observation networks. In the course of the project, we planned to use the pseudo CO2M data to be generated in the frame of Task T5.3. However, since these pseudodata could not be provided early enough for the modelling teams to be incorporated, we focused on the analysis of inversions with real observations from the current observation network and real satellite

observations from OCO-2 in 2018 and, to a lesser extent, in 2021. Furthermore, one OSSE was performed with *in situ* observations in the USA.

Another set of experiments was planned to assess the potential of using country-specific information, such as more detailed emission inventories or additional observations not included in operational networks. These experiments could not be performed as planned for two reasons: the first reason was that running the standard setup following the protocol was already challenging, using up a substantial amount of the available resources. The second reason was that the standard simulations indicated a rather poor constraint of anthropogenic CO_2 emissions, a conclusion that would not have changed by switching to more detailed prior inventory. As a countermeasure, some additional tests of the sensitivity of satellite versus surface observations and of the sensitivity to the prior estimates of anthropogenic emissions and biospheric CO_2 fluxes were performed.

Due to unexpected departures of postdocs, difficulties in replacing them and difficulties in hiring new personnel, some of the planned work was substantially delayed. This particularly affected the work of VUA, AGH, UEDIN, DLR and CEA, but in the end, all these groups contributed to this deliverable, though with some delays, or not to the extent that would have been possible otherwise. One of the inversion contributions initially planned in this task was in the end limited to model development and a demonstration of its capabilities with a test case instead of a setup following the inversion protocol (the CTDAS-WRF inversions by DLR). A bug in the new developments of that system affected the results of the VUA contribution, as it was discovered too late to rerun. On the other hand, several models were able to run larger domains, thereby covering more countries than originally planned.

When starting the CoCO2 project, there were some expectations that the first inversions in T4.4 could feed the CoCO2 report for the first GST in 2023. However, the first robust national scale inversion results came too late for this, for the reasons detailed above. Therefore, as an alternative, the extension of the European scale derivation of fossil fuel CO_2 emissions based on the NO_x and CO CIF-CHIMERE inversions from the H2020 VERIFY project and extended in WP6 of CoCO2 (see Section 3.2) were documented in the CoCO2 report for the GST, and a section detailing the plans in WP4 was added to this report (see the CoCO2 Deliverable 6.5).

3 The National-scale Inversions

In the following sections, the 11 contributing modelling systems are described, and their results are presented in the following order:

- the **CIF-CHIMERE CO_2 inversions** by CEA (including analysis of the potential of assimilating co-emitted CO and NO2 satellite data), covering **France**, and assimilating in situ and/or OCO-2 satellite data (see Section 3.2)
- the **ICON-ART-CTDAS CO_2 inversions** by Empa, covering **Northwest Europe**, and assimilating in situ and/or OCO-2 satellite data (see Section 3.3)
- the **ICON-ART-CTDAS CH_4 inversions** by Empa, covering **all of Europe**, and assimilating in situ data (see Section 3.3.3)
- the **ICON-ART-DWD CH_4 inversions** by DWD, covering **Germany**, and assimilating in situ data (see Section 3.4)
- the **WRF analytical CO_2 inversions** by AGH, covering **Germany and Poland**, and assimilating in situ data (see Section 3.5)
- the **LUMIA CO_2 inversions** by ULUND, covering **all of Europe**, and assimilating in situ data (see Section 3.6)
- the **TRACE CO_2 inversion OSSEs** by ULUND, covering **the contiguous USA**, and assimilating pseudo in situ data (see Section 3.7)

- the **LOTOS-EUROS CH_4 inversions** by TNO, covering **all of Europe**, and assimilating in situ data (see Section 3.8)
- the **GEOS-Chem CO_2 inversions** by UEDIN (co-assimilating CO_2 and CO data), covering **all of Europe**, and assimilating in situ or OCO-2 satellite data (see Section 3.9)
- the **CIF-FLEXPART CH_4 inversions** by FMI, covering **all of Europe**, and assimilating in situ data (see Section 3.10)
- the **WRF-CTDAS model development** by DLR and a test case assimilating in situ data (see Section 3.11) and preliminary **inversions** by VUA, covering **all of Europe** (with a zoom over North West Europe), and assimilating in situ data (see Section 3.11.4)

In an effort to make the contributions more comparable, a modelling protocol for the CO_2 inversions was developed, based largely upon the prior emissions dataset (PED) developed in WP2 of the CoCO2 project. To streamline the following descriptions, this protocol is briefly described here, such that only deviations from this need to be specified.

3.1 Modelling protocol

A comprehensive modelling protocol inspired by the TRANSCOM/IG3IS protocol for European CH_4 inversions was developed for the CO_2 inversions and shared with all modelling teams. For CH_4 inversions, the TRANSCOM/IG3IS protocol was to be followed.

The main elements of the CO_2 modelling protocol can briefly be summarised as follows:

- Target years: 2018 and 2021
- Domain: Left to the individual modelling groups, as different groups were targeting different countries. A high-resolution country mask was provided to ensure that national aggregation is taking place over common areas.
- Emission inputs (see Table 1)
- In-situ CO_2 observations: Primarily ICOS stations, additional stations in Poland UK, no strict list, but modellers are required to document exact list of stations. Use time-filtered observations 12:00 to 16:00 LT, except mountain sites (00:00-06:00 LT).
- Satellite CO_2 observations: OCO-2 ACOS v10r.
- Satellite observations of co-emitted species in case of join assimilation: Current operational version of NO_2 or CO from TROPOMI, apply quality filters as recommended in Product User Guide.
- Boundary conditions: CAMS inversion optimised fields. Although this data set is rather coarse in resolution, the biases are small. These were accessed from the CAMS Atmospheric Data Store (https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-greenhouse-gas-inversion?tab=form)
- Meteorological driving data: Use ERA5 whenever possible.
- Uncertainties: No strict rule, but all assumed uncertainties and covariance structures must be fully documented.
- State vector: No strict rule, but anthropogenic and biospheric fluxes must be separately estimated.
- Output: Gridded fluxes and uncertainties; national totals and uncertainties; mixing ratios and uncertainties.
- Documentation: Mandatory list of information points provided to modellers, e.g. name of contact person, name and version of inversion system, etc.

Table 1: CO_2 emissions and time factors to be used.

Category	Product	Resolution	Source	Mandatory
Anthropogenic CO_2	TNO	0.01° x 0.05°, annual with time factors, over Europe	PED	Yes
Anthropogenic CO_2	EDGAR	0.1°, annual with time factors	PED	No
Biogenic CO_2 (GPP & Re)	VPRM	1-km, hourly, over Europe	PED	Yes
Biogenic CO_2 (GPP & Re)	FLUXCOM	0.5°, hourly, global (for 2018: will be higher for 2021)	PED	No
Ocean fluxes	NEMO-PISCES	0.25 degree	PED	Yes
Biomass burning	GFAS	0.1°, daily, with emission heights	CDS	Yes
Lateral fluxes	from Frederic Chevallier	Gridded maps of displaced carbon at 8 km, monthly resolution.	FC	Yes*

*For the lateral fluxes, the recommendation was to use the source term as a (fixed/flexible) prior, but not the sink, so as to avoid double-counting. This should account for e.g. harvest from agricultural regions which is later respired elsewhere (e.g. in cities).

3.2 CIF-CHIMERE for CO_2 inversions from *in situ* and satellite observations (CEA)

This section describes the CIF-CHIMERE CO_2 inversion configuration for France, which was developed in the framework of task T4.4 of the CoCO2 project. We present the reference inversions for the year 2018 and additional one-month sensitivity tests.

This section also provides some insight on the current potential of the co-assimilation of species co-emitted with CO_2 during fossil fuel combustion, based on (i) NO_x and CO inversion results at the scale of France from other projects (the French ANR Argonaut and ADEME-AQACIA Lock'Air projects and Robin Plauchu's PhD thesis at LSCE) using a CIF-CHIMERE inversion configuration compatible with the one used for CO_2 inversions, and(ii) the conversion of results from European scale NO_x and CO inversions using a similar CIF-CHIMERE inversion configuration into CO_2 emission estimates, initiated in the H2020 VERIFY project and extended in the CoCO2 and ESA-World Emission projects, and whose implementation and results obtained in 2022 are documented in Deliverables 6.4, 6.5, 6.6 and in the CoCO2 GST leaflet of CoCO2 WP6.

Three reference inversions have been conducted to estimate the CO_2 anthropogenic and terrestrial ecosystem fluxes in France over the year 2018, assimilating the CO_2 observations (i) from the ICOS network and associated surface stations, (ii) from the satellite instrument OCO-2, and (iii) from both. The additional one-month tests of sensitivity have been focused on assessing the impact of the choice of the prior estimates of the anthropogenic and biogenic fluxes.

3.2.1 Model description

The system relies on the coupling between the variational mode of the CIF (Berchet et al., 2021), the regional chemistry transport model CHIMERE (Menut et al., 2013) and the adjoint of this model (Fortems-Cheiney et al., 2021).

3.2.1.1 Transport model

The configuration of CHIMERE (and of its adjoint code) for France covers the domain: 11°W-12°E; 39,5°N-54,5°N (cf Figure 1). Its zoomed grid has a 10 km horizontal resolution over France, and a 50 km horizontal resolution in the corners of this domain (cf Figure 1). It has 20 vertical layers, from the surface to 200hPa. The modelling of the CO_2 concentrations above 200 hPa for comparisons to CO_2 total column concentrations (XCO_2) from satellite instruments relies on the global product used for the model's CO_2 initial and boundary conditions. This CHIMERE configuration is driven by the ECMWF / IFS operational meteorological forecasts.

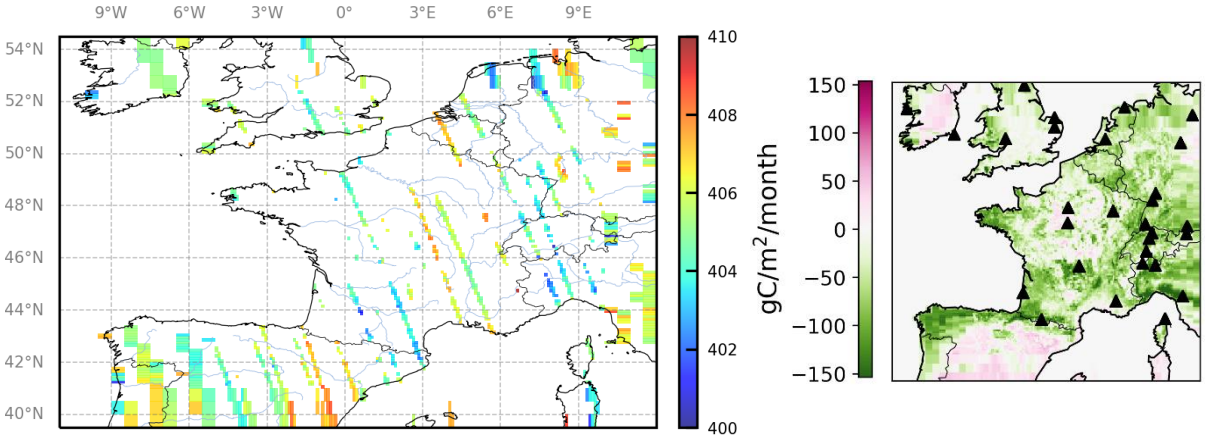


Figure 1: Domain of the CIF-CHIMERE inversions and 1) (left): binning into the CHIMERE zoomed grid of the OCO-2 v11 observations during the month of July 2018 2) (right): ground-based CO_2 measurement stations used for the inversions and prior estimate of the CO_2 NEE from VPRM interpolated on the CHIMERE zoomed grid during the month of July 2018.

The end of Section 3.2.2 documents some results of NO_x and CO inversions using the same CHIMERE configuration except that for those inversions, the MELCHIOR-2 chemical scheme (with more than 100 reactions; Menut et al., 2013) and its adjoint code are activated for the CHIMERE forward and adjoint simulations, since NO_x and CO are active species. The end of the Section 3.2.2 also recalls some results from Deliverables 6.4, 6.5 and 6.6 in WP6 based NO_x and CO inversions with a configuration of CHIMERE at 0.5° resolution covering Europe (see the details of this configuration in D6.4).

3.2.1.2 Prior fluxes and assimilated observations

Various CO_2 inversion experiments have been conducted with the CIF-CHIMERE configuration using different products for the prior estimates of the surface fluxes, and assimilating different observation datasets. This section details the set-up of these different experiments. This section also briefly indicates the main components of the NO_x and CO inversions mentioned in the result section. All the prior fluxes and the global product used to impose the boundary conditions are interpolated on the grid of the CHIMERE configuration for France and at one-hour resolution.

Prior fluxes for the reference CO_2 configuration

The reference configuration for the reference inversions relies on the prior estimate of the anthropogenic (fossil fuel and biofuel) emissions and terrestrial ecosystem fluxes from the TNO inventory (TNO_GHGco_6x6km_v4_0_year2018, at ~6-km resolution, with typical injection heights for the different sectors of activity) and the Vegetation Photosynthesis Respiration Model (VPRM) simulations (at 1 km resolution) delivered in the frame of WP2 (Denier van Der Gon et al., 2022) and listed as standard products in the Task T4.4 modelling protocol (Section 3.1).

In all experiments, the prior estimate of the initial and boundary conditions is derived from the CAMS global CO_2 inversions v20r2 (assimilating surface data) at 2.5° (longitude) x 1.27° (latitude) resolution. As explained above, this global inversion product is also used to

complement the vertical columns of CO_2 above the top boundary of CHIMERE when comparing the model to XCO_2 observations. However, unlike CHIMERE's initial and boundary conditions fields, this complement to the CO_2 partial columns from CHIMERE is not controlled by the inversion: it is not considered as a significant source of discrepancy between the observed and simulated XCO_2.

The prior estimate of sea/ocean fluxes within the CHIMERE domain is based on a hybrid product from the H2020 VERIFY project at 0.125° resolution combining the coastal ocean flux estimates from the University of Bergen and a global ocean estimate from MPI-BGC-Jena (Rödenbeck et al., 2014, McGrath et al., 2023). This product differs from the standard product of the Task T4.4 modelling protocol.

Two of the CO_2 flux components (for which standard products were proposed in the Task T4.4 modelling protocol) are ignored in both CHIMERE simulations and CIF-CHIMERE inversions: the biomass burning fluxes (assumed to be relatively small in France), and the source component of the "lateral fluxes" associated to human/animal respiration, wood decomposition and lake/river emissions. The use of these "lateral fluxes" appeared to be complicated, in particular due to the potential overlapping with the other flux components and due to the need to temporally disaggregate these fluxes at an hourly scale.

Alternative prior anthropogenic and terrestrial ecosystem CO_2 fluxes

For the one-month tests of sensitivity to the prior estimate of the anthropogenic and terrestrial ecosystem fluxes, the following three flux products are used:

- the COFFEE anthropogenic emission product combining the maps from EDGAR v4.3 at 0.1° resolution, the national/annual budgets from BP statistics, and the TNO temporal profiles (Steinbach et al., 2011, McGrath et al., 2023; available from ICOS-CP) , as an alternative to the TNO WP2 standard product
- an ORCHIDEE simulation at 0.25° resolution of the terrestrial ecosystem fluxes from the H2020 VERIFY project (McGrath et al., 2023), as an alternative to the VPRM WP2 standard product
- a C-TESSEL simulation of the terrestrial ecosystem fluxes from the ECMWF operational forecasts, as an alternative to the VPRM WP2 standard product

Prior fluxes for the NO_x and CO inversions

The prior estimates of the NO_x and CO anthropogenic emissions in the NO_x and CO inversions for France briefly mentioned in this report rely on the Inventaire National Spatialisé (INS) inventory from the INERIS agency in France. Prior natural NO_x and NMVOCs emissions estimates are extracted from simulations with the MEGAN model (as in Fortems-Cheiney et al., 2023). The prior estimate of the initial and boundary conditions are provided by European scale simulations with another configuration of CHIMERE (similar to that of Fortems-Cheiney and Broquet, 2022). In the European-scale NO_x and CO inversions documented in Deliverables 6.4, 6.5 and 6.6, the prior estimate of the NO_x and CO anthropogenic emissions is given by the TNO TNO-GHGco-v3 inventory (see the details in Deliverable 6.4).

Observations

The reference CO_2 inversions and the tests of sensitivity to the prior estimate of the fluxes assimilate alternatively:

- *in situ* hourly CO_2 observations from ground-based continuous measurement stations in France and in its vicinity (mainly from the ICOS network, all accessed from the ICOS carbon portal, https://data.icos-cp.eu/portal/, ICOS_ATC_OBSPACK-Europe-L2-2022, cf Figure 1). This set of stations does not include peri-urban and urban stations dedicated to the monitoring of a specific urban area, and in particular, here, the stations of the Parisian CO_2 network, which cannot be properly represented with a 10 km resolution model. Following the observation selection defined in Broquet et al. (2013) and kept for

the PYVAR-CHIMERE CO_2 NEE inversions in Monteil et al. (2020), observations at low altitude stations are assimilated during the afternoon (here 12:00-18:00 UTC) only, and observations at high altitude stations are assimilated during the night time (here 0:00-7:00 UTC) only. When several levels of measurements are available at a given station, the inversions assimilate the data from the highest level only.

- relatively high resolution (1.29 km x 2.25 km) satellite XCO_2 observations from the OCO-2 NASA-JPL mission (the v11 dataset, cf Figure 1); all individual observations are assimilated and compared to the grid cell of CHIMERE containing the centre of their corresponding OCO-2 ground pixel (there is no aggregation of the observations at the model resolution), computing the appropriate XCO_2 values from the model based on the OCO-2 averaging kernels and prior estimates for the retrieval of the XCO_2 values
- both of these observation datasets.

For both the *in situ* and satellite data assimilation, the inversion system accounts for both transport model and observation errors. The observation error covariance matrix of the system is set up as a diagonal matrix (without spatial or temporal correlation across the observations) with the observation error values provided in the observation products, and with values for the transport model error for the in situ and satellite CO_2 observations respectively taken from Broquet et al. (2013) and Potier et al. (2022).

The NO_x and CO inversions over France and for the period 2019-2021 have been conducted with the assimilation of the relatively high (3.5 x 5.5 km^2 and 7 x 5.5 km^2 respectively) resolution TROPOMI NO_2 tropospheric vertical columns densities (TVCDs, using the TROPOMI-PAL product) and TROPOMI CO columns (OFFLv2.4) from the Sentinel-5P satellite mission. The European-scale inversions documented in D6.4, D6.5 and D6.6 also assimilate satellite data only: the MOPITT “surface” multispectral MOPITTv8-NIR-TIR CO product (for CO inversions), and the OMI OMI-QA4ECV-v1.1 or TROPOMI PALv2.3 NO_2 TVCD products (for the NOx inversions; see the details in Deliverable 6.4).

3.2.1.3 State vector

Here, the different inversion experiments consist of either a one-month inversion or a series of independent one-month inversions: twelve one-month inversions to cover the full year 2018. For each month, the CO_2 inversions control separately the anthropogenic, terrestrial ecosystem and ocean CO_2 fluxes in addition to the model initial and boundary conditions. In particular:

- the anthropogenic (fossil fuel and biofuel) emissions are controlled at the scale of 5 aggregated sectors of activity (public power, industry, other stationary combustion, road transport, other) per administrative region (in France) and per country (outside of France), and at daily temporal resolution; however, when using the EDGAR-COFFEE product as a prior estimate of the anthropogenic emissions in the sensitivity tests, there is no sectoral resolution, so that the system controls the total emissions per administrative region/country at daily temporal resolution
- the ocean and terrestrial ecosystem fluxes are controlled at the model grid cell (i.e. 10-km over France) and 6-hour resolution

The setup for the part corresponding to the CO_2 natural fluxes in the prior uncertainty covariance matrix in the system is derived from that of the PYVAR-CHIMERE CO_2 NEE inversions in Monteil et al. (2020), albeit with 100-km scale spatial correlations for the terrestrial ecosystems (instead of 200-km spatial correlations, since the system operates at higher spatial resolution in this case) and some other slight differences.

The setup for the part corresponding to the CO_2 anthropogenic emissions in the prior uncertainty covariance matrix assumes a 50% one-sigma uncertainty in the total emissions per administrative region and day (i.e. a bit more than 100% 1-sigma uncertainty in the total

emissions per large sector of activity, administrative region and day). It ignores spatial correlations across the regions, and temporal day-to-day correlations.

The prior uncertainty in the boundary conditions is characterised by a 500 km horizontal correlation scale and by a 2 ppm 1-sigma uncertainty in the total columns in the prior error covariance matrix.

The prior uncertainty covariance matrix does not include any correlation between the different main components of the control vector (ocean, terrestrial ecosystem and anthropogenic fluxes, and boundary conditions).

In the NO_x and CO inversions for France, the logarithm of the anthropogenic and biogenic surface emissions are controlled separately, at the model grid cell (i.e. 10-km over France) and at daily resolution. The version of the European-scale NO_x and CO inversions documented in Deliverables 6.4, 6.5 and 6.6 controlled the anthropogenic and biogenic surface emission separately but with a Gaussian uncertainty in the prior estimate of these emissions, at 0.5° and daily resolution (the most recent European scale inversions now control the logarithm of the emissions as the configuration for France). The daily maps of NO_x or CO anthropogenic emissions from these European-scale inversions are converted into estimates of the fossil fuel CO_2 emissions at the national and monthly scale for five large groups of sectors of emitting activities. This conversion relies on the sectoral maps of emissions for the three species and, implicitly, on the emission ratios between the species for each sector, country and month from the TNO inventory used as a prior estimate of the emissions (see Deliverable 6.4 for the details).

3.2.1.4 Period of study

The reference inversions cover the full year 2018, while the one-month sensitivity tests are focused on July 2018, one of the two months of reference for Task 4.4.

3.2.2 Results

3.2.2.1 Reference CO_2 inversions

Fit to the assimilated observations

The reduction of the misfits between the simulation and the assimilated observations due to the corrections applied by the reference CO_2 inversions to their prior estimate of the fluxes and initial/boundary conditions is illustrated in Figure 2 and Figure 3.

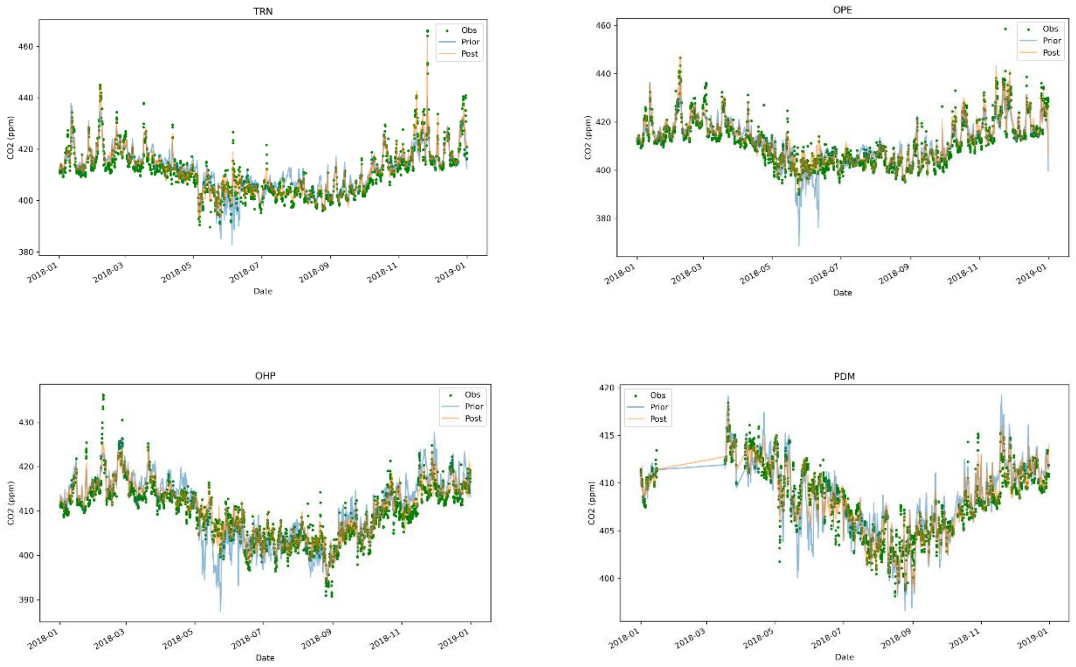


Figure 2: Timeseries of the hourly values of the CO_2 concentrations over the afternoon or night time assimilation window for three low-altitude stations (TRN, OPE and OHP) and one high-altitude station (PDM) over the year 2018: measurements (green dots) and prior (blue) / posterior (orange) simulations in the reference inversion assimilating surface data only.

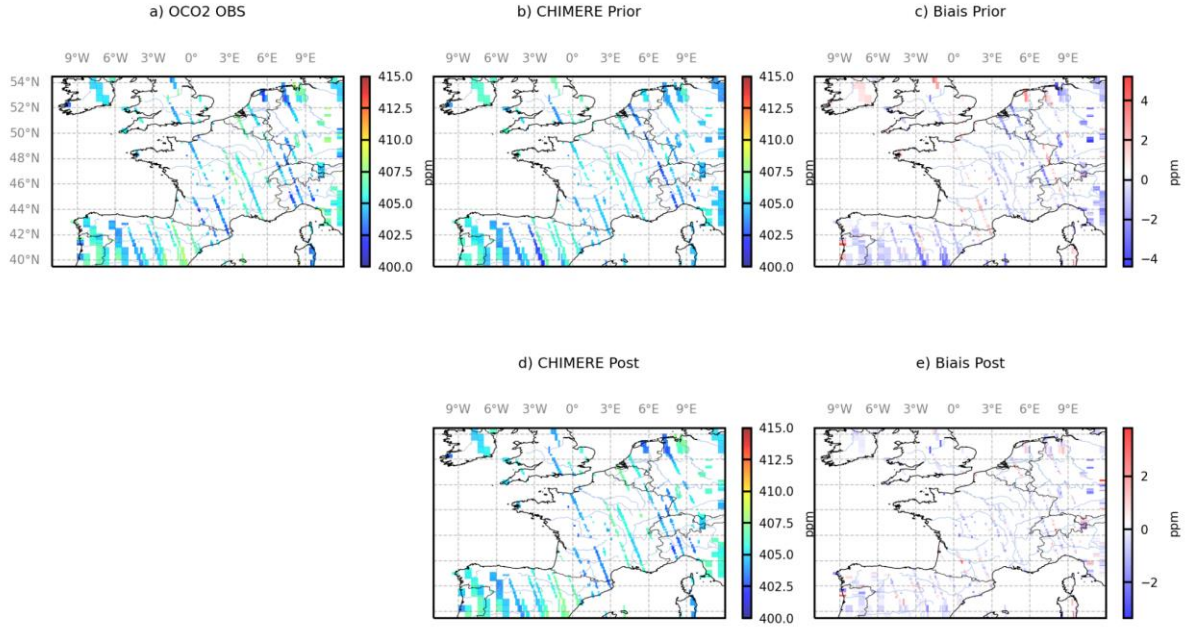


Figure 3: Comparison between the OCO-2 XCO_2 observations and the corresponding CHIMERE XCO_2 simulations in July 2018: averages over the month of the XCO_2 values per grid cell of the model (observations, prior and posterior simulations, and differences)

In the inversion assimilating CO_2 surface data only, the reduction (from the prior to the posterior simulations) of the monthly RMS misfits between hourly simulated vs. measured CO_2 during the assimilation windows from all stations range between 23%-34% in Jan-Mar and Aug-Nov, and between 38%-50% in Apr-Jul. The RMS reduction is significantly lower in

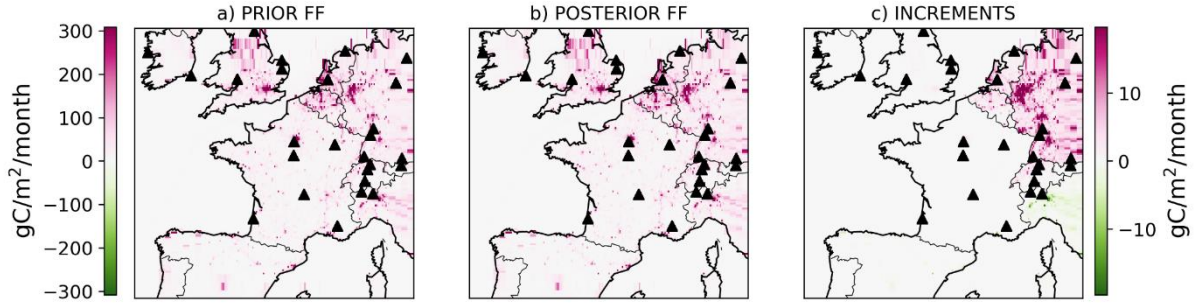
December (12%). There are both low-altitude and high-altitude stations, stations both in France and in neighbouring countries for which the monthly RMS misfit reductions are >50%, especially in spring. Similarly, there are both low-altitude and high-altitude stations, and in France and in neighbouring countries for which the monthly RMS misfit reductions are <15%, especially in winter. The reductions (from the prior to the posterior simulations) of the annual RMS misfits between hourly simulated vs. measured CO_2 during the assimilation windows from individual stations range from only 2 values below 10% i.e. 5% at TOH and 7% at GART, in Northern Germany, very close to the domain's eastern boundary, to a large set of values above 30% corresponding to all types of stations (high- and low-altitude, within or outside of France).

In the inversion assimilating OCO-2 observations only, the prior misfits (often large i.e. >1 or <-1 ppm) between these observations and the prior simulation are generally decreased by much more than 50%, with slight residual posterior misfits close to and often below the level of observation error (monthly posterior biases and RMS misfits to OCO-2 are systematically smaller than 0.1 ppm and 1.3 ppm respectively). Of note is the limited number of OCO-2 tracks for individual months over France illustrated by Figure 3, which helps fitting the available observations through the inversion.

The co-assimilation of surface and satellite observations does not significantly impact the posterior fit to the OCO-2 observations, which remains very close to that when assimilating OCO-2 observations only in terms of RMS errors (the monthly biases increase significantly but remain very low, i.e. <0.1ppm). It has a larger impact on the posterior fit to the surface observations with the reduction (from the prior to the posterior simulations) of the monthly RMS misfits between hourly simulated vs. measured CO_2 during the assimilation windows from all stations ranging from 7 to 39% when co-assimilating the satellite observations. This indicates some inconsistencies in the information brought by the two datasets, and of the larger weight of the satellite data in the corrections applied by the inversion whereas these data provide direct information on the fluxes.

Corrections applied to the prior anthropogenic emissions in France

Anthropogenic emissions, Feb 2018, assimilation of surface and OCO-2 observations



Anthropogenic emissions, July 2018, assimilation of surface and OCO-2 observations

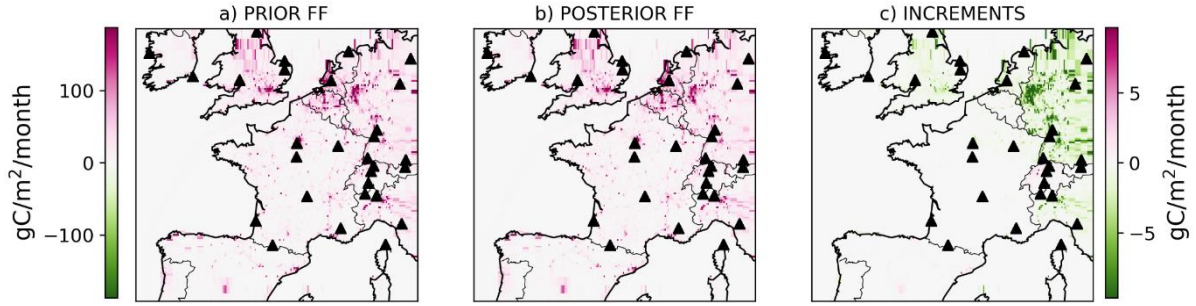
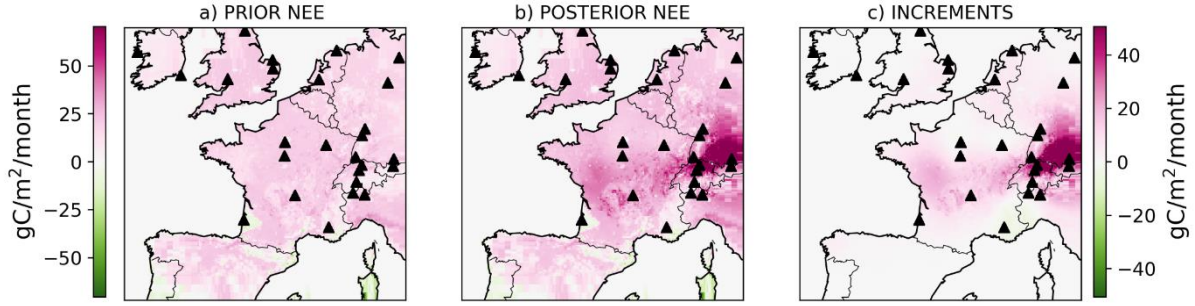


Figure 4: Maps of monthly mean anthropogenic emissions in the modelling domain in February (top row) and July (bottom row) 2018 when assimilating both surface and OCO-2 CO_2 observations: prior estimate, posterior estimate and differences (i.e. increments applied by the inversions to the prior estimate)

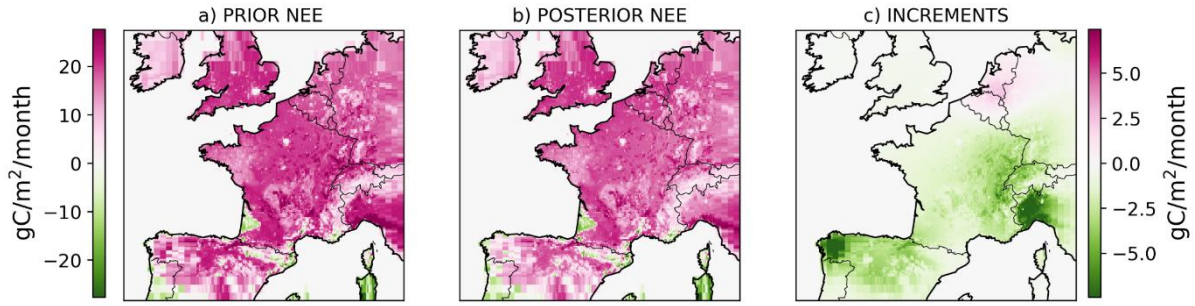
The most critical pattern from all three reference inversions with respect to the main objective of Task 4.4 is the lack of corrections from the inversion to the prior estimate of the anthropogenic emissions in France throughout the year, whether assimilating surface observation, OCO-2 observations, or both (see Figure 4). Some slight positive or negative corrections can be seen for the major French emission hotspots such as the Paris urban area and in north-eastern France, but this is very small in terms of relative differences to the prior estimate of the emissions. Larger corrections arise in neighbouring countries, especially in Germany, but they remain limited in amplitude, and can partially be explained by the need to handle residual errors in the background conditions despite the control of the boundary conditions. The variations of concentrations across the network of stations or along the OCO-2 tracks in the modelling domain is hardly sensitive to the uncertainties in the anthropogenic emission estimates whose signal is overwhelmed by that of the uncertainties in the terrestrial ecosystem fluxes. The inability to make use of existing peri-urban/urban stations dedicated to specific urban areas with the 10-km spatial resolution of the model exacerbates it: the sites for the surface stations used in the reference inversions have been primarily chosen to monitor natural fluxes, and are located relatively far from large urban and industrial areas. The OCO-2 tracks hardly catch major anthropogenic plumes from emission hotspots in France, as illustrated by the analysis of plume transects in the OCO-2 and OCO-3 data in WP6 and Task 4.2 (Chevallier et al., 2022). Finally, the accuracy of the estimates of national / annual budgets of the anthropogenic emissions in France from the TNO inventory (fed by the French National Inventory Report -NIR- to UNFCCC) may also partially explain the lack of corrections for these emissions.

Corrections applied to the prior terrestrial ecosystem fluxes in France

NEE, Feb 2018, assimilation of surface observations



NEE, Feb 2018, assimilation of OCO-2 observations



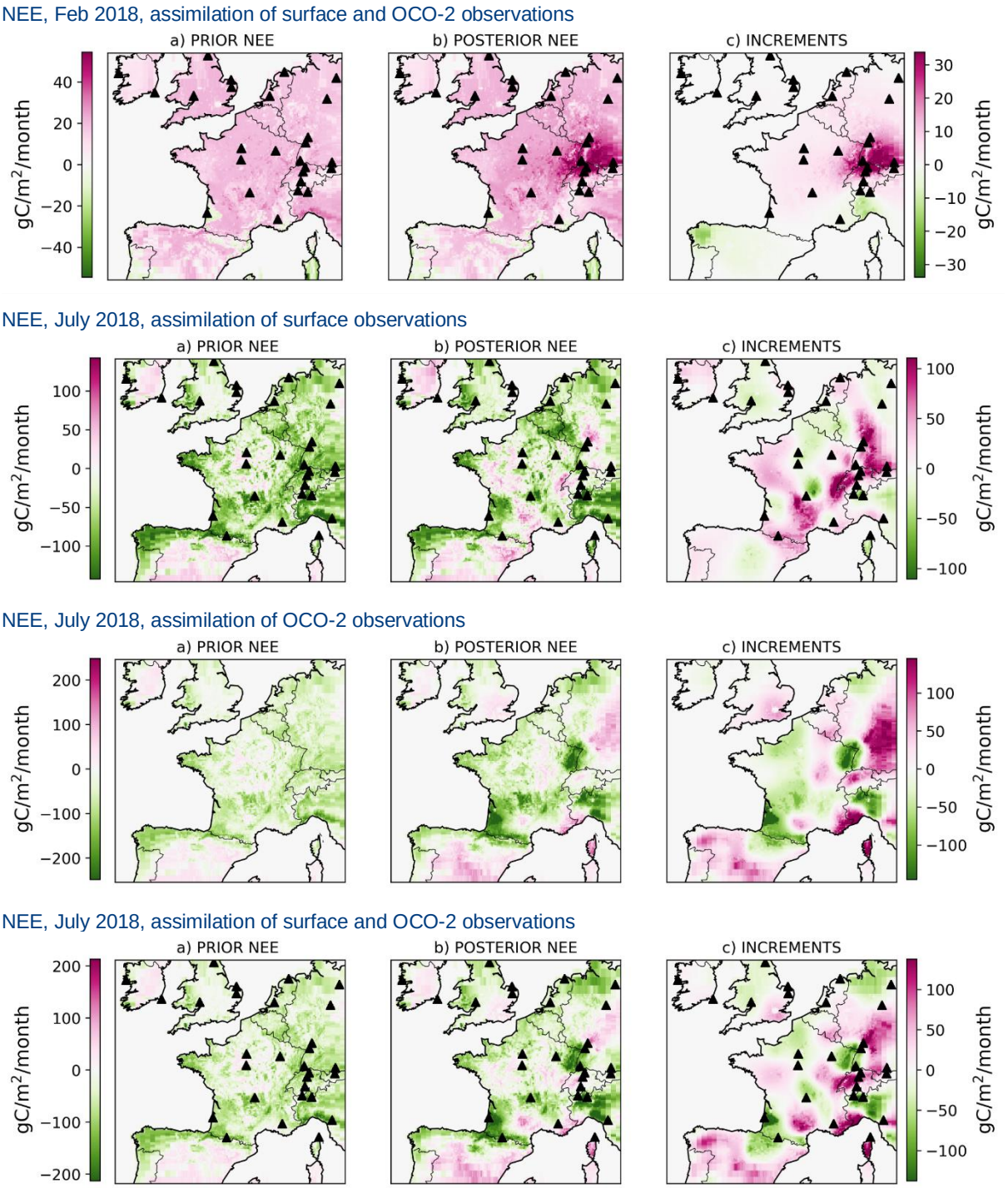


Figure 5: Maps of monthly mean NEE in the modelling domain in February (top 3 rows) and July (bottom 3 rows) 2018 when assimilating surface (1st and 4th rows), OCO-2 (2nd and 5th rows) and surface and OCO-2 (3rd and 6th rows) CO_2 observations: prior estimate, posterior estimate and differences (i.e. increments applied by the inversions to the prior estimate)

The terrestrial ecosystem fluxes, and more specifically here, the Net Ecosystem Exchange (NEE) was not the primary target of the relatively high-resolution inversion experiments in Task 4.4. However, even with the national-scale configuration for France at 10 km resolution over the country, it is the surface flux by far the main surface flux components controlled by the inversion assimilating observations from the existing surface and OCO-2 observations in any area of the country, including the Paris region (see Figure 5). In all three reference

inversions, the typical amplitude of the local corrections to the prior NEE is more than one order of magnitude larger than that of the corrections seen for the anthropogenic emissions in the Paris area.

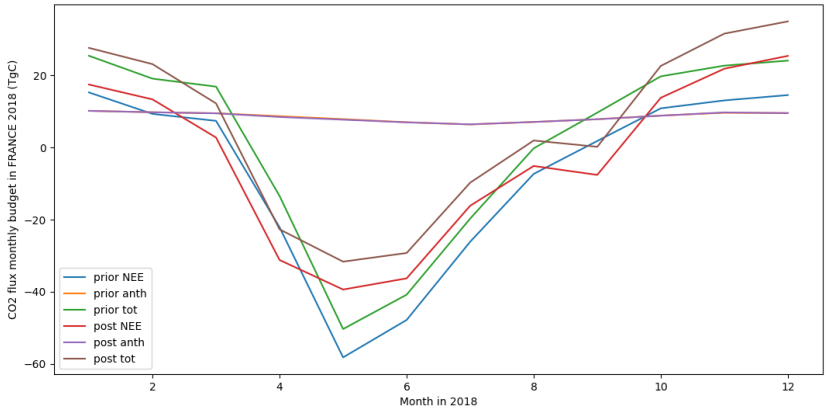
The corrections applied to the prior estimate of the NEE tend to be highly heterogeneous in space and time when assimilating the surface network and/or the OCO-2 satellite data as illustrated by Figure 5. A critical result is the large differences in terms of sign, spatial patterns, and amplitudes between the corrections (increments) from the assimilation of *in situ* data and from the assimilation of OCO-2 data. The lack of consistency between the constraints brought by the two datasets is a source of concern for the robustness of the NEE inversions: one can hardly assume that these differences are due to the fact that the two observing systems catch the signal from different areas, unless one assumes that their footprint would be too narrow to support a real national-scale monitoring of the fluxes. The surface networks have been used for more than a decade for the regional inversion of NEE (e.g. Broquet et al., 2011) while the assimilation of satellite data for such an activity remains relatively recent. Therefore, there is more confidence in the results from the assimilation of surface data.

When assimilating both datasets, the results in winter are dominated by the constraint from the surface network, as revealed by the similarity between these results and those when assimilating surface data only. This can be explained by the relatively low number of OCO-2 tracks and valid observations over France in winter when the cloud cover and solar zenith angles are higher (see the results for Feb 2018 in Figure 5). However, in spring-summer, there is a balance between the constraints from the two datasets, with the local corrections being alternatively dominated by one or the other dataset (as revealed by the similarity to the local corrections applied when assimilating the datasets separately). Along the main OCO-2 ground tracks (e.g. along the Normandy/Northwest - Provence/Southeast diagonal of France in July 2018, see Figure 5) the corrections are generally dominated by the constraint from the satellite, with the significant observation errors associated with the assimilation of these satellite data being balanced by the dense observation coverage of OCO-2 along its tracks. In a general way, the combination of the two datasets leads to maps of increments that significantly differ from those from the assimilation of a single dataset, so that this combination results in a complex synthesis.

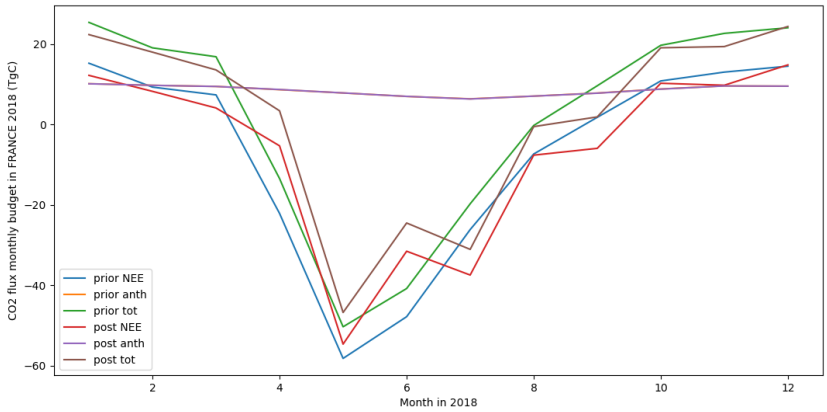
Monthly and annual budgets of anthropogenic and terrestrial ecosystem fluxes in France

The resulting monthly and annual budgets of anthropogenic emissions and NEE over France in 2018 from the three reference inversions are illustrated in Figure 6 and Table 2.

Assimilation of surface observations



Assimilation of OCO-2 observations



Assimilation of surface and OCO-2 observations

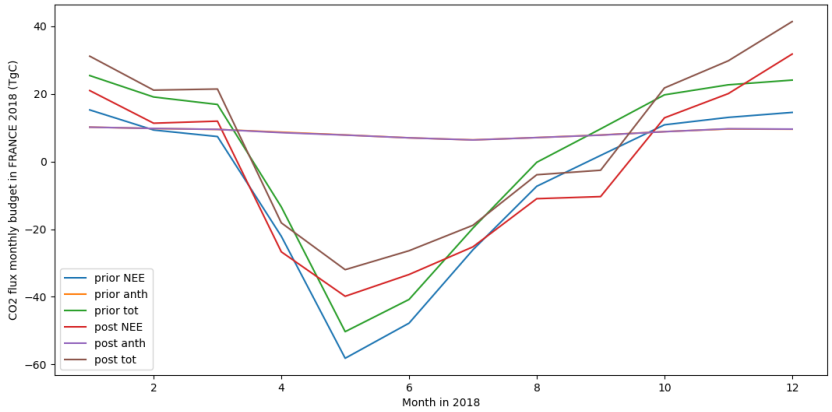


Figure 6: Monthly mean budgets of anthropogenic emissions, NEE and total CO_2 fluxes in France in 2018 when assimilating surface (1st subfigure), OCO-2 (2nd subfigure) and surface and OCO-2 (3rd subfigure) CO_2 observations: prior and posterior estimates. Note that the curves of the prior anthropogenic emissions are hidden by those of the posterior anthropogenic emissions (due to their similarity)

Table 2: Annual budgets of NEE and anthropogenic emissions in France in 2018 from the three reference inversions, in TgC

	NEE prior	NEE post	Anthropogenic emission prior	Anthropogenic emission post
Assim surf	-89	-41	102	102
Assim OCO-2	-89	-83	102	102
Assim surf+OCO-2	-89	-37	102	102

This figure and table confirm and summarise (i) the lack of correction applied by the reference inversions to the prior estimate of the anthropogenic emissions, (ii) the discrepancies between the corrections to the prior estimate of the NEE from the assimilation of surface observations

and from the assimilation of OCO-2 observations (iii) the complexity of the combination of the constraints on the NEE estimates from these two datasets when both are assimilated.

When assimilating OCO-2 observations only, the balance between positive corrections in spring and negative corrections in summer-fall-winter to the prior NEE monthly budgets leads to a weak correction of the annual budgets of NEE, and a posterior estimate of a sink of 83 TgC which stays close to the high prior value of a sink of 89 TgC over France in 2018. The resulting seasonal cycle is flattened compared to the prior one. The assimilation of the surface data with or without satellite observations tends to lead to positive corrections in fall-winter, but also in spring, leading to both a positive shift and a flattening of the seasonal cycle from the prior estimate of the NEE, and to an annual sink of 41 to 37 TgC in France, much smaller than in the prior estimate of the NEE. Overall, the weight of the surface network appears to be larger than that of the satellite observations in the correction of the annual budget of NEE, which can be explained by the limited number of satellite tracks and valid observations in France.

3.2.2.2 Sensitivity to the prior estimates of the anthropogenic emissions and NEE: one-month tests for July 2018

The list of one-month tests of the sensitivity of the result to the prior fluxes and the corresponding prior flux products are given and labelled in Table 3. The labels of these tests can correspond to inversions assimilating the surface CO_2 observations only, the satellite data only, or both datasets (as indicated in the figures or in the text).

Table 3: The one-month CO_2 inversion in July 2018: reference inversion and sensitivity tests

CO_2 inversion experiment	**Prior estimate of the anthropogenic emissions**	**Prior estimate of the NEE**
Reference	TNO	VPRM
A	TNO	ORCHIDEE
B	TNO	C-TESSEL
C	EDGAR-COFFEE	VPRM
D	EDGAR-COFFEE	ORCHIDEE
E	EDGAR-COFFEE	C-TESSEL

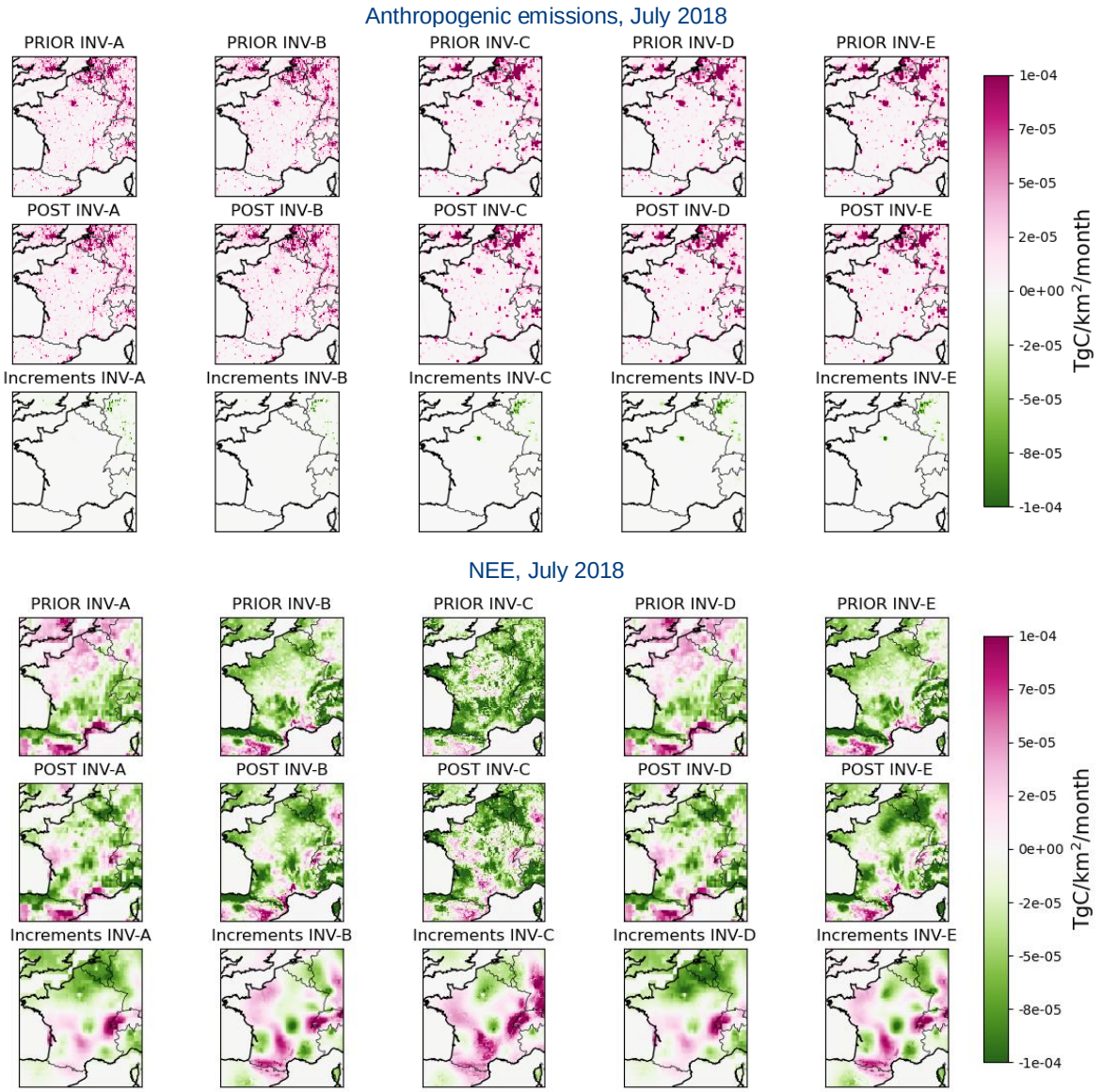
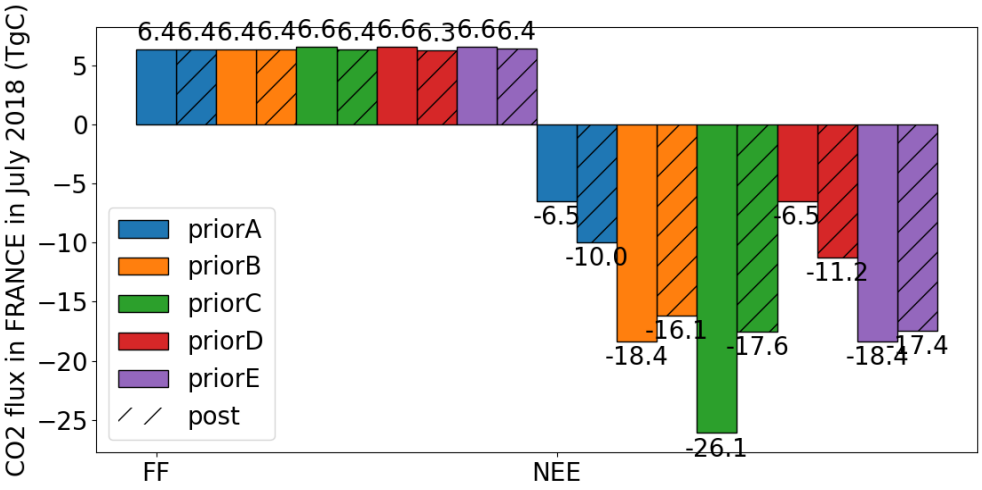
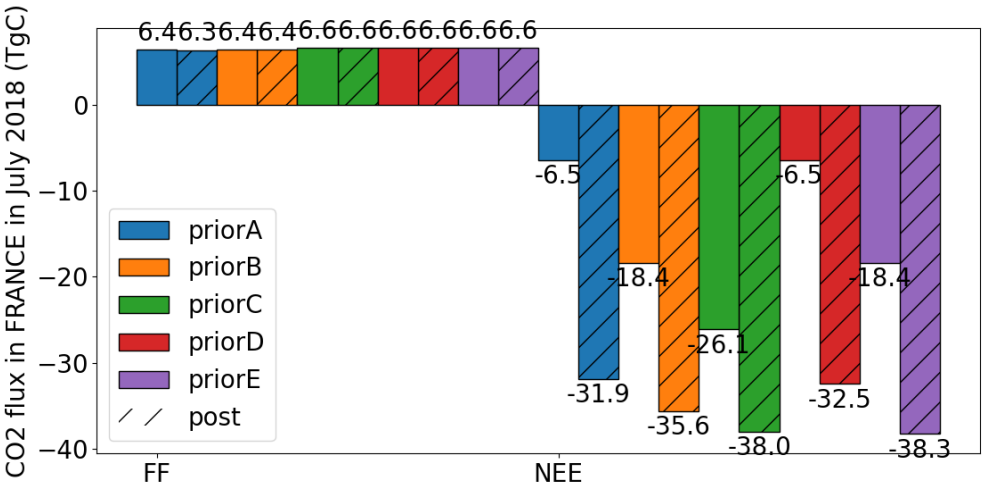


Figure 7: Maps of monthly mean anthropogenic emissions (top 3 rows) and NEE (bottom 3 rows) in the modelling domain in July 2018 when assimilating surface CO_2 observations only: prior estimate, posterior estimate and differences (i.e. increments applied by the inversion to the prior estimate) for the different tests of sensitivity (see Table 3 and comparisons to Figure 5 for the reference inversion)

Assimilation of surface observations



Assimilation of OCO-2 observations



Assimilation of surface and OCO-2 observations

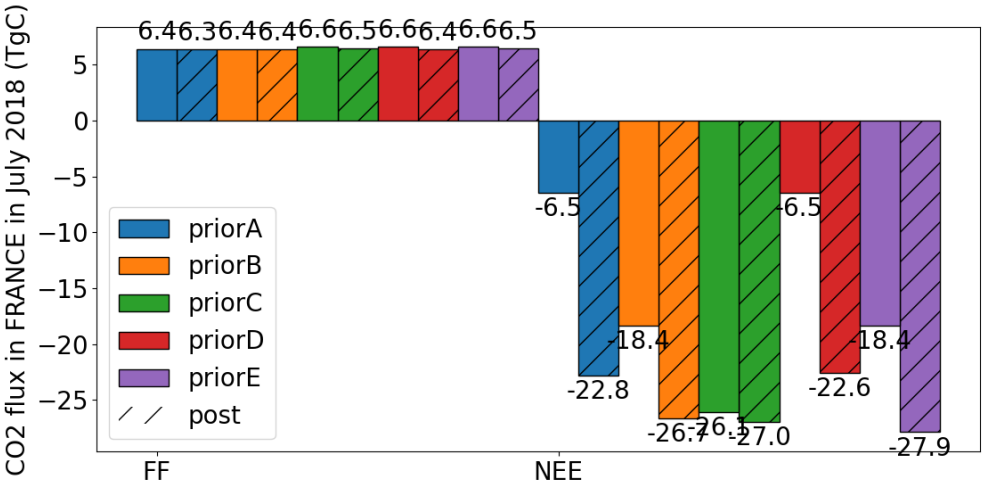


Figure 8: Monthly budgets of anthropogenic emissions (denoted FF like fossil fuel but containing the biofuel emissions) and NEE in France in July 2018 when assimilating surface (1st subfigure), OCO-2 (2nd subfigure) and surface and OCO-2 (3rd subfigure) CO_2 observations: prior and posterior estimates for the different tests of sensitivity (see Table 3 and comparisons to Figure 6 for the reference inversions)

The main conclusions arising from this series of sensitivity tests are illustrated in Figure 7 and Figure 8:

1) There is a lack of correction to the anthropogenic emissions in all cases. Using the EDGAR-COFFEE product as a prior estimate of the emissions (with a control of the total rather than sectoral budgets of the anthropogenic emissions in the administrative regions) does not alter this behaviour, even though the corrections are slightly larger when using the EDGAR-COFFEE product than when using the TNO inventory when assimilating surface observations. There is a strong consistency at the national scale between the two inventory products, and the higher corrections from the inversion to the EDGAR-COFFEE product than to the TNO inventory actually reduce the slight (0.2 TgC for France in July 2018) misfit between these two products when assimilating surface observations or both surface and satellite observations. As a result, all inversions assimilating the same set of observations have similar posterior estimates of the national budgets of anthropogenic emissions in France in July 2018 (6.3-6.4 TgC when assimilating surface observations only, 6.3-6.6 TgC when assimilating satellite data only, and 6.3-6.5 TgC when assimilating both datasets). Such a result could strengthen the assumption that the lack of corrections to the anthropogenic emissions in the inversion is due to the high accuracy of the prior emission estimates. However, the spatial distribution of these emissions is relatively uncertain and differs between the two inventories *a priori* and *a posteriori* (the distribution of the emissions in the TNO inventory being more diffuse than in the EDGAR-COFFEE inventory). Therefore, the convergence of the difference inestimates of anthropogenic emissions from 0.2 TgC (prior) to less than 0.1 TgC (posterior) in July 2018 when assimilating surface data only could simply be fortuitous, especially since the stronger corrections to the TNO inventory than to the EDGAR-COFFEE inventory leads to an increase in the spread of posterior estimates compared to the prior estimates when assimilating satellite data only.

The results for NEE are not impacted much by the change of prior estimate of the anthropogenic emissions, which strengthens the conclusion of the lack of sensitivity to these prior emissions. They differ by ~1 TgC in terms of the budget of NEE for France in July 2018 when using the TNO vs EDGAR-COFFEE inventory as a prior estimate of the anthropogenic emissions when assimilating surface observations (see A vs. D, B vs. E and C vs. the reference inversions when assimilating surface observations only or both surface and satellite observations), which is significantly higher than the differences between the two

anthropogenic emission inventories, even though it is relatively small compared to the typical NEE budgets obtained with the different inversions (see below). This may reveal some sensitivity at some of the surface stations to the local anthropogenic emissions (impacting the corrections to the NEE when changing the inventory for the prior anthropogenic emission estimate), despite the general lack of sensitivity of the surface network to the uncertainties assigned to the anthropogenic emissions in the inversion system. This would relate to the limited capability to account for the large uncertainties in the local distribution of the anthropogenic emissions when controlling these emissions at regional and not at local scale. This may, however, require some precise knowledge of the local uncertainties and of their spatial and temporal correlations, which is challenging (Super et al. 2020). The problem is a bit more significant when assimilating only the limited amount of available OCO-2 observations, with differences between the posterior estimates of the NEE at the monthly and national scale reaching 2.7 TgC between test cases B and E (i.e. with C-TESSEL outputs as the prior estimate of the NEE).

2) There is a convergence of the estimates of NEE, illustrated by the comparison of the spread of posterior estimates of NEE when assimilating a given set of observations (over France in July 2018: -10 to -17.6 TgC when assimilating surface data only, -31.9 to -38.3 TgC when assimilating satellite data only, and -22.6 to -27.9 TgC when assimilating both datasets) to the spread of the prior estimates of the NEE (-6.5 to -26.1 TgC) in this ensemble of experiments. While the ensemble of prior NEE maps is heterogeneous, the posterior maps of NEE when assimilating surface observations along with satellite data or not all exhibit consistent large-scale patterns, such as the sinks in northeast France, Brittany, and southwest France, and the sources in southern France, the northern Alps and western France. This provides confidence in the robustness of the inversions, especially when assimilating surface data. The ensemble of posterior maps from the tests assimilating satellite data only are also consistent between themselves, but they bear some significant differences locally to those when assimilating surface data. Furthermore, the significant residual discrepancies between the posterior estimates of the NEE from the different inversions, especially when considering the full ensemble of posterior estimates from all the tests assimilating different datasets (with a spread of values between -10 to -38.3 TgC for France in July 2018) reveal the limits of the constraint from the existing observing system and the corresponding uncertainties in the NEE estimates.

3.2.2.3 Insights on the potential of the co-assimilation of satellite NO_2 and CO observations from NO_x and CO inversions

The following consists of a brief summary of some of the current conclusions from Robin Plauchu's PhD research at LSCE (in the frame of the French ANR ARGONAUT and ADEME-AQACIA Lock'Air projects) which has not been finalised yet, and from the European-scale experiments initiated in the H2020 project VERIFY, extended and documented in CoCO2 WP6 (D6.4, D6.5, D6.6) and in the ESA World Emission project.

The analysis of the TROPOMI CO measurements over France does not indicate a sufficient sensitivity of the current TROPOMI CO observation products to the anthropogenic emissions in the country. The images do not show clear patterns of plumes from the major French urban areas. When applying the corresponding averaging kernels to the vertical columns of CO simulated by CHIMERE, the resulting simulations of satellite images do not show such spatial patterns either, while the model exhibits large plumes of CO near the surface downwind of major urban and industrial areas. Consequently, the CIF-CHIMERE CO inversions assimilating the TROPOMI CO observations do not apply significant corrections to their prior estimate of the CO anthropogenic emissions. Instead, corrections are applied mainly to the initial and boundary conditions of the model.

At European scale, the assimilation of the MOPITT CO surface product, whose sensitivity to the signature of the anthropogenic emissions near the surface appear to be larger than that of the current TROPOMI CO product over France, but whose spatial resolution is coarser than

that of TROPOMI, brings a stronger but still limited constraint for the derivation of the anthropogenic emissions, so that the corrections applied by the CO inversions to the prior estimate of the emissions from the TNO inventory at the European to national scales is moderate (see D6.5, D6.6 and Fortems-Cheiney et al. 2023).

The NO_x inversions for France or at the European scale demonstrate some capabilities to control the NO_x anthropogenic emissions at national scale based on TROPOMI and even the OMI NO_2 TVCDs. However, the non-linearity of the chemistry raises computational challenges for the iterative scheme of the variational inversion. This and the large observation errors associated with the current NO_2 data limit the ability to fit these data and the amplitude of the corrections to the prior estimate of the emissions (see D6.5). This challenge adds to that of the uncertainties in the NO_x-to-CO_2 anthropogenic emission ratios in the context of the co-assimilation of NO_2 and CO_2 observations for the monitoring of the CO_2 anthropogenic emissions (Fortems-Cheiney and Broquet, 2022, Deliverables 6.4, 6.5 and 6.6).

The current results, including those from Deliverable 6.5, raise some positive perspectives regarding the use of co-emitted species to increase the ability to control the anthropogenic CO_2 emissions in national scale inversions, starting from the overall consistency between the spaceborne observations and simulations of the NO_2 and CO concentrations in European countries where the anthropogenic emission estimates are relatively accurate. They show that the satellite NO_2 observations are currently more promising than the satellite CO observations over a country like France. However, the detailed analysis of the NO_x inversion behaviour and results indicate that much developments and work are still needed to refine the NO_x inversions themselves before tackling the co-assimilation of NO_2 and CO_2 observations.

3.3 ICON-ART-CTDAS for CO_2 and CH_4 inversions from in situ and satellite observations (Empa)

3.3.1 Model description for the CO_2 inversions

3.3.1.1 Transport model

The atmospheric transport model in this inversion system is ICON-ART, which is composed of the icosahedral non-hydrostatic weather and climate model ICON and the ART extension for the simulation of aerosols and reactive trace gases. ICON is a highly versatile model for global and regional weather and climate simulations developed jointly by the German Weather Service (DWD) and the Max Planck Institute of Meteorology. Tracers are transported with perfect mass conservation by solving the continuity equation of mass for each tracer consecutively in the vertical with a finite volume method and in the horizontal direction with a simplified flux-form semi-Lagrangian method. Here, ICON-ART was used in a limited area configuration with a grid spacing of about 13 km and with meteorological initial and boundary conditions obtained from ERA5 reanalyses (Figure 9).

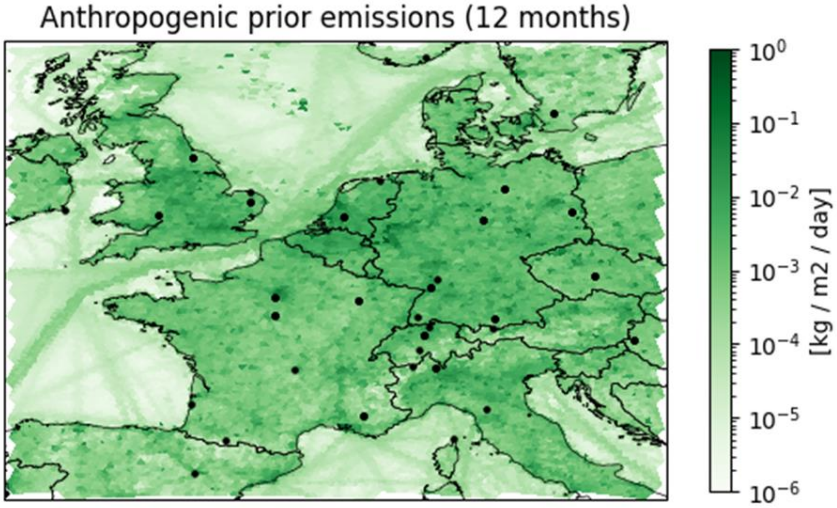


Figure 9: ICON-ART simulation domain with a priori anthropogenic CO_2 emissions and the locations of in-situ CO_2 measurement stations used in the inversion. ICON-ART has a semi-structured grid composed of triangles. The average distance between triangles (resolution) was about 13 km.

The ART-model was developed as an extension for ICON at the Karlsruhe Institute of Technology (KIT) with the aim of simulating aerosols as well as passive and chemically reactive trace gases. The ART module is coupled online with ICON and allows a flexible definition of tracers and processes to be included. Since only CO_2 was simulated in the present study, all chemistry and aerosol processes were switched off and all CO_2 tracers were treated as fully passive with no feedback on meteorology.

To simplify and accelerate the treatment of emissions during the simulations, we implemented the Online Emissions Module (OEM) into ICON-ART, which was originally developed for the regional weather and climate model COSMO (Jähn et al. 2020). OEM allows source-specific temporal and vertical emission profiles to be applied online, which greatly reduces I/O during the simulation. Furthermore, the Vegetation Photosynthesis and Respiration Model (VPRM) was integrated in OEM in order to simulate the exchange of CO_2 between the atmosphere and the land vegetation online during the simulation, driven by near-surface temperature and radiation from ICON and by two satellite indices (enhanced vegetation index EVI and the Land Surface Water Index LSWI) from MODIS satellite observations. The standard land cover data set used in ICON is GlobCover, which proved to be insufficiently accurate for representing the vegetation over Europe. We therefore exchanged the land cover data set with CORINE, which is limited to Europe but describes vegetation cover and land use much more accurately than the global data set GlobCover. Although not identical, the fluxes computed online in this way are largely consistent with the offline VPRM product provided through the modelling protocol (e.g. the same set of VPRM parameters was used).

For inverse modelling, we coupled ICON-ART with the CarbonTracker Data Assimilation Shell (CTDAS), which is an Ensemble Kalman Smoother originally developed for the estimation of biospheric CO_2 fluxes at the global scale (Peters et al., 2005). To couple ICON-ART with CTDAS in a robust and efficient way, we made a few adaptations to ICON-ART and the simulation setup. CTDAS requires a large ensemble of tracers to be simulated, each ensemble member corresponding to one specific perturbation of the state vector, e.g., fluxes and boundary conditions (see next section). Instead of generating these perturbed fields offline and reading them in during simulation, as is usually done, we extended the OEM module with the option to generate the ensemble of perturbed fluxes and boundary conditions and corresponding tracers online during the simulation. With this extension, the only input required

at the start of a simulation is the ensemble of perturbed scaling factors provided by CTDAS, which greatly simplifies and accelerates the simulations. Each scaling factor scales the flux of one emission category (one tracer can experience the emissions of multiple categories) in one region. The regions can be any combination of grid cells (including individual cells) and are defined by a region mask provided as an input for OEM. In order to keep the simulation close to the real meteorology, ICON was weakly nudged towards ERA5 meteorology in the whole model domain. This was achieved by providing the meteorological forcing data from ERA5 not only at the domain boundaries but in the whole model domain and defining a minimum nudging strength in the inner parts. Further details of the ICON-ART-CTDAS system and a first application to the estimation of European CH_4 emissions are presented in Steiner et al. (2023). Any results submitted using ICON-ART-CTDAS regarding CH_4 come from the results of Steiner et al. (2023). In the following, we describe the newly obtained CO_2 inversion results in detail while the CH_4 results are summarised only briefly.

For the CO_2 simulations, the following model setup was used:

- **Region**: [-8.639°, 17.856°] lon, [40.621°, 58.962°] lat; 21 184 horizontal cells in total; effective horizontal resolution is 13.15 km. Extracted from the ICON R6B3 parent grid.
- **Meteorological IC/BC**: ERA5 (reanalysis-era5-complete, https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-complete?tab=overview). For ICON, we furthermore post-process the soil parameters following the script `mars4icon` that is shipped with ICON.
- **Atmospheric IC/BC**: CAMS global inversion-optimised greenhouse gases (https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-greenhouse-gas-inversion?tab=overview). The CAMS data were interpolated onto the ERA5 dataset using linear interpolation on the log pressure axis vertically and linear interpolation horizontally.

3.3.1.2 Prior fluxes and assimilated observations

Prior fluxes were chosen according to the protocol. As will be mentioned again later, only the first two (anthropogenic and biospheric fluxes) were optimised, the others were fixed.

- **Prior anthropogenic emission fluxes**: TNO_GHGco_6x6km_v4_0_year2018 inventory for the year 2018, with vertical profiles and temporal profiles as specified in the TNO dataset.
- **Prior biospheric fluxes**: computed by VPRM online during the ICON-ART simulation, driven by MODIS/Terra surface reflectance imagery version 6.0 (since superseded by version 6.1 at https://e4ftl01.cr.usgs.gov/MOLT/MOD09A1.061/), and the CORINE land cover dataset at 100 m resolution (https://land.copernicus.eu/en/products/corine-land-cover/clc2018). The VPRM parameters used are those from Gerbig et al., https://meta.icos-cp.eu/objects/JmRwpxQzfbHzwzH26abwt10z.
- **Prior forest fire fluxes**: Global Fire Emissions (GFAS, https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-fire-emissions-gfas?tab=overview). The data were averaged over the 10-day interval we simulate over. Emissions are from the surface only, the temporal profile is constant over time.
- **Prior ocean emissions**: from Cyril Germineaud at Mercator (as per the protocol). The data are averaged over the 10-day simulation interval. Emissions are from the surface only, the temporal profile is constant over time. An error was made in our setup, such that no negative ocean fluxes were included. Since the contribution of ocean fluxes is very small, the impact on our results is expected to be negligible.
- **Prior lateral fluxes**: from Frederic Chevallier (as per the protocol); computed as: allcropsource-biofuelcropsource+allwoodsource-biofuelwoodsource+lakeriveremiss. Again, the data were averaged over the 10-day simulation interval. Emissions are from the surface only, the temporal profile is constant over time.

In terms of assimilated observations, we optimised the following set of data:

- **ICOS ObsPack ground station observations**: ICOS_ATC_OBSPACK-2022 (https://meta.icos-cp.eu/collections/w7dcCpKXsyn-NuVlzXhcX5qN). When multiple inlet heights were available (e.g., the Cabauw station has inlets at 27, 67, 127 and 207 metres above ground level), only the data from the highest level were assimilated. The exact list of stations and the values used for model data mismatch (MDM) are given in Table 4 below. The MDM is based on a prior run for a full year, then taking 2.5+standard-deviation-at-the-station in ppm. For a *normal ground station*, we took the daytime mean observation from 12-17 UTC; for a *mountain station* we took the night time mean observations from 0-7 UTC. Furthermore, for a *normal ground station*, we sampled ICON-ART at an altitude corresponding to the inlet height over the model topography (which does not necessarily correspond exactly to the inlet height over the *true* ground level); for a *mountain station* we sampled ICON-ART at the true elevation over sea level.
- **OCO-2 total column observations**: OCO2_L2_Lite_FP_11r (https://disc.gsfc.nasa.gov/datasets/OCO2_L2_Lite_FP_11r/summary). Only observations with "XCO2_Quality_Flag==0" were used. Simulated ICON-ART columns (to about 25 km altitude) were vertically extended to the top of the atmosphere with CAMS data over the highest available pressure level. The model data mismatch was set to 2 ppm plus the reported 1-sigma uncertainty of the column-averaged dry air mole fraction (typically smaller to 1 ppm).

Horizontal sampling of the ICON-ART data towards the observations was done using an inverse-distance-weighting interpolation method.

Table 4: The ICOS ground station name and inlet height above the ground level in metres (after the final underscore), the assigned model data mismatch (MDM), and indication if a station is treated as a mountain station (indicated with an X in the column).

Station Name and inlet height (m)	MDM (dry air ppm)	Mountain station
Beromunster_212	8.24	
Bilsdale_248	5.85	
Biscarrosse_47	5.60	
Cabauw_207	8.61	
Carnsore Point_14	4.19	
Ersa_40	4.40	
Gartow_341	6.57	
Heidelberg_30	10.66	
Hohenpeissenberg_131	6.06	X
Hyltemossa_150	5.49	
Ispra_100	11.61	
Jungfraujoch_5	3.08	X
Karlsruhe_200	10.05	
Kresin u Pacova_250	5.83	
La Muela_80	5.21	
Laegern-Hochwacht_32	11.56	
Lindenberg_98	7.44	
Lutjewad_60	7.65	
Monte Cimone_8	3.73	X
Observatoire de Haute Provence_100	6.15	
Observatoire perenne de l'environnement_120	8.89	
Pic du Midi_28	3.22	X
Plateau Rosa_10	3.32	X
Puy de Dome_10	5.45	X

Ridge Hill_90	7.09	
Saclay_100	8.87	
Schauinsland_12	5.79	X
Tacolneston_185	6.67	
Torfhaus_147	6.62	
Trainou_180	7.82	
Weybourne_10	6.47	
Zugspitze_3	3.68	X

3.3.1.3 State vector for the CO_2 inversions

We used CTDAS for the inversion, using cycles of 10 days and 2 lags. Each state vector element is thus optimised twice, first based on observations in the current 10-day window and then based on observations in the next 10-day window. The state vector has size $n_{cells} \times n_{categories} \times n_{lags} + 8 = 84\,744$ per cycle, where:

- $n_{cells} = 21\,184$, that is, we optimised each cell of the ICON-ART simulation.
- $n_{categories} = 2$, that is, we optimised the anthropogenic fluxes (assuming a prior uncertainty of 50%, and a specific length of 200 km assuming exponential decay in the covariance matrix) and the net biospheric fluxes consisting of the sum of the respiration and the gross photosynthesis production (assuming a prior uncertainty of 100%, and a specific length of 300 km assuming exponential decay in the covariance matrix).
- $n_{lags} = 2$, that is, we optimised each time period of 10 days twice.
- $+8$, that is, we had 8 inflow regions controlling the background CO_2 concentration as taken from CAMS, which were optimised separately (see Steiner et al. (2023) for more details), with an assumed prior uncertainty of 1.5% (roughly 6 ppm) and a 25% covariance between neighbouring inflow regions.

No correlations between the categories were assumed.

We applied localization in the inversion, by multiplying the result for each assimilated observation with an exponential decay function of 400 km centred on the observation location.

We generated an ensemble of 180 members for each cycle, with CO_2 = λ_{anth} $CO2_{anth}$+λ_{VPRM} $CO2_{VPRM}$+λ_{BG} $CO2_{BG}$+ forest fire emissions + ocean emissions + lateral fluxes, where λ_{anth} and λ_{VPRM} are varying spatially, while λ_{BG} consists of only 8 values for each of the 8 inflow regions).

3.3.1.4 Period of study

02/01/2018 – 27/12/2018 (i.e., January 2, 2018, until December 27, 2018). The first and last 10-day cycle are only optimised once. Due to the limited resources, no simulations for 2021 could be conducted.

3.3.2 Results of the CO_2 inversions

Two inversions were carried out, differing in the observations that were assimilated. Inversion (1) assimilated only in-situ ICOS stations; inversion (2) assimilated in-situ ICOS stations and OCO-2 satellite observations. The results section is split into three subsections. The first subsection shows the comparison of inversion 1 to observational data and inversion results on the model scale; the second subsection shows similar results for inversion 2; the third subsection summarizes the results on a country-scale basis, and discusses the overall results.

3.3.2.1 Assimilating ICOS station data only

Figure 10 shows the timeseries of hourly CO_2 dry mole fractions at station 'Bilsdale' at 42 m altitude (https://meta.icos-cp.eu/objects/lVo3gLCopZerMjEogqkRAns0), subsampled to the assimilated times (12-17 UTC for normal stations, 23-5 UTC for mountain stations) for real observations, a prior simulation and the posterior simulation ('optimised'). This station was chosen merely as an example. We can see that the posterior simulation (in black) moves the simulated values closer to the observed values compared to the prior. This result is also quantified in the error plot below the timeseries, with the root mean square error (RMSE) and bias both decreasing, and the Pearson product-moment correlation coefficient ('r') rising.

Figure 11 summarises the quantitative results now for all ICOS stations in a bar plot. We can generally see similar trends as for the Bilsdale 42 m altitude station, with RMSE values dropping, biases moving closer to 0, and R-squared ('r^2') values rising. Exceptions exist for the Křešín u Pacova and La Muela stations which have a more modest increase in their R-squared scores, but these are stations at the edge (in the Czech Republic and Spain, respectively) which suggests that we are dealing with edge artefacts. Figure 12 shows a Taylor diagram summarizing the results for all the stations, similarly showing and suggesting an improved performance when moving from the prior to posterior simulations.

Figure 13, Figure 14 and Figure 15 show a map of the yearly average anthropogenic, biospheric and background fluxes and associated λ values with which the prior fluxes (i.e., the state vector elements) are multiplied. It is quite clear to see that the λ values are determined at the observation station locations and decay away from there, such that the final map of λ values is something like an interpolated map between the values at various station locations. For the background values of λ, we note that the mean value is approximately 0.99 which on a background of about 400 ppm corresponds to a 4 ppm difference. Indeed, we can see that when averaged over a full year, the surface-level of the simulation has changed by about -2 ppm after optimization compared to the prior simulation run. Clearly, the CAMS product (optimised with *in situ* data) needs to be adjusted further downwards when optimised with the *in situ* data used in our inversion setup.

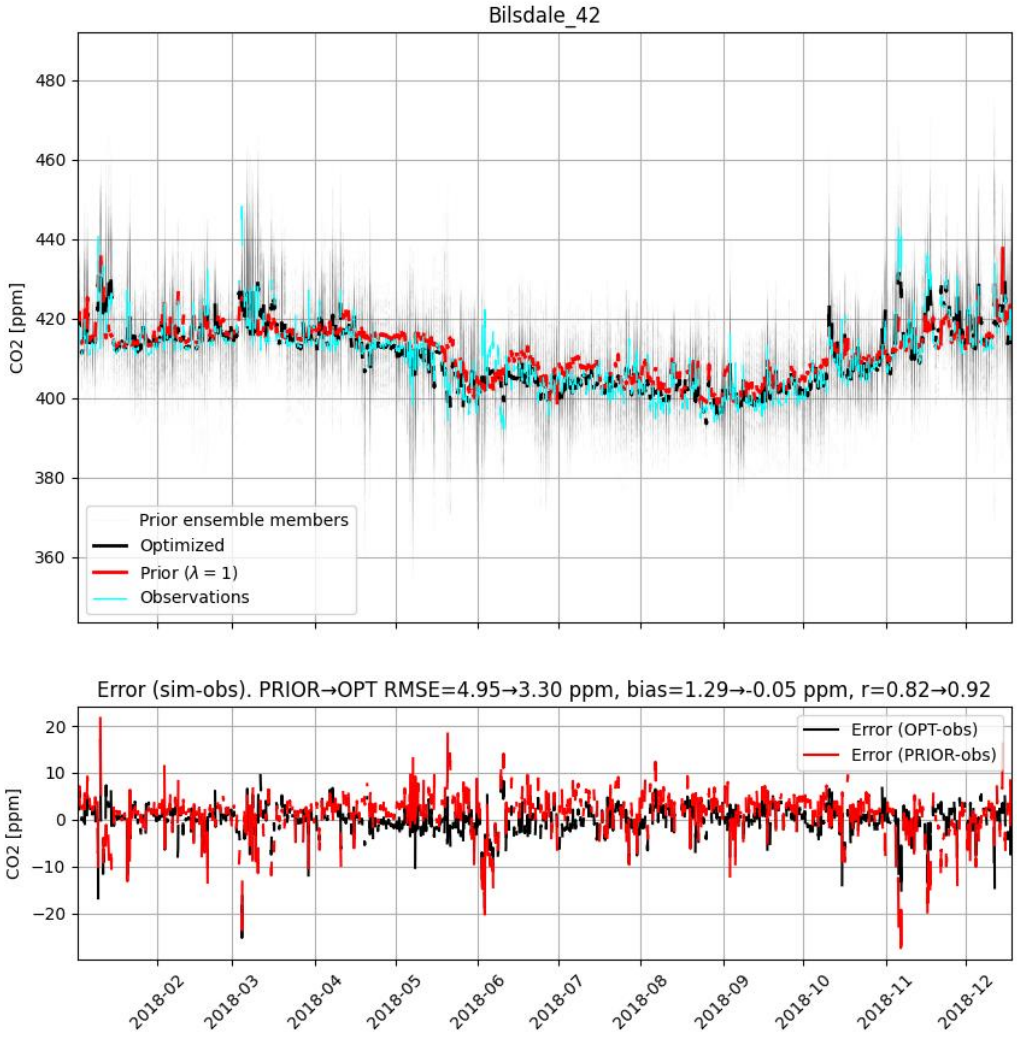


Figure 10: Example of simulated timeseries for the station Bilsdale (UK) with ICON-ART, when assimilating ICOS data only. Note the discrepancy between simulations and observations in early June that is not well captured in the inversions (just right of the 2018-06 line).

Figure 11: Inversion results with ICON-ART when assimilating ICOS station data only. Note that the full range of ICOS OBSPACK stations in the domain is shown here, at multiple inlet heights if applicable, this range is larger than the set of stations used for inverting the data.

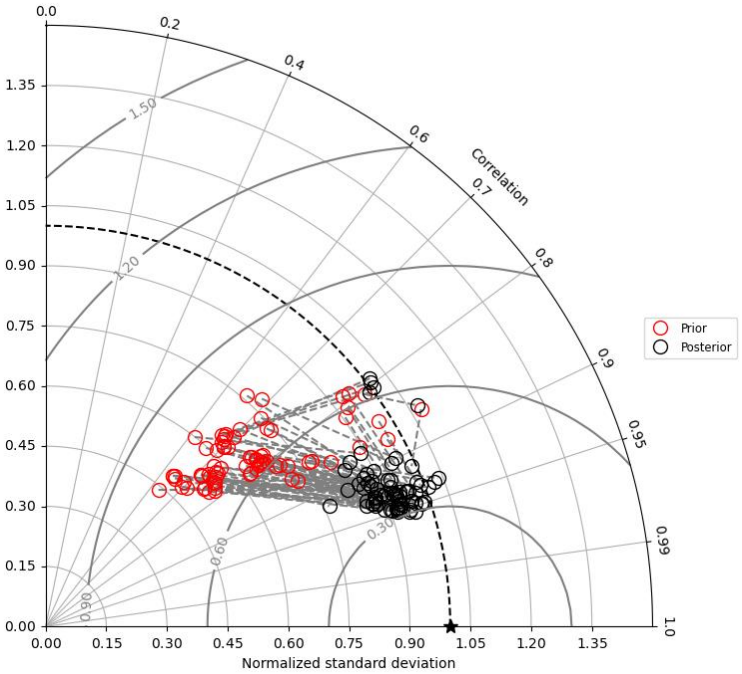


Figure 12: Taylor diagram for the inversion results with ICON-ART assimilating ICOS station data only; results improve when moving towards the star on the bottom horizontal axis. Each dot represents the performance at an ICOS station for a year of data (i.e., it represents the same information as the previous figure).

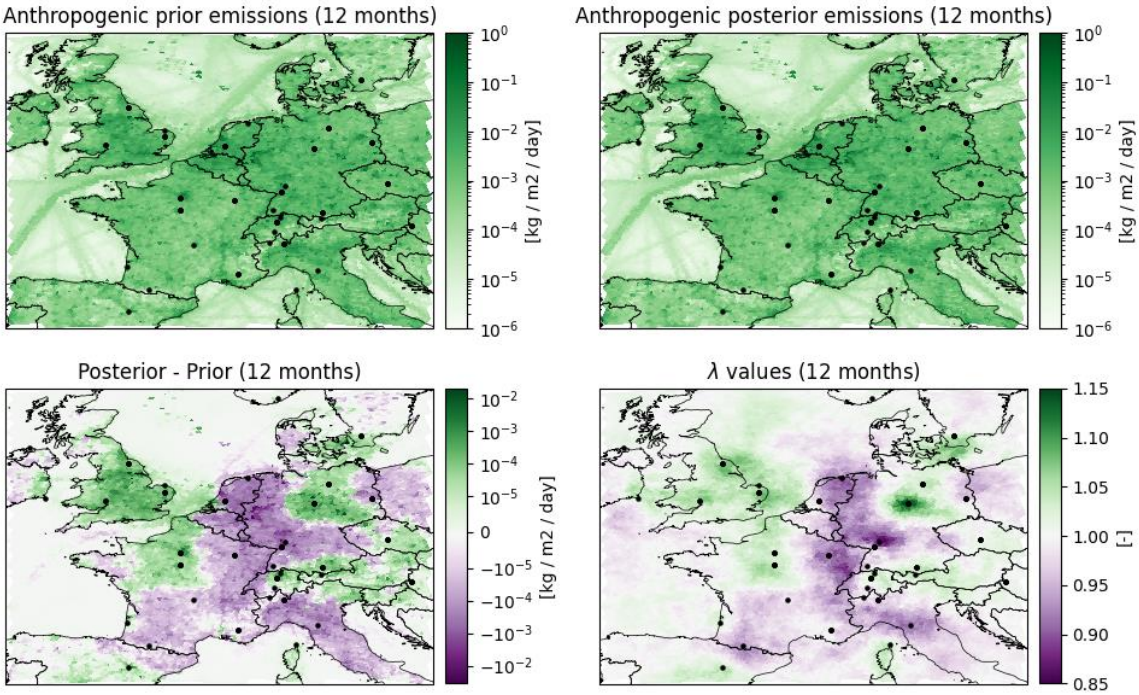


Figure 13: Full-year inversion results with ICON-ART assimilating ICOS station data only, for the anthropogenic component. Black dots on the map correspond to ICOS stations.

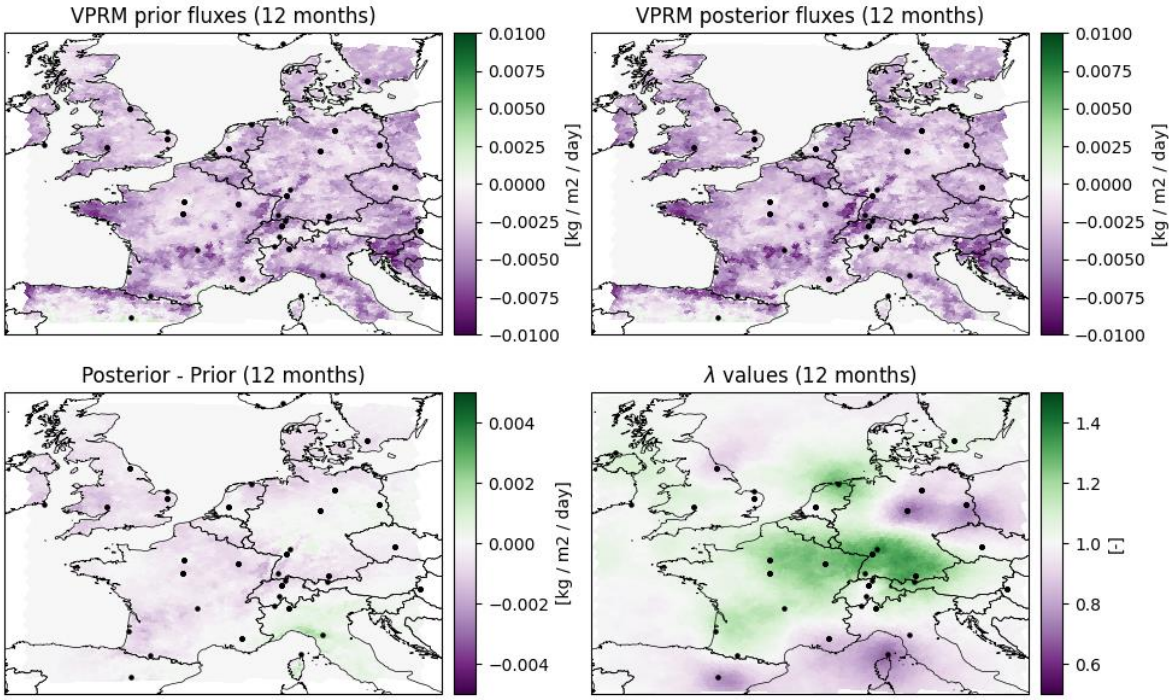


Figure 14: Full-year inversion results with ICON-ART assimilating ICOS station data only, for the biospheric component. Black dots on the map correspond to ICOS stations.

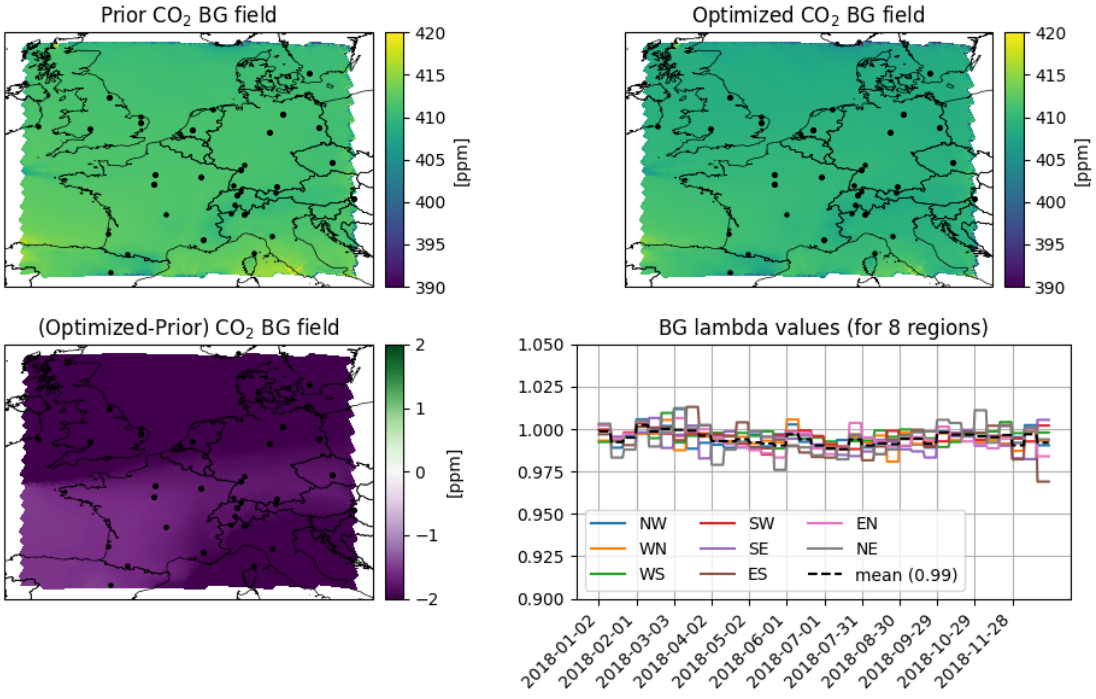


Figure 15: Full-year inversion results with ICON-ART assimilating ICOS station data only, for the average background component. Inflow regions are denoted following major side - minor half side (e.g., 'NW' means from the top, on the left half, while 'WN' means from the left, on the top half).

3.3.2.2 Assimilating both ICOS station and OCO-2 data

Figure 16 shows the timeseries of hourly CO_2 dry mole fractions at station 'Bilsdale' at 42 m altitude (https://meta.icos-cp.eu/objects/lVo3gLCopZerMjEogqkRAns0), subsampled to the assimilated times (12-17 UTC for normal stations, 23-5 UTC for mountain stations) for real observations, a prior simulation and the posterior simulation ('optimised'). We can see that posterior simulation (in black) moves the simulated values closer to the observed values compared to the prior. This result is also quantified in the error plot below the timeseries, with the root mean square error (RMSE) and bias both decreasing, and the Pearson product-moment correlation coefficient ('r') rising. However, compared to the inversion with only ICOS stations, the improvement in RMSE and bias is a lot more modest. Clearly, a positive bias remains when moving from prior to posterior data.

Figure 17 summarises the quantitative results now for all ICOS stations in a bar plot. We can generally see similar trends as for Bilsdale, with RMSE values dropping, and R-squared ('r^2') values rising. Exceptions are again the stations Křešín u Pacova and La Muela, which do not improve at all upon doing the inversion in terms of RMSE scores. What is notable, is that nearly all stations end up with a considerable positive bias after doing the optimization. This is contrary to the inversion with only ICOS stations – clearly now that ICOS stations and satellite data are available, the fit to the ICOS stations is not as good as it was before. The results point at systematic differences between the model-observation mismatches when comparing the simulations with ground-based and with satellite observations. This could originate from systematic differences between satellite and in-situ observations, but it could also hint at an incorrect representation of the vertical distribution of CO_2 in ICON-ART (which is closely linked to the vertical distribution of CO_2 in CAMS). Figure 18 shows a Taylor diagram summing up the results for all the stations, similarly showing and suggesting an improved performance when moving from the prior to posterior simulations; but a lesser improvement than was found with the ICOS-stations-only inversion.

Figure 19 presents one example of a full 10-day cycle of OCO-2 column observations, as compared to the prior and posterior simulation runs. Only a very limited number of stripes are actually present even for a 10-day cycle. It is clear to see that, prior to the optimization, the spread in the simulated values does not correspond well to the spread in the observed columns, with a clear positive bias (the simulated column being larger than the observed column). After doing the inversion, the posterior is better able to capture the spread in the columns, with a clear improvement in performance. Figure 20 repeats this plot for all 10-day cycles, to give (1) an overview of the total coverage of OCO-2 data points for a full year and to give (2) an idea of the over-all reduced model-data-mismatch which appears to be around 2 to 2.5 ppm nominally after optimization, which corresponds to the model data mismatch defined for the OCO-2 data. Furthermore, the observations drawn in the description of the previous figure hold here too: the spread in the error is reduced, and a positive bias is eliminated. Figure 21 summarises the misfit of the prior and posterior simulations for all 10-day periods that were considered. In the violin plots we record the overall distribution of misfits, which is typically normally distributed *but* with a small positive bias in the prior that is moved towards zero in the posterior simulations. The average RMSE value and bias over each 10-day period is reduced; the Pearson product-moment correlation coefficient is increased in all-but-one cases. We can, moreover, see that the number of observations is not spread evenly over the full year – winter months have considerably less observations, likely due to cloud cover and higher solar zenith angles.

Figure 22, Figure 23, and Figure 24 show a map of the yearly average anthropogenic, biospheric and background fluxes and associated λ values with which the prior fluxes (i.e., the state vector elements) are multiplied. Compared to the inversion with ICOS stations only, the map of λ values is not like an interpolated map between the values at various ICOS station locations. Instead, it does not seem to follow a particular spatial distribution at all. Moreover, compared to the ICOS-station-only inversion, the updates for both the anthropogenic and biospheric fluxes are considerably larger. For example, in the French Burgundy region, there

is a net release of CO_2 following the online VPRM calculations, whereas in the previous inversion there was purely uptake of CO_2. And, we see for example that the shipping fluxes are now modified where they weren't before, as we now have access to data over sea. For the background values of λ, we note that the mean value is approximately 1.00 instead of 0.99. We can see over a full year average that the surface-level of the simulation has changed about -1 ppm after optimization compared to the prior simulation run. Clearly, the CAMS product (optimised with in-situ data) needs to be adjusted downwards when optimised with our in-situ and satellite data and our inversion set-up; but less so than when using only in-situ data. This also explains why a positive bias remains in the bar plot of the in-situ data (Figure 17), as this inversion simply corrected the background down a bit less than the previous inversion. However, we see a lot more variation in the background field in this inversion compared to the last one, with larger updates to the scaling factor in every 10-day cycle.

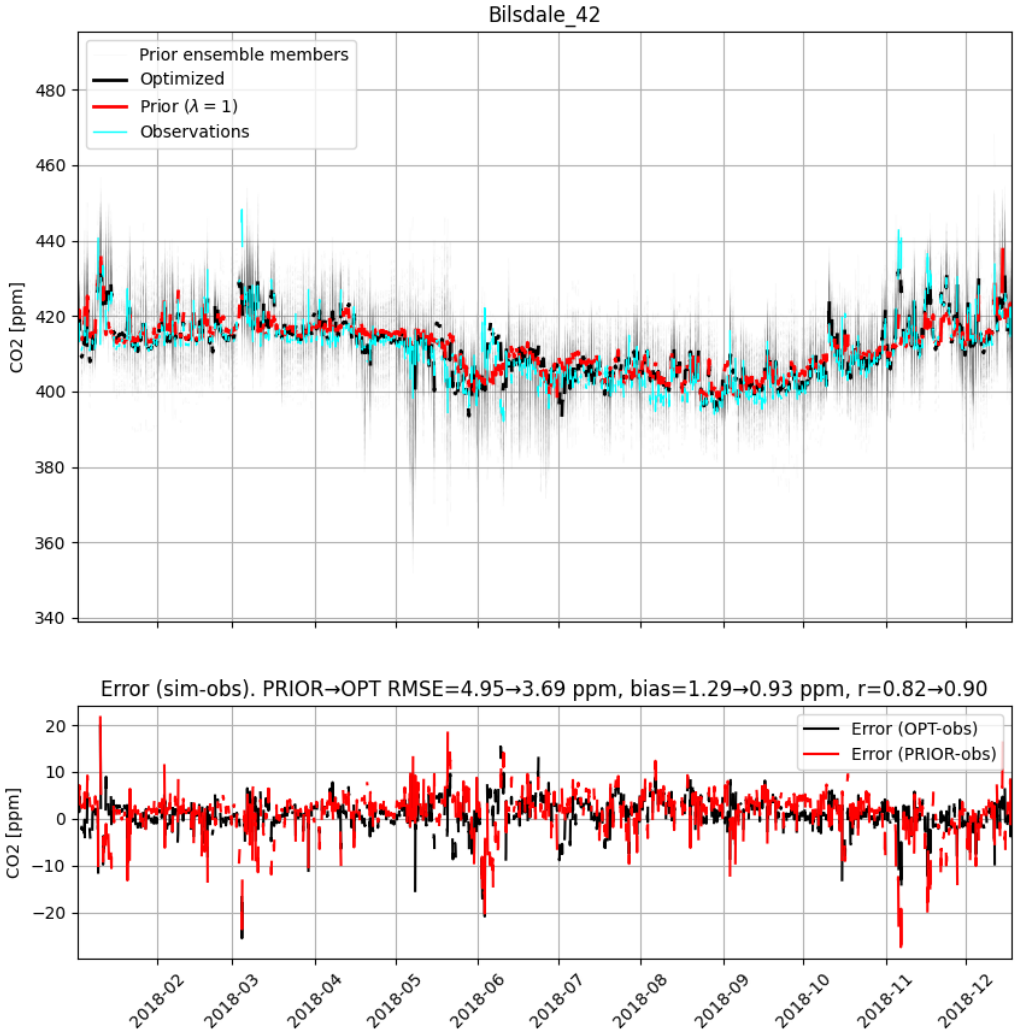


Figure 16: Example of simulated time-series for station Bilsdale (UK) with ICON-ART, when assimilating ICOS and OCO-2 data. Note how the satellite data improves the fit between simulations and observations in early June compared to the inversion with ICOS data only (just right of the 2018-06 line).

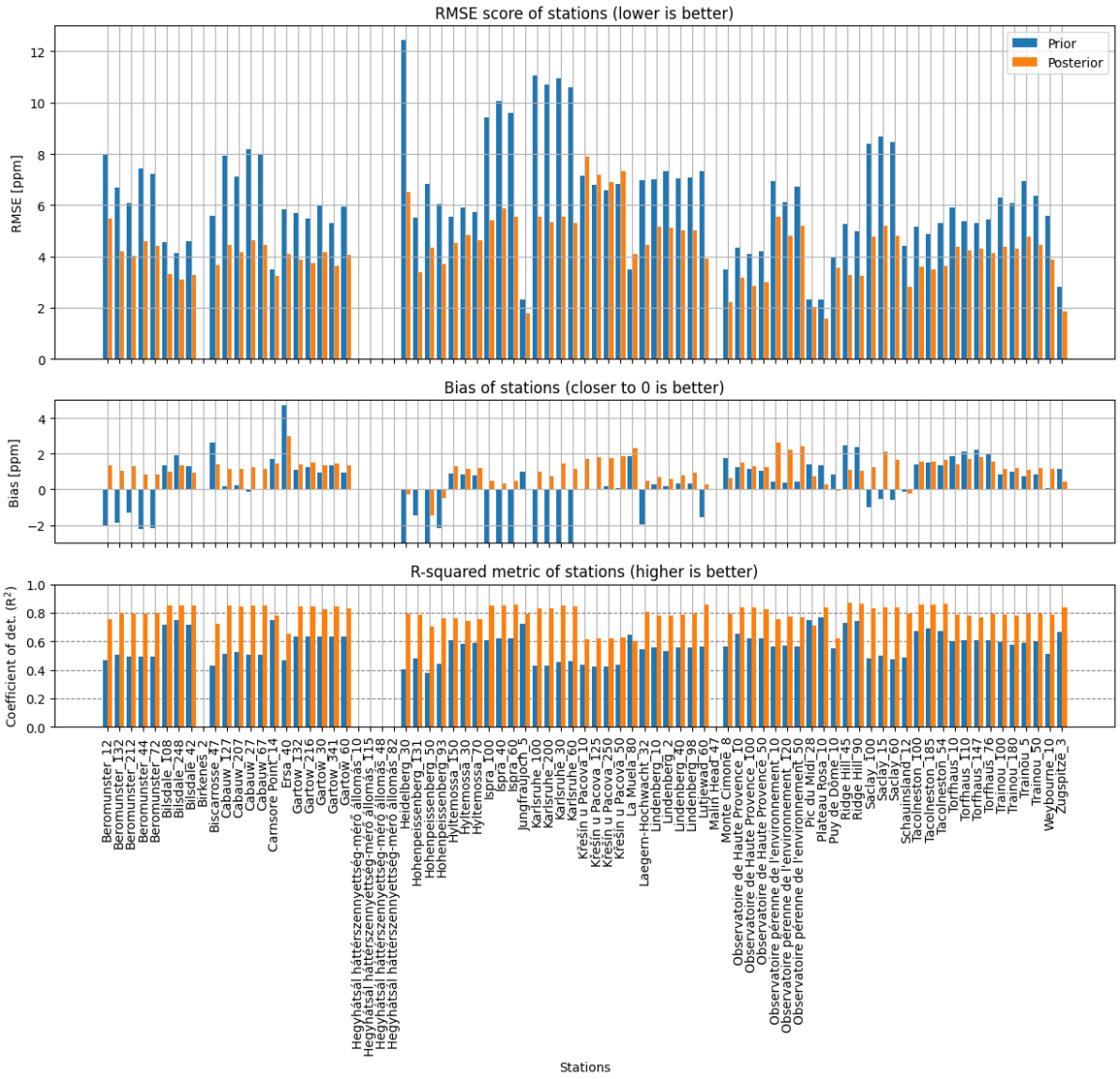


Figure 17: Inversion results with ICON-ART when assimilating ICOS station and OCO-2 data. Note that the full range of ICOS OBSPACK stations in the domain is shown here, at multiple inlet heights if applicable, this range is larger than the set of stations used for inverting the data.

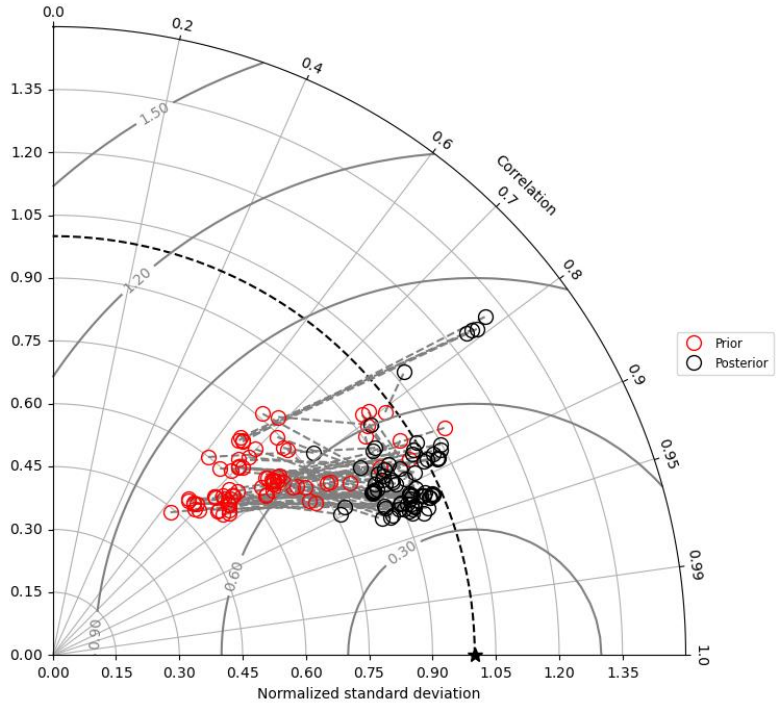


Figure 18: Taylor diagram for the inversion results with ICON-ART assimilating ICOS and OCO-2 data; results improve when moving towards the star on the bottom horizontal axis. Each dot represents the performance on an ICOS station for a year of data (i.e., it represents the same information as the previous figure).

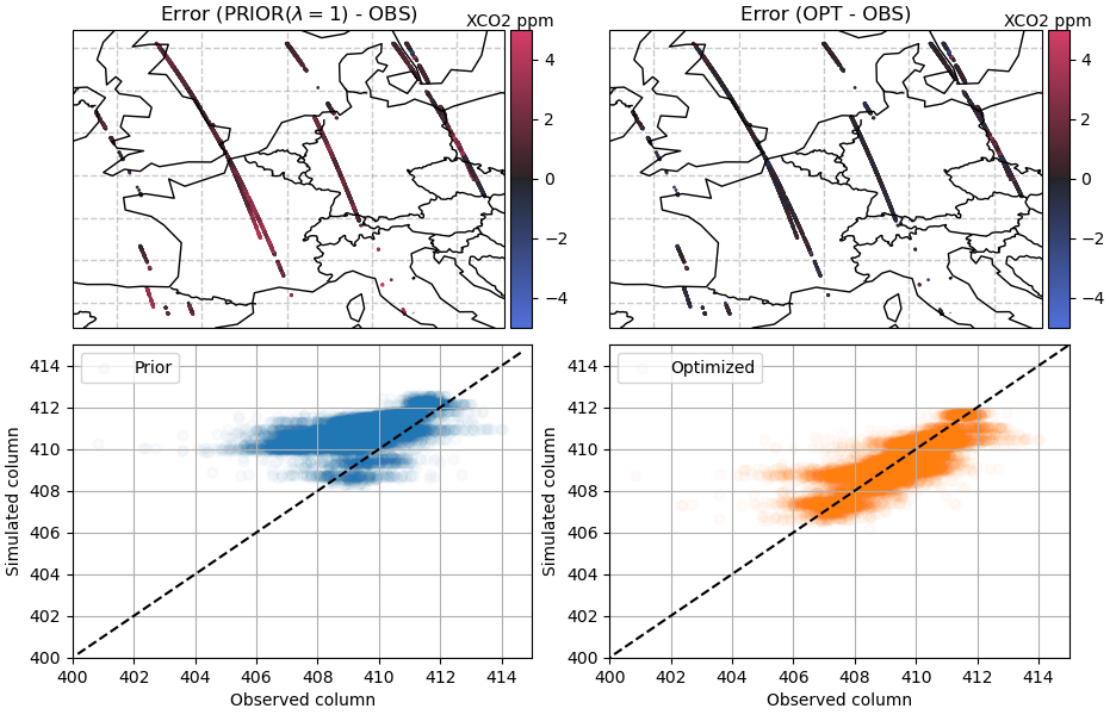


Figure 19: One example of OCO-2 observations made during a complete 10-day cycle starting at 2018-05-02, for which observations are compared to ICON-ART prior (left) and optimized (right) total CO$_2$ columns. The top set of plots shows the error between simulation and observed data, where a good match would correspond to a black drawn pixel. The bottom set of plots are scatter plots between the simulated and observed data, where a good match would place the point on the dashed black 1:1 line. It is clear that the prior simulated data has a positive bias that is largely eliminated after the optimization.

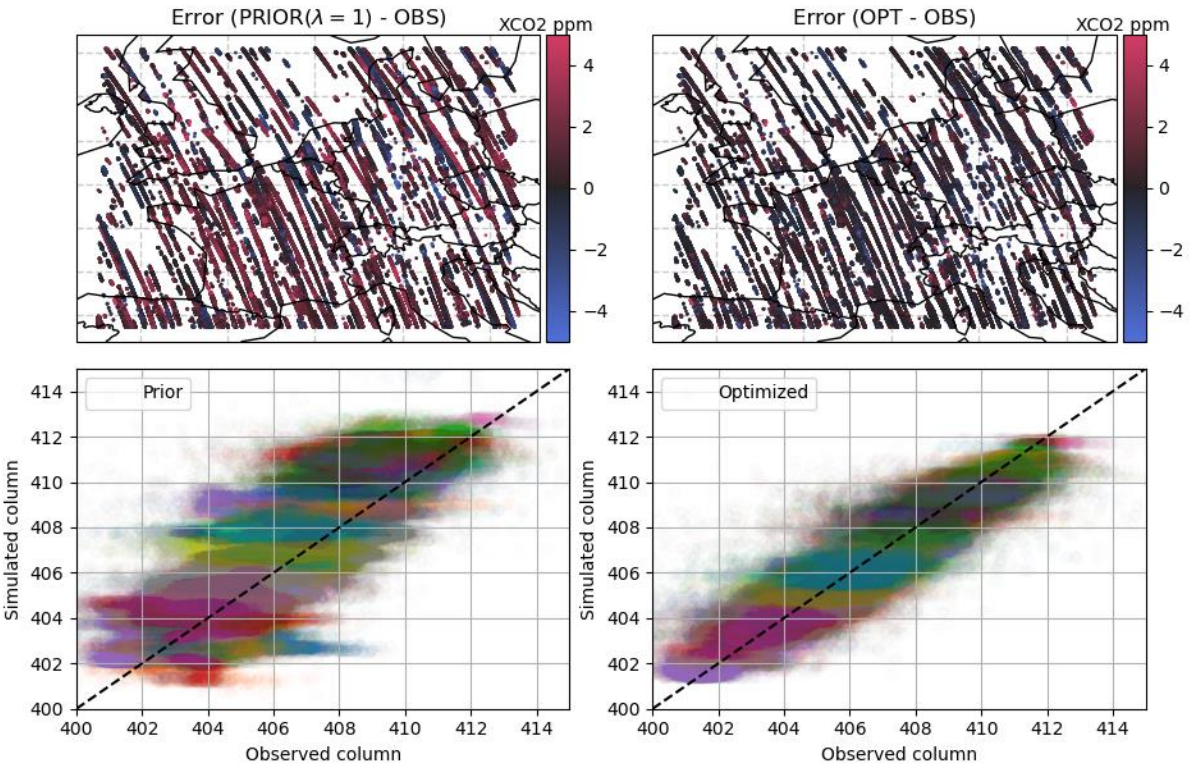


Figure 20: All OCO-2 observations made during all 10-day cycles, for which observations are compared to ICON-ART prior (left) and optimised (right) total CO_2 columns. The top set of plots shows the error between simulation and observed data, where a good match would correspond to a black dot. The bottom set of plots are scatter plots between the simulated and observed data, where a good match would place the point on the dashed black 1:1 line. Every 10-day cycle is given a unique colour. It is clear that the prior simulated data has a positive bias and large spread, while after optimization the values lie closer to the 1:1 line.

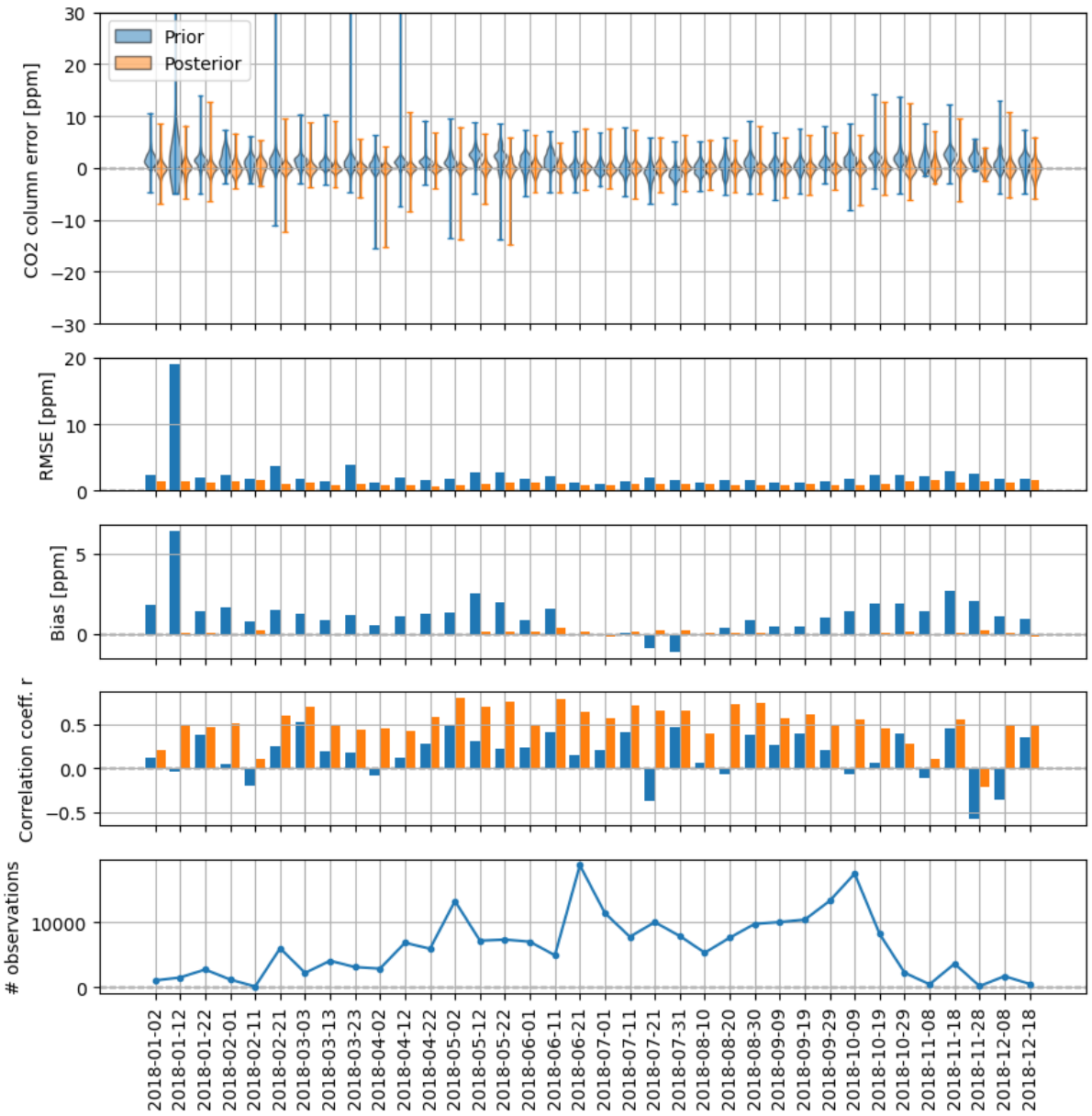


Figure 21: OCO-2 misfits collected for each assimilation cycle with ICON-ART given ICOS and OCO-2 data (the start of each 10-day cycle is indicated on the horizontal axis). On the top we show violin plots of the prior and posterior error, on the other plots we show the RMSE, bias, correlation, and number of OCO-2 observations assimilated per cycle. It is clear that the column bias is virtually eliminated as a result of the data assimilation. Note how more observations are available over the summer months compared to the winter months. Note that, in this figure, the correlation coefficient is not related to what we saw for the ICOS stations (which showed the temporal correlation for a single station) but simply the correlation over all total columns for the assimilated 10-day period.

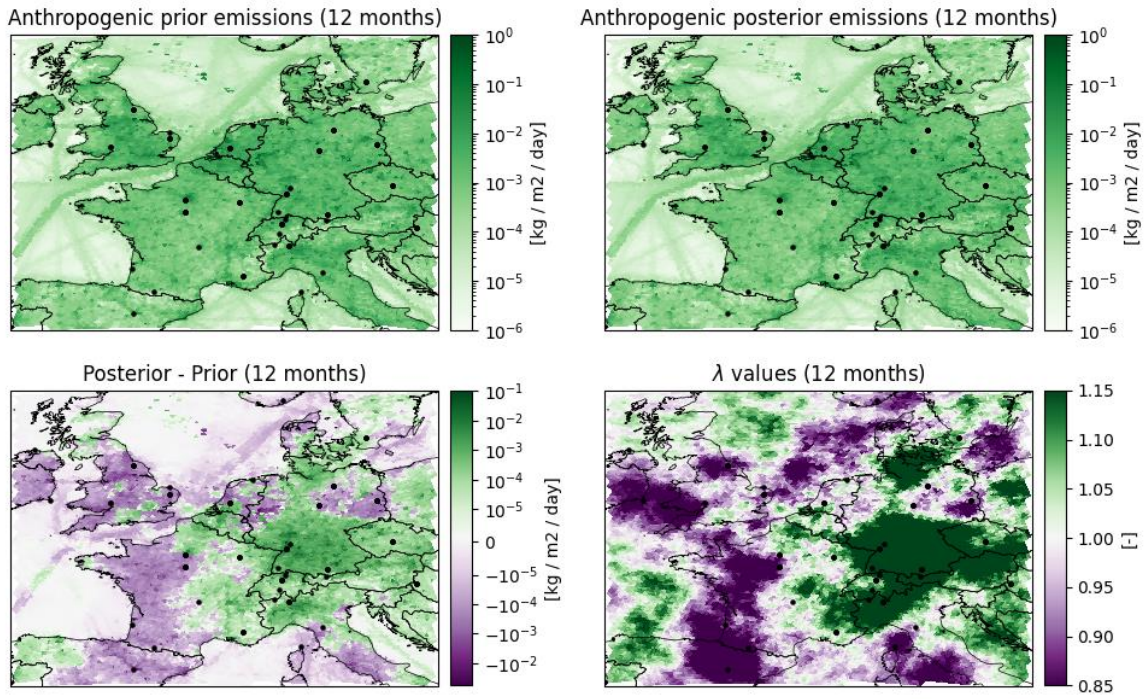


Figure 22: Full-year inversion results with ICON-ART assimilating ICOS station and OCO-2 data, for the anthropogenic component. Black dots on the map correspond to ICOS stations.

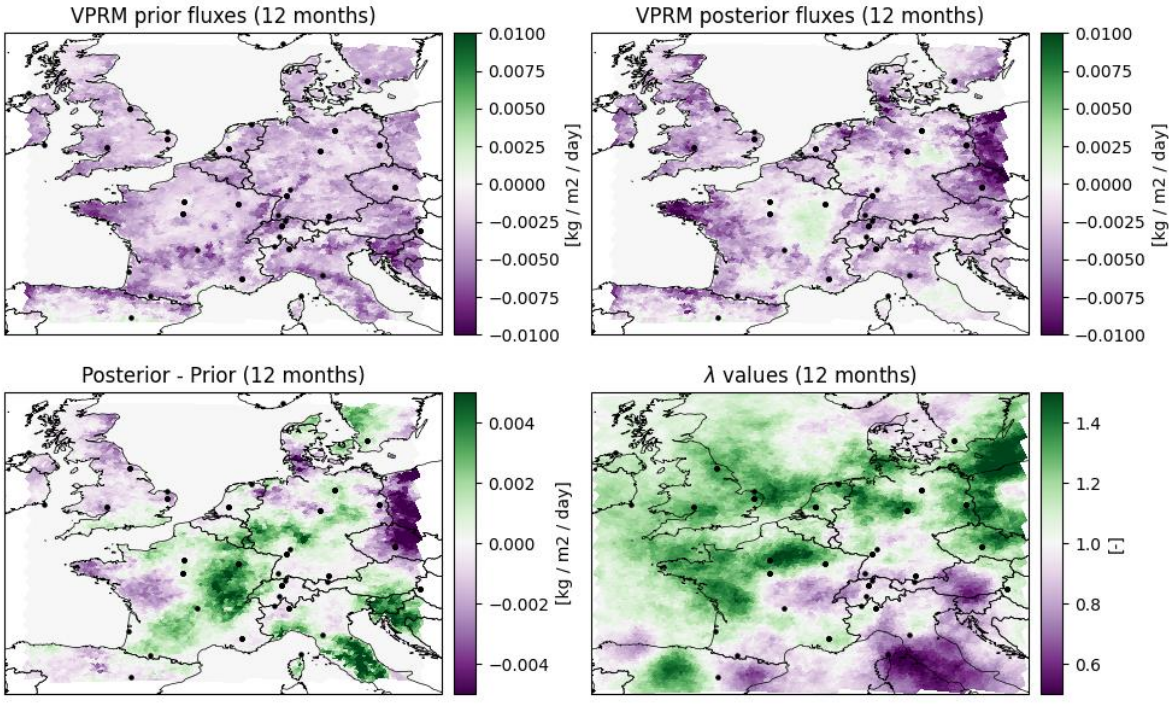


Figure 23: Full-year inversion results with ICON-ART assimilating ICOS station and OCO-2 data, for the biospheric component. Black dots on the map correspond to ICOS stations. Note that some areas (e.g., the Burgundy region in France) exhibit yearly net positive vegetation fluxes after the inversion, unlike the case where we worked with just the ICOS stations.

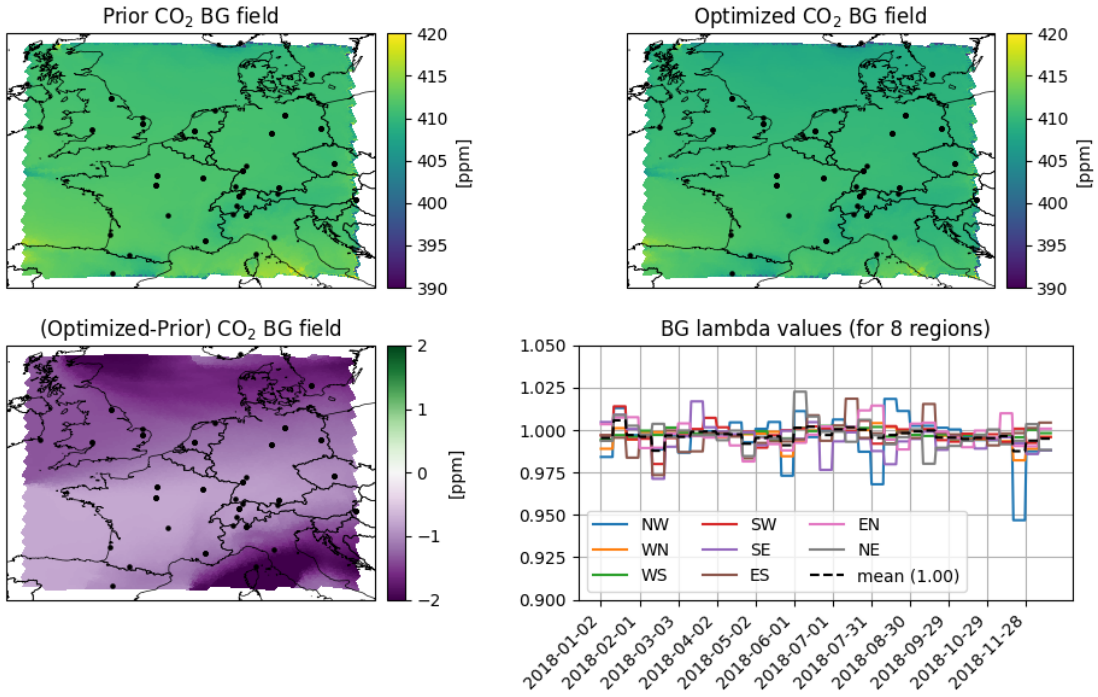


Figure 24: Full-year inversion results with ICON-ART assimilating ICOS station and OCO-2 data, for the average background component. Inflow regions are denoted following major side - minor half side (e.g., 'NW' means from the top, on the left half, while 'WN' means from the left, on the top half).

3.3.2.3 Gathering country-total results

After carrying out the inversions with only in-situ ICOS stations and with in-situ ICOS stations plus OCO-2 total column observations, we can compute yearly country-scale emissions. Those results are given in Table 5 and graphically displayed in the bar plot of Figure 25. The country-total emissions are computed by multiplying the prior emissions with the estimated λ values. The uncertainty for each country-total result is computed through $\sqrt{\frac{\sum_{n=1}^{36} g^T C_n g}{36^2}}$, where C_n is the posterior covariance matrix of the state vector for lag n, and g is a vector where each element corresponds to the proportion for which each cell belongs to a certain country.

Table 5: The country-total inversion results with ICON-ART, in Mt CO_2/yr. The value after ± is a very rough estimate of the uncertainty. ANTH stands for "anthropogenic emissions", while NAT stands for "biogenic" emissions which are the VPRM computations plus forest fire fluxes, lateral fluxes and (positive) ocean fluxes.

Country	Prior ANTH	POST ANTH ICOS	POST ANTH ICOS+OCO2	PRIOR NAT	POST NAT ICOS	POST NAT ICOS+OCO2
Austria	88.2±7.3	89.6±6.9	104.5±6.6	-77.7±20	-81.8±17.7	-62.5±12.6
Belgium	109.3±10.6	107±9.7	115.1±7.8	-11.2±10.7	-16.9±7.7	-1.6±6.9
Croatia	17.5±1.5	17.6±1.5	19.2±1.5	-44±14.3	-42.2±13.4	-15.6±11.4
Czechia	96±8.4	96.7±7.7	110.9±7.3	-53.3±21.2	-56.8±18.2	-112.5±13.2
Denmark	44.5±3.8	45±3.5	46±3.6	-15.3±12.5	-23.3±11.5	-14.8±7.9
France	385.8±20.6	383.2±17.6	390.2±17.0	-452.2±138.6	-537.3±103.3	-268±66.6
Germany	855±54.2	835.3±46.6	960.1±38.7	-216.3±89.3	-250.4±60.6	-164.3±41.4
Netherlands	156.6±14.5	152.5±13.0	157.6±10.2	9.6±9.8	4.6±6.8	8.2±6.2
Slovenia	18.2±1.8	18.4±1.7	19±1.7	-40.7±10.5	-39.3±9.4	-16.9±7.6
Switzerland	44.2±4.2	44.6±3.3	49.7±3.2	-33.2±9.1	-38.8±6.2	-26.9±4.0
UK	376.8±27.2	390.1±24.1	353.1±23.0	-140.6±53.8	-187.9±43.6	-188.7±31.6

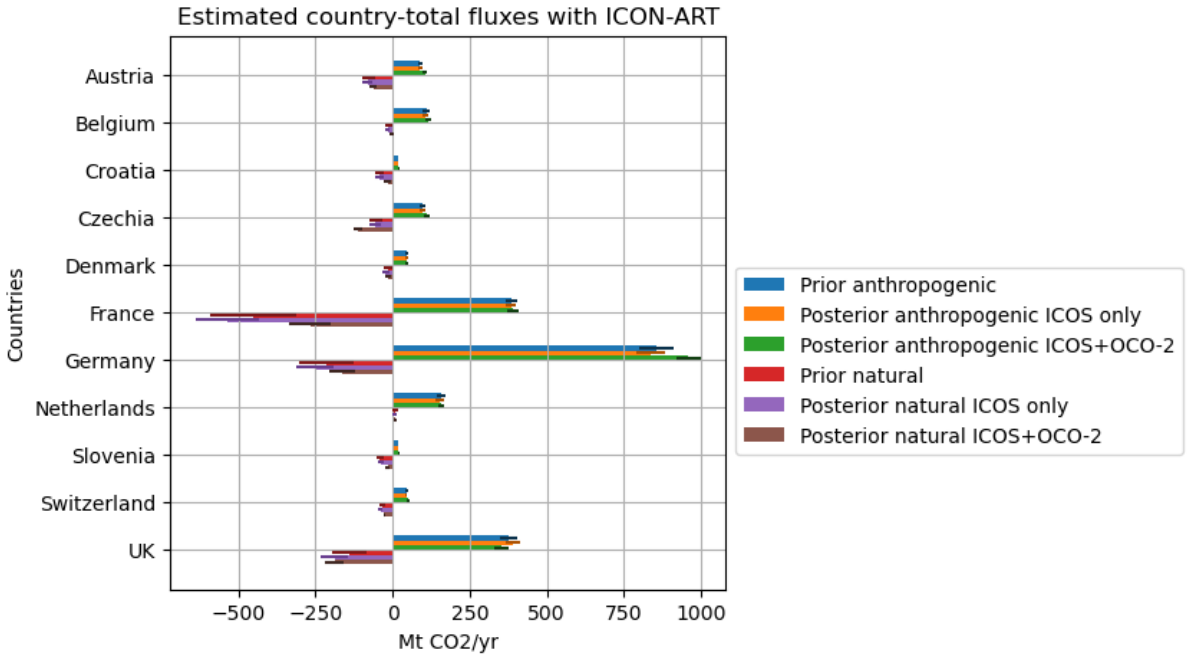


Figure 25: Bar-plot of country-total inversion results using the ICON-ART CTDAS setup, in Mt CO_2/yr. The lines in black are a rough estimate of the standard deviation.

3.3.2.4 Discussion

We note a few points of discussion below.

- The biospheric uptake as modeled here is too strong and does not seem accurate. The VPRM formulation implemented online in ICON-ART (in fact any version of VPRM) is not made to be neutral on a yearly basis (which is what one would, roughly, expect). Some comparisons with other models were performed that also simulate the net ecosystem exchange. The performance of the VPRM was rather similar to other models when looking at *instantaneous* or *daily* flux values (e.g., in comparison with SiB4 fluxes from the Carbon Tracker Europe which is close to neutral on a yearly basis by design, https://www.icos-cp.eu/data-products/high-resolution-near-real-time-co2-fluxes-over-europe-carbon-tracker-europe-2017-2023). However, the VPRM fluxes have a small bias in the form of too little respiration, and this bias becomes large when summing the emissions over a large area and over a full year, leading to a large seeming uptake of CO_2 where little to none would actually be expected. For example, over a yearly basis one can find a difference to a standard VPRM product for Europe (https://www.icos-cp.eu/data-products/biosphere-atmosphere-exchange-fluxes-co2-vegetation-photosynthesis-and-respiration) as shown in Figure 26. What is clear is that the VPRM line (in blue) has a very similar order of magnitude, but too little respiration both in summer and winter days (or, too much gross photosynthetic production). But whereas the SiB4 product has a yearly flux over Europe that sums to about -0.0055 PgC/yr, the VPRM product has a yearly flux that sums to -1.4871 PgC/yr. Hence, the small differences add up to a large number.
- The fact that the VPRM product leads to a large over-estimation of the CO_2 uptake is not really corrected for by the inversions, although the inversion with the OCO-2 data appears to generally lead to less uptake. Still, it appears that (1) the uncertainties set to the prior VPRM fluxes are much too conservative, (2) it appears that the standard VPRM product is capable of producing realistic instantaneous fluxes but not too realistic yearly fluxes, (3) the multiplicative λ values on the net ecosystem are not capable of fixing this error (as increasing the nighttime respiration would require increasing the λ value, but this would also increase the daytime photosynthetic production), such that a better way

forward is perhaps to split respiration and photosynthesis parts into separate components of the state vector, and (4) perhaps the fact that we mostly have daytime observations both in the in-situ and total column data means we are somewhat biased to emissions during the afternoon hours, where the photosynthetic part generally dominates.

- In general, we do not observe a 'dipole'-like effect, where updates in the anthropogenic and biogenic fluxes are counteracting or compensating for each other – we generally observe that both anthropogenic and biogenic fluxes are moving similar directions, which builds some confidence in the system.
- Both inversion results require us to decrease the background CO_2 field. As the background field comes from the CAMS "inversion-optimised" product which was optimised with in-situ data, the surprise is that the update to the background field is largest in the optimization with only in-situ data. We have not looked into the reasons behind this feature in the data. It suggests, however, that there is some inconsistency between the OCO-2 and in-situ data.

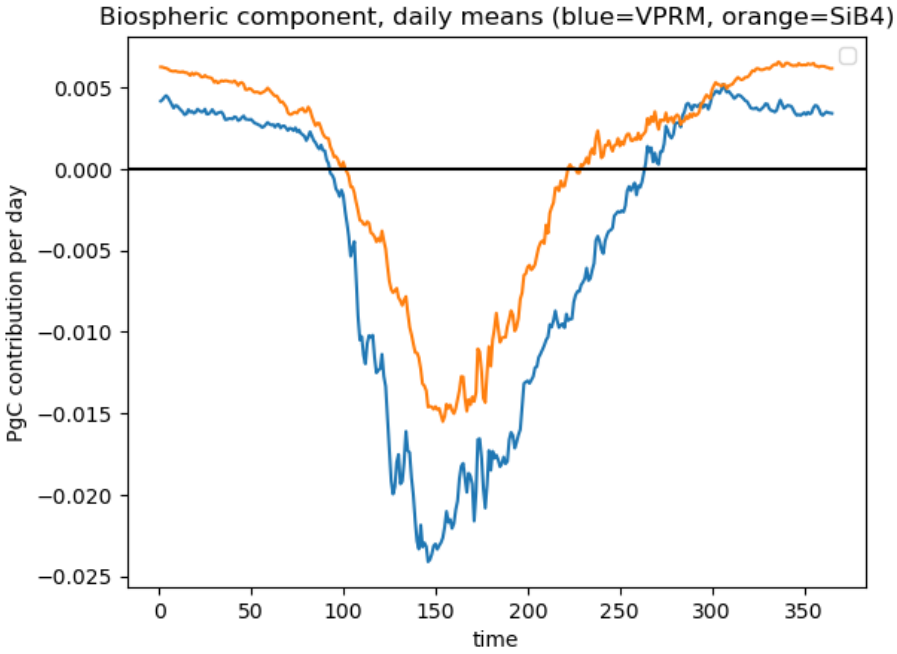


Figure 26: Plot of the daily biospheric carbon (C) flux summed over Europe for days in 2018, with the horizontal axis representing the day of the year, comparing ICOS-SiB4 (https://www.icos-cp.eu/data-products/high-resolution-near-real-time-co2-fluxes-over-europe-carbon-tracker-europe-2017-2023) to ICOS-VPRM (https://www.icos-cp.eu/data-products/biosphere-atmosphere-exchange-fluxes-co2-vegetation-photosynthesis-and-respiration).

3.3.3 Methane inversion description

In addition to the CO_2 inversion described above, we shortly summarise results of a European methane (CH_4) inversion performed with ICON-ART by (and documented in greater detail in Steiner et al., 2023). The inversion was performed in the context of the TransCom modelling community intercomparison. Concerning the model domain, the resolution is coarser by a factor of two (i.e., an effective grid spacing of about 26 km) which covers the entirety of Europe and neighbouring regions, resulting in a total of 21 344 cells, roughly the same as for the CO_2 inversion. The meteorological and atmospheric IC/BC come from the ERA-5 and CAMS inversion-optimised products (same as for the CO_2 inversion). Two categories are optimised separately: anthropogenic and natural fluxes; in addition to that, 8 background components

are optimised for the same 8 inflow regions as for the CO_2 inversion. The prior anthropogenic fluxes (agriculture, waste, and fossil fuels) come from the EDGAR v6.0 inventory. The natural fluxes correspond to peatland and mineral soils, inland water, termites, ocean, biofuels and biomass burning and geological emissions. Fluxes are considered constant for each 10-day window that is assimilated, and emitted in the lowest model level (0 to 20 m altitude). Like the CTDAS set-up for CO_2, we consider overlapping cycles of 2 10-day windows. We use 100% uncertainty for the fluxes in the prior error covariance matrix, and an exponentially decaying correlation length scale of 200 km, and generate the 192 ensemble members for the anthropogenic and natural emissions, while using an uncertainty of 0.05% uncertainty for the background (corresponding to roughly 1 ppb CH_4). The observation stations are ICOS stations with the addition of 10 flask sample locations. The model-data mismatch (i.e. the model plus observation error) is set to 10 ppb + 30% of the yearly mean signal from the anthropogenic and natural emissions (as modelled with ICON-ART).

The results of the 2018 inversion are shown in Figure 27. The anthropogenic emissions show a strong upward correction in northwestern Europe of up to 25 mg/m^2/day in the Benelux countries, and more moderate upward corrections in northwestern France and southern England. In terms of annual means, natural fluxes are corrected downward almost everywhere, particularly over Italy (-22%). The national scale inversion results for 2018 are shown in Table 6, the results for 2021 are shown in Table 7, and both are plotted in Figure 28. The trends for the 2021 inversion are fairly similar to those found for 2018. The downward correction of the natural fluxes in Italy (very pronounced in the 2021 results) is most likely due to the very high a priori geological emissions in these regions (it is clearly visible in Figure 27 that the prior emissions are markedly stronger for Italy and Romania compared to other European countries). We refer again to Steiner et al. (2023), where the CTDAS set-up is validated with further synthetic tests, and results are presented (also for the years 2008 and 2013) in greater detail.

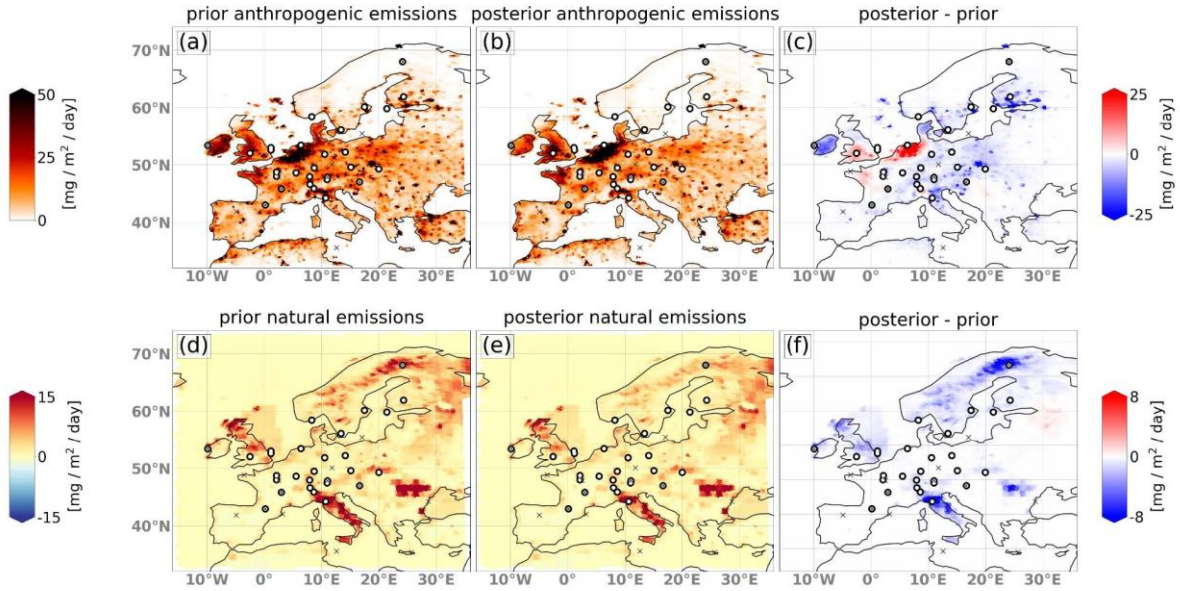


Figure 27: Yearly (prior and posterior) CH_4 emissions over Europe as derived with ICON-ART with ICOS in-situ and flask sampling stations. Figure taken from Steiner et al. (2023).

Table 6: The country-total inversion results for 2018 with ICON-ART, in Tg CH_4/yr. The value after ± is a very rough estimate of the uncertainty. ANTH stands for "anthropogenic emissions", while NAT stands for natural fluxes. The results are described in more detail in Steiner et al. (2023).

Country	Prior ANTH	POST ANTH	PRIOR NAT	POST NAT
Austria	0.39±0.05	0.26±0.02	0.09±0.04	0.08±0.03
Belgium	0.65±0.07	0.71±0.04	0.02±0.01	0.02±0.01
Bulgaria	0.33±0.05	0.31±0.05	0.09±0.05	0.09±0.05
Switzerland	0.19±0.03	0.13±0.01	0.05±0.02	0.04±0.02
Cyprus	0.02±0.01	0.02±0.01	0.00±0.00	0.00±0.00
Czechia	0.49±0.06	0.35±0.03	0.06±0.03	0.05±0.03
Germany	2.43±0.24	2.55±0.10	0.22±0.10	0.20±0.09
Denmark	0.24±0.05	0.15±0.02	0.05±0.04	0.05±0.03
Estonia	0.10±0.02	0.05±0.01	0.04±0.03	0.04±0.03
Greece	0.36±0.08	0.31±0.08	0.05±0.05	0.05±0.05
Spain	1.53±0.20	1.39±0.18	0.03±0.09	0.03±0.09
Finland	0.65±0.10	0.23±0.03	0.42±0.11	0.31±0.08
France	2.50±0.27	2.40±0.13	0.17±0.11	0.16±0.10
Croatia	0.20±0.03	0.15±0.02	0.03±0.02	0.03±0.02
Hungary	0.31±0.05	0.28±0.04	0.04±0.03	0.04±0.03
Ireland	0.66±0.09	0.39±0.05	0.08±0.04	0.05±0.03
Italy	1.35±0.10	1.05±0.05	1.14±0.21	0.84±0.16
Lithuania	0.15±0.04	0.13±0.03	0.02±0.02	0.02±0.02
Luxembourg	0.02±0.00	0.02±0.00	0.00±0.00	0.00±0.00
Latvia	0.08±0.02	0.07±0.02	0.03±0.02	0.03±0.02
Malta	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
Netherlands	0.66±0.07	1.01±0.04	0.04±0.02	0.04±0.02
Poland	2.25±0.22	1.58±0.11	0.26±0.10	0.24±0.09
Portugal	0.32±0.06	0.31±0.06	0.01±0.03	0.01±0.03
Romania	0.77±0.07	0.67±0.06	0.85±0.18	0.73±0.16
Sweden	0.38±0.06	0.13±0.02	0.55±0.14	0.33±0.10
Slovenia	0.13±0.02	0.08±0.01	0.01±0.01	0.01±0.01
Slovakia	0.22±0.03	0.15±0.02	0.04±0.02	0.04±0.02

United Kingdom	1.49±0.14	1.59±0.08	0.50±0.16	0.40±0.13

Table 7: The country-total inversion results for 2021 with ICON-ART, in Tg CH_4/yr. The value after ± is a very rough estimate of the uncertainty. ANTH stands for "anthropogenic emissions", while NAT stands for natural fluxes. The results are described in more detail in Steiner et al. (2023).

Country	Prior ANTH	POST ANTH	PRIOR NAT	POST NAT
Austria	0.39±0.05	0.28±0.02	0.09±0.04	0.07±0.03
Belgium	0.65±0.07	0.59±0.03	0.02±0.01	0.02±0.01
Bulgaria	0.33±0.05	0.25±0.04	0.09±0.05	0.08±0.05
Switzerland	0.19±0.03	0.17±0.01	0.05±0.02	0.04±0.02
Cyprus	0.02±0.01	0.02±0.01	0.00±0.00	0.00±0.00
Czechia	0.49±0.06	0.42±0.03	0.06±0.03	0.05±0.03
Germany	2.45±0.24	2.44±0.08	0.23±0.10	0.19±0.08
Denmark	0.24±0.05	0.16±0.02	0.05±0.04	0.04±0.03
Estonia	0.10±0.02	0.07±0.01	0.05±0.03	0.05±0.03
Greece	0.36±0.08	0.24±0.06	0.05±0.05	0.04±0.05
Spain	1.54±0.22	0.95±0.14	0.01±0.08	0.01±0.07
Finland	0.66±0.09	0.26±0.03	0.43±0.11	0.28±0.08
France	2.51±0.28	2.10±0.11	0.16±0.10	0.14±0.09
Croatia	0.21±0.03	0.18±0.02	0.03±0.02	0.03±0.02
Hungary	0.31±0.05	0.21±0.03	0.04±0.03	0.04±0.03
Ireland	0.66±0.09	0.60±0.05	0.07±0.04	0.05±0.03
Italy	1.36±0.10	0.67±0.05	1.14±0.21	0.34±0.11
Lithuania	0.15±0.04	0.13±0.03	0.02±0.02	0.02±0.02
Luxembourg	0.02±0.00	0.02±0.00	0.00±0.00	0.00±0.00
Latvia	0.08±0.02	0.07±0.02	0.03±0.02	0.03±0.02
Malta	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
Netherlands	0.67±0.07	0.85±0.03	0.04±0.02	0.04±0.02
Poland	2.26±0.22	1.24±0.09	0.27±0.10	0.23±0.09
Portugal	0.32±0.07	0.29±0.06	0.01±0.02	0.01±0.02
Romania	0.77±0.07	0.61±0.05	0.85±0.18	0.74±0.16
Sweden	0.39±0.06	0.19±0.03	0.58±0.14	0.32±0.10

Slovenia	0.13±0.02	0.08±0.01	0.01±0.01	0.00±0.00
Slovakia	0.22±0.03	0.06±0.01	0.05±0.02	0.03±0.02
United Kingdom	1.50±0.15	1.53±0.07	0.48±0.15	0.36±0.12

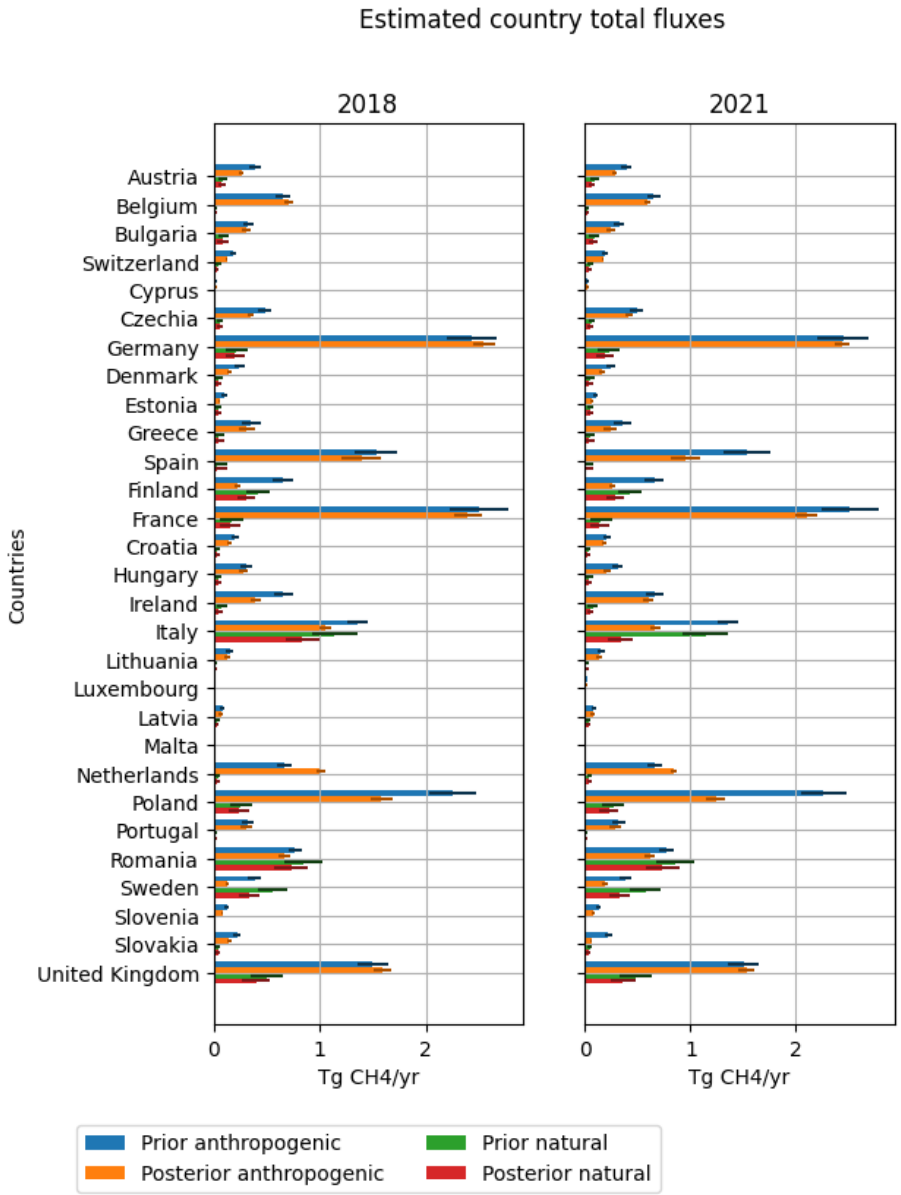


Figure 28: Bar-plot of country-total inversion results using the ICON-ART CTDAS setup, in Tg CH$_4$/yr. The results are described in more detail in Steiner et al. (2023).

3.4 ICON-ART-DWD for CH$_4$ inversions from in situ observations (DWD)

Different aspects of the development of a greenhouse gas (GHG) inversion system at DWD have been pushed forward as part of this project. A general concept described in Section 3.4.1 combines the assimilation of GHG concentrations and the inversion of GHG emissions. This concept serves as a guideline on the path towards a flexible GHG inversion system integrated into the weather prediction service at DWD. As a special case of this concept, a simplified

inversion of the emissions without a simultaneous concentration assimilation has been implemented. This test case is described in Section 3.4.3.

3.4.1 Model description

The modelling of the CH_4 concentrations is carried out using the ICON (ICOsahedral Non-hydrostatic)-ART (Aerosols and Reactive Trace gases) model in the limited area mode. The Copernicus Atmosphere Monitoring Service (CAMS) provides the necessary initial and boundary concentration fields for the simulation setup. The modelled methane concentrations are compared to in-situ observations at the ICOS towers.

3.4.1.1 Transport model

ICON is the numerical weather prediction model used by DWD. It is developed in collaboration with the Max-Planck Institute for Meteorology (MPI-M) and other partners. The ICON model solves the fully compressible nonhydrostatic atmospheric equations of motion.

The discretization of the model equations of ICON is performed on an icosahedral-triangular Arakawa-C grid, with a choice of resolutions obtained by successive refinement of a spherical icosahedron. The mass points are located in the circumcenter of each triangular cell. To optimise runtime efficiency, there is a distinction between so-called fast-physics processes (e.g. cloud microphysics, turbulence), which are calculated at every physics time step, and slow-physics processes (e.g. radiation), which may be called at longer time steps and provide tendencies to the dynamical core that remain constant between two successive calls of the parametrization. Fast-physics processes are treated with operator splitting, which means that they act on an atmospheric state that has already been updated by the dynamical core, horizontal diffusion and the tracer transport scheme.

A mass conserving transport module ART (Aerosols and Reactive Trace gases) is developed in collaboration with KIT Karlsruhe. The tracers in ICON-ART are transported in the same way as ICON tracers, ensuring tracer mass continuity, and it predicts the large-scale redistribution of tracer in the atmosphere due to air motion. Tracer transport is accounted for in a time-split fashion, this means that vertical and horizontal transport is treated separately.

ICON-ART provides an XML interface which adds a-priori emissions from available point and area sources. While the area emission sources have to be stored as mass flux densities in units of kg m^{-2} s^{-1} in a specific directory structure, the strength of point sources has the unit of kg s^{-1}. The ART XML interface allows for the distinction between different transport types (stdchem, stdaero) and properties (chemical, passive, aerosol) of ART tracers. The emission datasets for area sources must be provided on the ICON grid.

While the initial data here denote the state of the atmosphere (meteorological and CH_4 concentration fields) at the start of the model run, the boundary conditions shall denote the data in the lateral boundary zone where the model is forced by the meteorological and CH_4 concentration data outside the domain. We used DWD's operational numerical weather prediction output as meteorological boundary conditions.

The CAMS data used for our initial and boundary conditions for CH_4 concentrations are provided with vertical coordinates on a hybrid sigma-pressure system. Thus, these data need to be interpolated horizontally and vertically to the height-based SLEVE coordinate system used by ICON. A linear interpolation was used to generate hourly data for boundary conditions from the CAMS fields which have a 3h temporal resolution.

The fields for meteorological parameters are initialised daily by using DWD's operational data. The atmospheric CH_4 concentrations are initialised with the CAMS data only for the first day and then propagated using the ICON-ART Limited Area Mode.

To run a simulation setup with hourly output for longer time periods, the Basic Cycling environment (BACY) tool is used.

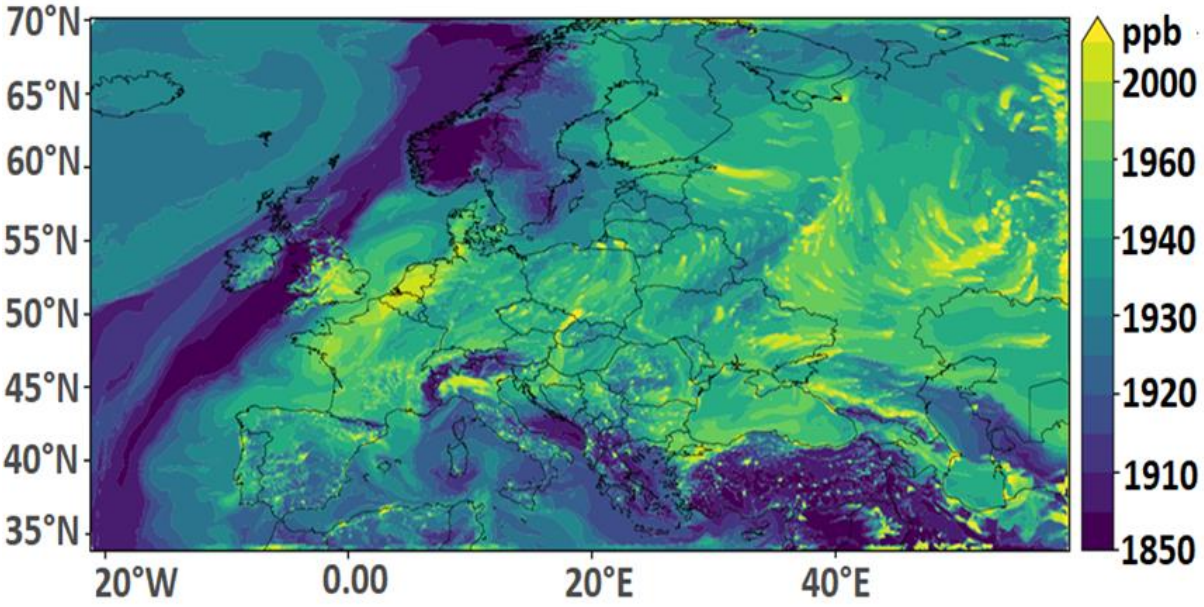


Figure 29: Example CH_4 concentration output of the ICON-ART forward run in limited area mode fed by CAMS boundary conditions and CAMS emissions for the lowest model level.

3.4.1.2 Prior fluxes

We use anthropogenic greenhouse gas emissions provided by Copernicus/TNO (CAMS-REG-GHG_v5_1_emissions_year2018.nc). As they are on a different grid, we have to preprocess these emissions to use them in the ICON model. They need to be mapped onto the ICON triangular grid in a mass-conserving manner. The emissions within a single ICON grid cell are summed up and assigned to the latitude and longitude of the ICON cell midpoint (see Figure 30). For the results presented here, we considered time-constant fluxes.

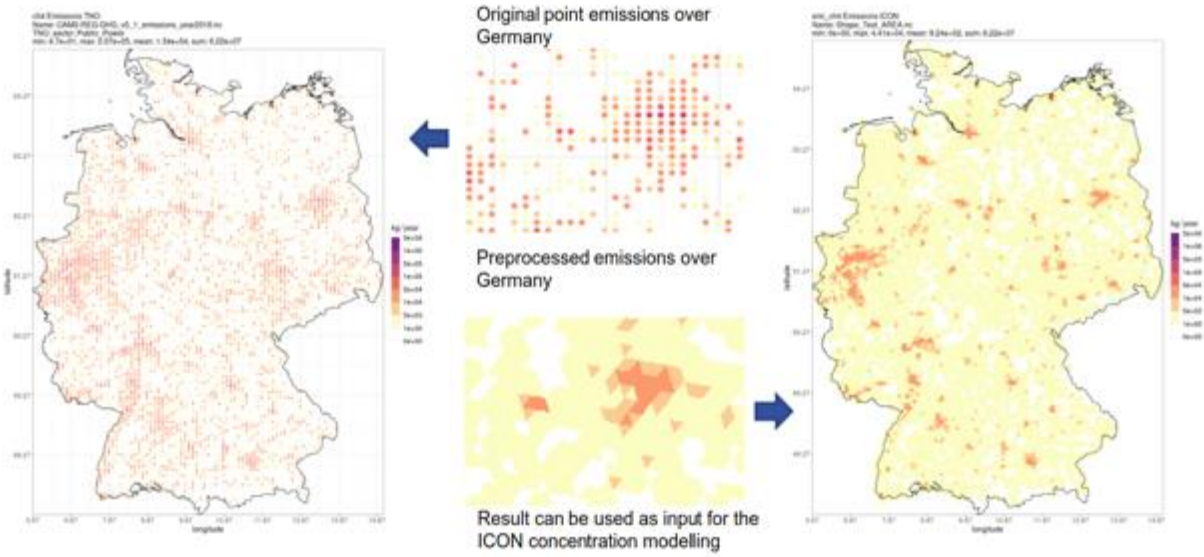


Figure 30: Preprocessing from the TNO grid to the ICON grid conserves mass.

3.4.1.3 Assimilated observations

For the inversion we use hourly in-situ methane concentration observations from 19 ICOS stations available from the ICOS-Portal. The stations that provide data for summer 2018 include ICOS towers with multiple sampling heights and are listed in Table 10. Importantly, the selection of 19 stations was made only for a minimal test case as more data—including the ICOS obspack stations—are available and currently being integrated in our inversion system.

3.4.2 Data assimilation and inversion concept

The general framework that is being implemented at DWD is designed as a weakly-coupled GHG data assimilation scheme combining meteorology and GHG concentrations. It combines the three tasks of (1) describing the current meteorological state of the atmosphere, (2) describing GHG concentration fields in the atmosphere, and (3) estimating the GHG fluxes. GHG concentrations and fluxes are assimilated jointly, but the meteorological state is treated separately (see Figure 31). The ensemble-based approach builds upon the data assimilation system used for the numerical weather prediction (NWP) at DWD.

Each ensemble member in an ICON-ART run consists of both the meteorological and concentration/flux fields. For each ensemble member, the output (first guess) is split into two parts: the meteorological fields and the part containing the concentrations/fluxes. Each part is separately presented as an input to the meteorological and GHG data assimilation systems, respectively. Each of these two systems generates a new meteorological and GHG analysis ensemble. The ensembles are then merged together (thereby maintaining the previous one-to-one relationship among the members) to set up the ensemble for the next ICON-ART run. This procedure is cyclically repeated.

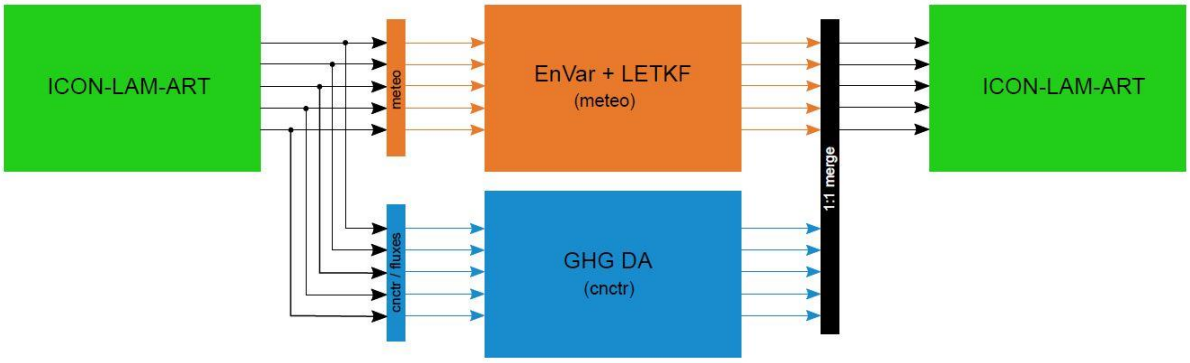


Figure 31: The weakly-coupled GHG data assimilation scheme (here with an illustrative ensemble size of five).

3.4.2.1 General data assimilation setup

We now take a closer look at the blue “GHG DA” block of Figure 31. In a first implementation, the GHG data assimilation system will be based on the OmniVAR with different realisations of the B-matrix. Moreover, the OmniVAR is to be extended in such a way that the simultaneous treatment of the flux inversion problem is possible.

The OmniVAR is part of DWD’s Data Assimilation Coding Environment (DACE) and was designed for a generic and efficient treatment of the minimization part which is common for most variational algorithms for data assimilation: Determine the analysis x_a as best state approximation by minimising the distance to a first guess x_b and the deviation from observations y, i.e.,

$$x^a := \operatorname{argmin}\left(\|x - x^b\|^2_{B^{-1}} + \|y - H(x)\|^2_{R^{-1}}\right)$$

with observation operator *H*, background error covariance matrix *B* and observation error covariance matrix *R*. Setting the gradient of the expression in brackets to zero, the solution can be obtained in two steps (as done for 3DVar and EnVAR): First, solve

$$(R + HBH^\top)z = y - H(x^b)$$

for *z* and then do the post-multiplication

$$x^a - x^b = BH^\top z.$$

Moreover, the OmniVAR provides a framework for the whole variational data assimilation setup by defining interfaces to other parts of the DACE code and preparing code templates, data structures and procedures for the post-multiplication as well.

3.4.2.2 Extensions to GHG data assimilation and flux inversion

So far, the state vector *x* for atmospheric data assimilation only contained atmospheric variables which were chosen to provide the NWP model with initial values. Of course, these initial values should represent the atmospheric state as accurately as possible (being the main goal of atmospheric data assimilation). Now, as we are also interested in flux inversions, the state vector is extended as described in the following lines.

We consider a fixed point in time. Let x_c denote the concentration of a GHG species in the atmosphere and let x_e denote the fluxes (e.g. emissions) from the Earth's surface in some bounded domain. We introduce

$$\vec{x} := \begin{pmatrix} x_c \\ x_e \end{pmatrix} \in \mathbb{R}^{n_c + n_e}$$

as state vector (with some natural numbers n_c, $n_e > 1$) and observe that its background error covariance matrix takes the form

$$\vec{B} := \begin{pmatrix} B_{cc} & B_{ce} \\ B_{ec} & B_{ee} \end{pmatrix} \in \mathbb{R}^{(n_c+n_e)\times(n_c+n_e)},$$

where B_{cc} and B_{ee} describe the background error covariance matrices for the concentration and fluxes, respectively, and $B_{ce} = B^T_{ec}$ describe the covariances between concentrations and fluxes mutually. Furthermore, we observe that the fluxes do not have a direct impact on the observations, but indirectly via the concentrations built up so far. As a consequence, the observation operator takes the form

$$\vec{H} = (H_c \quad 0)$$

with

$$H_c : \mathbb{R}^{n_c} \to \mathbb{R}^{m_c}$$

being the linear interpolation operator which maps the species concentration to the pointwise spatial locations of the observations; and for the observation error covariance matrix we have

$$R_{cc} \in \mathbb{R}^{m_c \times m_c}.$$

Now, by substituting the quantities with their vector counterparts, the update equation reads

$$\vec{x}^a = \vec{x}^b + \begin{pmatrix} B_{cc} \\ B_{ec} \end{pmatrix} \underbrace{H_c^\top (R_{cc} + H_c B_{cc} H_c^\top)^{-1} (y - H_c x_c^b)}_{=:z},$$

with

$$\vec{x}^a := \begin{pmatrix} x_c^a \\ x_e^a \end{pmatrix} \qquad \text{and} \qquad \vec{x}^b := \begin{pmatrix} x_c^b \\ x_e^b \end{pmatrix}$$

respectively. Note that thanks to the vector

$z \in \mathbb{R}^{n_c}$,

the implementation of the extension of the OmniVAR can be made in such a way that besides a simultaneous calculation of both the concentration and flux updates it is also possible to calculate only one of them.

3.4.3 Simplified inversion

While the concept described in Section 3.4.2 is not yet fully implemented, we use simpler test cases. Here we describe an inversion of methane emissions over a two-month period aiming at determining scaling prefactors for emissions from different categories. This method—called synthesis inversion—is the simplest special case for the previously described concept as it is technically implemented like a GHG inversion with a static ensemble.

We choose June and July 2018 as a study period to facilitate comparisons with other project partners, and focus on emissions from Germany. However, this test run mainly aims at demonstrating the main ideas and learning about possible difficulties, and it is not expected to provide meaningful results.

3.4.3.1 Synthesis inversion method

We split the a priori emission estimate into a handleable set of categories. The inversion task is to adjust scaling prefactors for emissions from these different categories. Here we only categorise emissions by area, but separation by sectors is planned in the future. Since we aim to validate the overall emissions from Germany, we select these emissions as one category. Emissions from outside Germany are split in the four categories North-West, North-East, South-West, and South-East (see Figure 32). We optimise the scaling factor for each category to match the observed methane concentrations, computing also the uncertainties and correlations of all scaling factors.

Besides the emission categories, a global, time-constant bias in the methane concentrations was allowed in the inversion as a simplified approach to account for uncertainties in the initial and boundary conditions.

For the inversion, we need to assign uncertainties to the comparison of observations and model results. We assume that all considered observations—described in Section 3.4.1.3 — are independent. To compare the in-situ observations to the concentrations on our model grid, we use a simple inverse-distance-weighted horizontal interpolation and a linear interpolation in the vertical direction. To account for uncertainties arising from this interpolation, the choice of model-equivalent heights, transport uncertainties, and for other uncertainties in our model, we assume uncertainties of 100, 150, and 200 ppb in the model-observation mismatch. A comparison of results for these different uncertainties is provided in Table 8.

We compute posterior scaling factors for multiple time series, each consisting of multiple 24h inversion time windows. More precisely, we consider 10 partially overlapping time series starting every 5 days and each lasting 15 days, which in combination cover the 60 days of our

forward simulation. Within each time series, a new inversion is calculated for each day (starting at 0 UTC). The posterior emissions of the previous day are used as prior for the next day. When starting a new time series, the prior is reset to the initial estimate with all emission scaling factors set to 1.

At the beginning of each assimilation time window, a so-called background concentration is defined as the methane that is currently in the atmosphere. To calculate this background, we use the scaling factors obtained for the previous day in the same time series. Here we assume that the emissions are constant in time and that the transport model is linear. Thus, we adjust all emissions that were released into the atmosphere before the considered time window. Once the background is fixed, we can assimilate the scaling prefactors for the emissions that occurred during the given assimilation window using all observations during this time window.

3.4.3.2 Results

We present preliminary results for a synthesis inversion with very simplified emission categories. These results are not expected to yield an improvement of the prior emissions, but shall rather help identify obstacles and potentials for the further development of the inversion method.

In our analysis we first note that the synthesis inversion yields temporal variations of the scaling factors together with an overall trend. For different measurement and prior uncertainties, this trend yields the estimates in Table 8 and Table 9 which are compatible with the prior of value 1 for each area.

Table 8: Scaling factors with their uncertainty estimates for different uncertainties of the model-observation comparison (left column). The area emissions for Germany (DE) and the remaining areas (NW, NE, SE, SW) have a prior of 1 with uncertainty 10%, and the concentration bias offset (BO) has a prior of 0. The global concentration bias is obtained by multiplying its scaling factor with 50 ppb. The uncertainty is determined by the variation of the results from different time series.

Uncertainty	DE	NW	NE	SE	SW	BO
100 ppb	1.065(55)	1.029(22)	1.017(11)	1.32(15)	1.21(13)	0.679(96)
150 ppb	1.036(30)	1.014(10)	1.0070(51)	1.183(85)	1.118(77)	0.681(92)
200 ppb	1.024(18)	1.0080(60)	1.0043(29)	1.115(52)	1.075(49)	0.670(88)

To obtain reasonable results, we need to assume a comparably large model-observation uncertainty of 150 ppb. This high uncertainty signifies the relevance of estimating uncertainties in the meteorology or in the initial concentrations for the improvement of the inversion method. We assume that this global uncertainty estimate can be reduced significantly when explicitly accounting for the transport uncertainty.

To estimate the uncertainty in the computed scaling factors, we compare 10 different time series. Since we use overlapping time series with correlated results, we simply take the standard deviation of the 10 different scaling factors for each area.

Table 9: Scaling factors for different prior uncertainties. The prior uncertainties of the area scaling factors are given in the left column. All other parameters are as in Table 8, with a measurement uncertainty of 150 ppb.

Prior uncertainty	DE	NW	NE	SE	SW	BO
5%	1.0091(87)	1.0035(27)	1.0013(13)	1.053(25)	1.035(24)	0.694(90)
10%	1.036(30)	1.014(10)	1.0070(51)	1.183(85)	1.118(77)	0.681(92)
20%	1.116(76)	1.050(36)	1.041(19)	1.47(21)	1.31(18)	0.648(96)

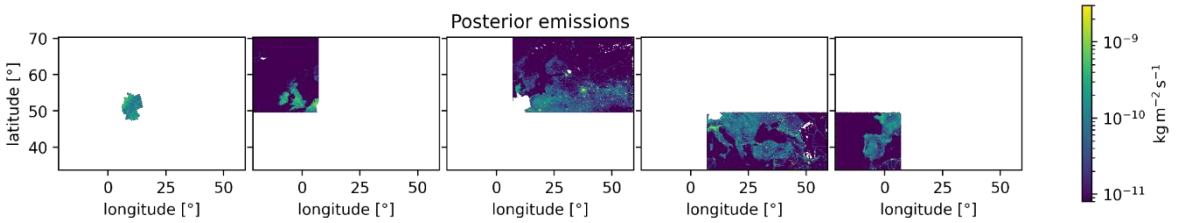


Figure 32: According to region, the emissions from Germany, from the North-West, North-East, South-East and South-West sectors were scaled separately with a model-observation mismatch uncertainty of 150 ppb and a prior uncertainty of 10%.

To validate our results and check the dependence on individual stations, we compare inversion results when excluding individual stations in Table 10. This shows a comparably strong effect of the station Ispra (IPR) on the posteriori emission estimate not only in the south-east region, but also in Germany.

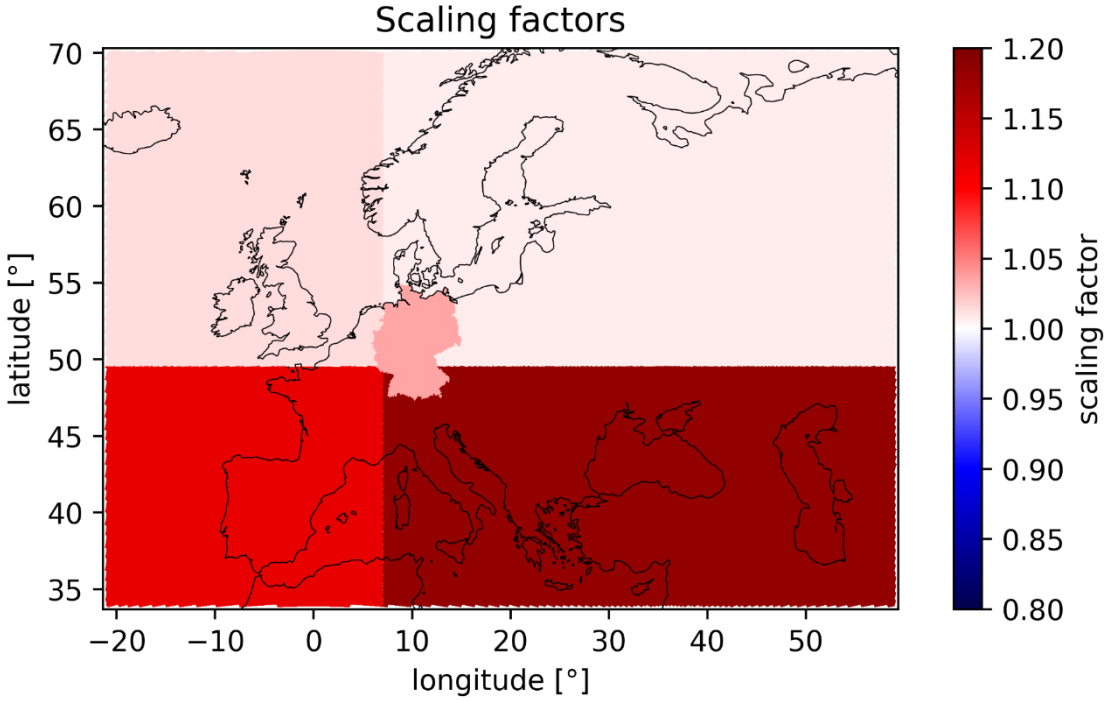


Figure 33: Scaling factors of the five areas obtained with a model-observation mismatch uncertainty of 150 ppb and a prior uncertainty of 10%.

The presented results highlight the relevance of various possible improvements of the method, which we summarise in the following:

Uncertainty estimation and validation: One main focus of the further development of the method lies in the better assessment and the reduction of uncertainties. To begin, our model equivalents for the measurements can be improved by a better interpolation between grid cells. Also, an optimal choice of model heights corresponding to the observation heights may improve the comparison between model and observations. A major improvement is expected from systematically taking into account uncertainties in the transport model by using meteorological ensembles.

Longer analysis times: Running an inversion with a long assimilation time window for a forward model like the ICON—for which no adjoint model exists—requires a careful treatment of temporal correlations. The synthesis inversion is particularly well suited for this task and can in principle be used for long time windows. Choosing a good time window for an inversion requires an understanding of uncertainties of the transport model, which will be obtained using meteorological ensembles. Further improvement of longer inversions is expected when considering uncertainties in the boundary conditions of our limited area model. A clear goal is to leverage the potential of long assimilation windows in a longer run of the forward model for a whole year. A comparison of our forward model results to an independent concentration assimilation may help understanding the potential and risks of such long-time calculations.

Resolution and choice of emission areas: A detailed analysis requires more emission categories from different sources. These can be given by splitting the covered area in many patches, and by using multiple emission categories within the same area, associated with different sectors. In any case, this means including more emission categories in each run of the forward model. This also opens a path to assessing the uncertainty associated with this parametrization of the emissions, as we can check how the results change if we lower the number of parameters by coupling scaling factors for different categories.

More input data: Both the concentration observations and the prior emission estimates could be improved by using more information. Using more observations and possibly other observation types is expected to make our inversion more stable. For the emission estimates, taking into account modelled time-dependence may improve the results.

Lessons learnt: In addition to the ideas for improvement summarised above, this demonstration of our model has drawn our attention to many important technical details that will lead to an improvement of the accuracy and reliability of our results. For example, we have noted the relevance of nonlinearities in the transport model, which can be corrected for by a simple linearization in further calculations.

Table 10: Dependence of scaling factors on individual stations. Given are scaling factors for an inversion excluding the station in the left column. All parameters are as in Table 8, with a measurement uncertainty of 150 ppb. We find that leaving out the station Ispra (IPR) has the strongest overall effect on the scaling factors.

Excluded station	DE	NW	NE	SE	SW	BO
GAT	1.023(24)	1.0127(92)	1.0070(60)	1.183(83)	1.118(80)	0.680(94)
HPB	1.043(35)	1.014(10)	1.0067(48)	1.183(86)	1.117(76)	0.682(93)
HTM	1.029(29)	1.013(10)	1.0102(47)	1.180(85)	1.113(77)	0.695(90)
SMR	1.027(27)	1.013(10)	1.0090(56)	1.179(86)	1.112(75)	0.700(86)
IPR	1.061(27)	1.016(10)	1.0138(53)	1.0045(97)	1.134(74)	0.630(94)
JFJ	1.030(29)	1.013(10)	1.0055(50)	1.188(85)	1.115(77)	0.692(91)
KIT	1.044(25)	1.014(10)	1.0077(47)	1.183(84)	1.120(78)	0.675(95)
KRE	1.043(31)	1.014(10)	0.9892(35)	1.178(84)	1.124(78)	0.662(93)
LIN	1.009(23)	1.014(11)	1.0072(59)	1.185(83)	1.120(79)	0.674(93)
CMN	1.032(30)	1.013(10)	1.0061(51)	1.186(84)	1.116(77)	0.687(94)
NOR	1.026(28)	1.013(10)	1.0075(57)	1.178(85)	1.112(76)	0.701(87)
OPE	1.040(30)	1.013(11)	1.0085(50)	1.186(84)	1.105(73)	0.669(92)
PAL	1.033(29)	1.013(10)	1.0062(51)	1.182(85)	1.116(77)	0.687(89)
PUY	1.034(30)	1.013(10)	1.0065(50)	1.182(85)	1.127(82)	0.685(92)
SAC	1.047(31)	1.0048(32)	1.0100(50)	1.189(82)	1.030(15)	0.656(97)
SVB	1.026(28)	1.013(10)	1.0048(49)	1.178(86)	1.111(76)	0.702(89)
TOH	1.034(31)	1.013(10)	1.0056(59)	1.180(84)	1.114(77)	0.691(92)
TRN	1.042(31)	1.0119(94)	1.0088(49)	1.186(83)	1.113(77)	0.666(94)
UTO	1.033(29)	1.013(10)	1.0064(50)	1.182(85)	1.116(76)	0.687(89)

3.5 WRF analytical inversion for CO_2 using in situ observations (AGH)

3.5.1 Model description

3.5.1.1 Transport model

We use WRF v.4.4.1. as the transport model. WRF-ARW is an Eulerian NWP model developed in a collaboration headed by the US National Center for Atmospheric Research (NCAR). It integrates the non-hydrostatic, fully compressible Euler equations in flux form on a terrain-following mass-based vertical coordinate. The WRF model has been successfully applied from global to microscale, with effective downscaling possible through one- or two-way nesting.

In our configuration, the model is run as a limited area model. The spatial extent of the model domain covers Central Europe, as displayed in Figure 34. Two nested high-resolution domains were primarily run for CoCO2 tasks associated with other deliverables, thus their respective description is omitted. The primary domain (d01) is run at 5 km x 5 km spatial resolution with 80 vertical levels, spanning vertically up to 50 hPa (approximately 21 km m.a.s.l.). The vertical levels are variably distributed, with their density increasing non-linearly towards the surface (with typically 30 levels below 3000 m.a.g.l.). Temporal resolution of the model is 20 seconds.

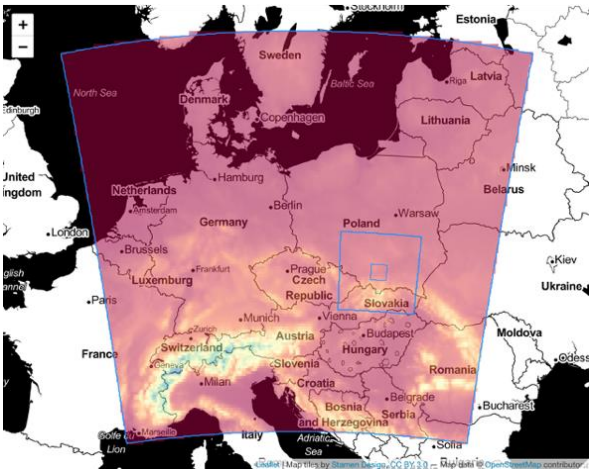


Figure 34: Spatial extent of the WRF-GHG modelling domain framework applied by AGH for CoCO2. Internal domains (blue rectangles centered around Kraków, southern Poland) were used for city-scale simulations within Task 4.3.2.

We use ERA5 data for meteorological boundary conditions, downloaded from its official repository (Climate Data Store) at 0.25° x 0.25° horizontal and 137 pressure-level vertical resolution.

For our experiment, WRF was configured with the Thompson microphysics scheme, RRTMG schemes for longwave and shortwave radiation, revised MM5 scheme for surface layer physics, and the Noah-MP land surface model. The Grell 3D cumulus parameterization was enabled in the parent domain. Due to our internal domains running in grey-zone horizontal resolutions, we have applied the Shin Hong PBL scheme in all nests. Full instantaneous model output was saved every 15 minutes. For comparisons against hourly observations, averages were calculated using respective instantaneous values from the model.

To simulate atmospheric mole fractions of CO_2, we have used the GHG module from the WRF-Chem suite, distributed with the model code, which was developed specifically for greenhouse gas simulations (this framework is referred to as WRF-GHG, (Beck et al., 2011)). The tracers in the model are run as fully passive and advected internally every model timestep. To represent influences from beyond modelled domain, a single background tracer was advected through the modelling system (co2_bck). Its initial and lateral boundary conditions were assigned from the interpolated CAMS product “Global atmospheric concentrations inferred by atmospheric inversion”, as per the modelling protocol. As the CAMS product and our WRF domain were available on different spatial grids, we used bilinear regression for horizontal interpolation, and linear interpolation in both vertical and temporal dimensions in order to provide the necessary input for our model framework.

In order to minimise the discrepancy between the modelled and actual meteorology without using the formal meteorological data assimilation (or nudging), we have applied a 24-hour meteorological restart system. The simulated periods were executed as a sequence of partially overlapping 30-hr model runs starting at 18:00 UTC, with the first 6 hours being used for meteorological spinup. At midnight between day 0 and day 1, the CO_2 tracers were copied from the previous day’s simulation – except for the very first day in the sequence, when the previous day’s data were not available. The advantage of this approach is that every 24 hours, the model is reinitialized from ERA5 initial conditions, thus preventing a large departure of meteorological variable values from the observations. This approach has been tested in past scientific studies, first adopted by Ahmadov et al. 2007.

3.5.1.2 Prior fluxes and assimilated observations

As input, we’ve used the following prior fluxes:

- **Prior anthropogenic emission fluxes**: based on TNO_GHGco_6x6km_v4_0_year2018 inventory for the year 2018, and TNO_GHGco_6x6km_v5_0_year2021 for year 2021. Vertical profiles and temporal profiles were used as specified in the TNO dataset. No explicit plume-rise mechanism was applied. Following category aggregates (chosen based on the annual fluxes in the region of interests, so that the emissions followed by each tracer were of the same order of magnitude) were used as priors:
- gnfr_a – corresponding to GNFR A category only (power generation, public)
- gnfr_b – GNFR B (industry)
- gnfr_c – GNFR C (other stationary combustion)
- gnfr_fgi - sum of emissions from GNFR categories F1–F4, G and I (transport, shipping, offroad)
- gnfr_other – sum of all remaining emissions (fugitive emissions, solvents, aviation, waste, agriculture)
- **Prior biospheric fluxes**: pre-computed offline VPRM fluxes as provided within the T4.4 protocol,
- **Prior forest fire fluxes**: Global Fire Emissions (GFAS, https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-fire-emissions-gfas?tab=overview). Daily emissions used. Surface emissions only, temporal profile constant over the day.
- **Prior ocean emissions**: provided by Cyril Germineaud at Mercator (as per protocol). Monthly values used. Surface emissions / uptake only, no vertical profile, the temporal profile is constant over time.
- **Prior lateral fluxes**: from Frederic Chevallier (as per the protocol); computed as sum of the following variables (from the input file): *'allcropsource'*, *'allwoodsource'*, *'biofuelcropsource'*, *'lakeriveremis'*, *'biofuelwoodsource'*. No vertical profile, surface emissions only. The temporal profile is constant over time.

In order to allow the optimization of subregions of our domain, we have applied spatial masks to the anthropogenic and biogenic fluxes for optimization of national fluxes of Poland (PL) and Germany (DE). Other countries were not considered due to computational constraints.

A total of 25 tracers were simulated for the purpose of this deliverable. These were:
- co2_bck: Background tracer
- co2_ocean: Tracer of ocean emissions
- co2_lateral: Tracer for lateral CO_2 fluxes
- co2_bbu: Tracer for biomass burning
- co2_vprm_gpp_[region]: CO_2 flux of Gross Primary Production (photosynthetic uptake) from VPRM model, geographically divided into 3 regions of interest: PL, DE, and remaining part of the domain (rest)
- co2_vprm_resp_[region]: CO_2 flux of heterotrophic respiration from VPRM model, geographically divided into 3 regions of interest: PL, DE, rest
- co2_gnfr_[anth_group]_[region]: 15 tracers of anthropogenic CO_2 fluxes, five for each region of interest (PL, DE, rest), based on TNO_GHGco_6x6km_v4_0 product provided within the Work Package.

All the other tracers were emitted exclusively into (or subtracted from) the first model layer, between 0 and 25 m agl.

For each geographic region, five anthropogenic tracers were created in order to follow emissions from one or more subcategories of the input data.

We have assimilated surface station data from the European Obspack compilation product, based on GlobalviewPlus_v8 (ICOS RI, 2023), which contains stations from the ICOS network supplemented by data collected at two surface sites located in Poland (BIK - Białystok, and KAS - Kasprowy Wierch). We have also excluded stations located close to our domain boundaries (to avoid boundary effects). Finally, in the course of the analysis we have excluded two stations heavily influenced by local sources (HEI, ISP), as our model was not able to represent the variability well, and thus these affected the results disproportionately. The full list of stations and inlet altitudes (highest available) from which data was assimilated in the runs is as follows:

BIK 300 m, BRM 212 m, GAT 341 m, HEL 110 m, HPB 131 m, HUN 115 m, HTM 150 m, JFJ 5 m, JUE 120 m, KAS 7 m, KIT 200 m, KRE 250 m, LIN 98 m, LUT 60 m, OXK 163 m, PRS 10 STE 252 m, TOH 147 m, WES 14 m, ZSF 3 m.

The exact locations of the measurement sites are available with the dataset. Local solar time was calculated for each station. Afternoon (LT 11-16) data was selected in the inversions, except mountain stations, for which night time data was selected (LT 23-04).

3.5.1.3 State vector

The state space in our framework was defined as linear scaling factors of the a-priori emissions, with a single value for each state space member across the modelled period (i.e. we considered 4 identical and independent state spaces for our four modelled periods). We optimize them using the analytical inversion formula that optimizes the cost function

$$J(\lambda) = (y - \boldsymbol{K}\lambda)^T \boldsymbol{S}_\epsilon^{-1}(y - \boldsymbol{K}\lambda) + \left(\lambda - \lambda_{prior}\right)^T \boldsymbol{S}_{prior}^{-1}\left(\lambda - \lambda_{prior}\right)$$

by means of resolving the linear system of equations:

$$\hat{\lambda} = \lambda_{prior} + \left(\boldsymbol{K}^T \boldsymbol{S}_\epsilon^{-1} \boldsymbol{K} + \boldsymbol{S}_{prior}^{-1}\right) \boldsymbol{K}^T \boldsymbol{S}_\epsilon^{-1}\left(y - \boldsymbol{K}\, \lambda_{prior}\right)$$

The notation above follows Rodgers (2000).

In all our inversions, we assume perfect knowledge of ocean, lateral and biomass burning fluxes. These were subtracted from the observational data prior to running the inversion.

To represent uncertainty of the transport model and observations, we construct the transport error covariance matrix from two components assuming a model-data mismatch of 8 ppm, with roughly a third of the mismatch (3 ppm) temporally correlated (exponentially decaying with τ of 180 minutes).

We assume *a priori* emission uncertainty as fully uncorrelated and equal: 10% for anthropogenic emissions and 50% for biogenic components, regardless of the level of aggregation. The inflated values of the anthropogenic fluxes (as compared to the values reported e.g. within UNFCCC National Inventory Reports, where they are reported as close to 2-3% depending on the emission sector) is due to two main considerations:

a) we take into the account hourly fluxes over limited time periods; annual fluxes can be much better constrained, as in many cases the emission factors used for emission estimations rely on accurately monitored proxies, for which higher resolution data are not available,
b) we use temporal factors provided by TNO that are, again, based on standardised activity data not based on actual activity information. The uncertainty of these factors was not provided and is difficult to assess.

The uncertainty values for biogenic emissions were assigned based on consensus within the task group, literature overview and finally results of the sensitivity tests performed within the scope of the project (values of 50%, 100% and 200% were tested).

Due to the structure of the framework and high speed of computations, we were able to test multiple configurations of the state vector of various complexities. The simplest were constructed with the state space consisting of aggregated biogenic (GPP + RES) and anthropogenic (sum of all components) fluxes, and the most complicated built from the full suite of available biogenic and anthropogenic tracers (21 members) plus an extra state space element for background adjustment

Following these tests, we have opted on running the system in two modes for the purpose of this report:

- full state space with variable background, and
- reduced state space with variable background.

In the full state space, we adjust scaling factors associated with all the tracers described in the previous section, allowing for optimization across both emission sectors and geographical regions.

In the reduced state space configuration, we aggregated all the anthropogenic tracers into their respective sum, allowing us to optimise the full anthropogenic flux from Poland and Germany. We also aggregated the respiration and net ecosystem exchange fluxes, forcing the system to effectively optimise the net CO_2 flux. This second, simpler configuration is considered closest to classical, previously published studies, as those former works primarily dealt with optimization at this level of detail at most (in terms of treating natural and anthropogenic fluxes).

3.5.1.4 Period of study

Due to the limits of the computational costs, primarily stemming from necessary high-cost computations of city-scale emissions, we have limited the period of study for the national scale inversions to the four separate monthly periods defined as reference months in the T4.4 protocol.

The simulated analysis periods were as follows:

- winter 2018: Feb 1st, 00:00 UTC – Mar 1st 00:00, 2018
- summer 2018: Jul 1st, 00:00 UTC – Aug 1st 00:00, 2018
- winter 2021: Feb 1st, 00:00 UTC – Mar 1st 00:00, 2021
- summer 2021: Jul 1st, 00:00 UTC – Aug 1st 00:00, 2021

For each of the aforementioned periods, separate inversions were run, independent of each other. A 10-day spin-up period was executed prior to the listed analysis periods in order to flush the model domain and reduce the influence of initial conditions on the optimization process.

3.5.2 Results

3.5.2.1 Full state space

We performed the inversion using the full (22-member) state space for all study periods. Figure 35 and Figure 36 present results for two study periods as examples (Feb 2018 and Jul 2018, respectively), with aggregated results from all periods summarised in Table 11.

What is apparent when using full state space is that the flux adjustments in the inversion system affect mostly the values of biogenic fluxes, with significant adjustments only in GPP and respiration fluxes of VPRM and very minor readjustments to the anthropogenic fluxes over the domain.

The inversion predicts slightly elevated net CO_2 fluxes in the winters of both 2018 and 2021, with slightly higher emissions in the latter in both Germany and in the domain fringes (rest), with average fluxes of CO_2 of approximately 330 – 340 Tg-CO_2/yr in Germany. In Poland the model does predict smaller net emission for Feb 2018, which the model attributes exclusively to reduced respiration (Figure 35, upper panel; prior = 239 ± 119 Tg-CO_2/yr, posterior = 181 ± 42 Tg-CO_2/yr).

In July of both years, the inversion consistently predicted decreases in net uptake across the modelled domain, with significant adjustment in CO_2 flux in Germany in both. In 2018, the predicted flux was increased from -751 ± 531 Tg-CO_2/yr *a priori* to -216 ± 40 Tg-CO_2/yr *a posteriori* (71% decrease of uptake), while in 2021 the *a priori* monthly average flux of -1148 ± 811 Tg-CO_2/yr was adjusted to -593 ± 89 Tg-CO_2/yr. Interestingly, in 2021 the optimization algorithm yielded simultaneous increases of both respiration flux and uptake, the latter being adjusted weaker than the former (not shown). Similar predictions were observed also for other spatial regions that year, suggesting an underestimation of the strength of the diurnal cycle by VPRM. That same pattern, albeit weaker, was present also in 2018 in Poland and in the domain fringes, but not in Germany, where a slight increase of respiration (14%, to 1323 ± 234 Tg-CO_2/yr) was calculated together with a decrease of photosynthetic flux (22%, to -1489 ± 136 Tg-CO_2/yr).

In terms of anthropogenic fluxes, the results point consistently to a small reduction (as compared to the prior) of anthropogenic emissions in Germany, by -3.5% for Feb 2018 and -4.2% in July 2018, and a stronger one (by -14.8% and -24.5%) in 2021. However, a large portion of this reduction seems to be caused by significant adjustment to the power generation (GNFR_A), which is usually considered to be well known. One of the hypotheses being tested is the vicinity of a single strong point source, which in the analytical inversion setup might cause elevated model-data mismatches. However, the tests performed so far remain inconclusive.

Only negligible adjustments were observed in Poland, underlining the overall lack of observational sites. The Białystok (BIK) station, located in north-eastern Poland, close to the border with Belarus, is located far away from main industrial and population areas, thus the probability of sampling anthropogenic signals is low under the most frequent westerly synoptic air flow. To further aggravate the matter, the data from BIK for July 2021 were missing altogether due to equipment malfunction at the station, forcing the model framework to rely on the mountain station of Kasprowy Wierch (KAS) and the German ICOS network to constrain the national emissions. As is immediately visible, the inversion framework clearly suggest that this network configuration was not sufficient for that purpose.

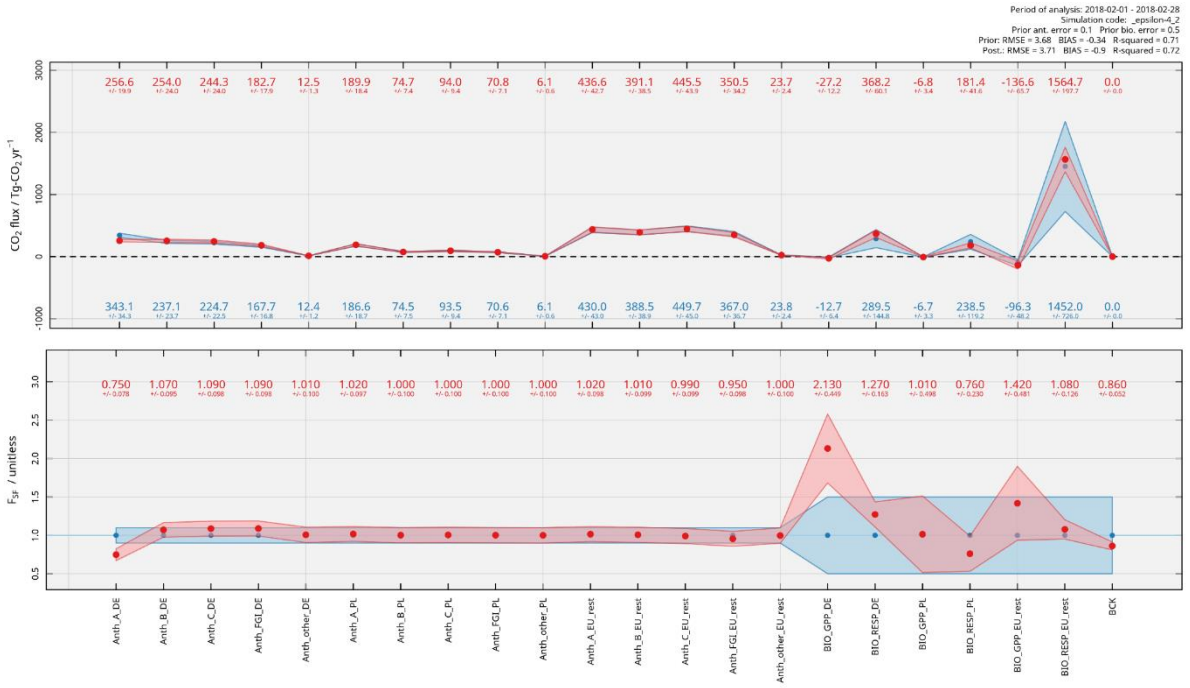


Figure 35: Results for the full state space, February 2018. Blue: a priori. Red: a-posteriori. Top panel: absolute flux value for each state vector element. Emissions for BCK have no physical meaning and are set to 0. Bottom panel: adjustment of scaling factors.

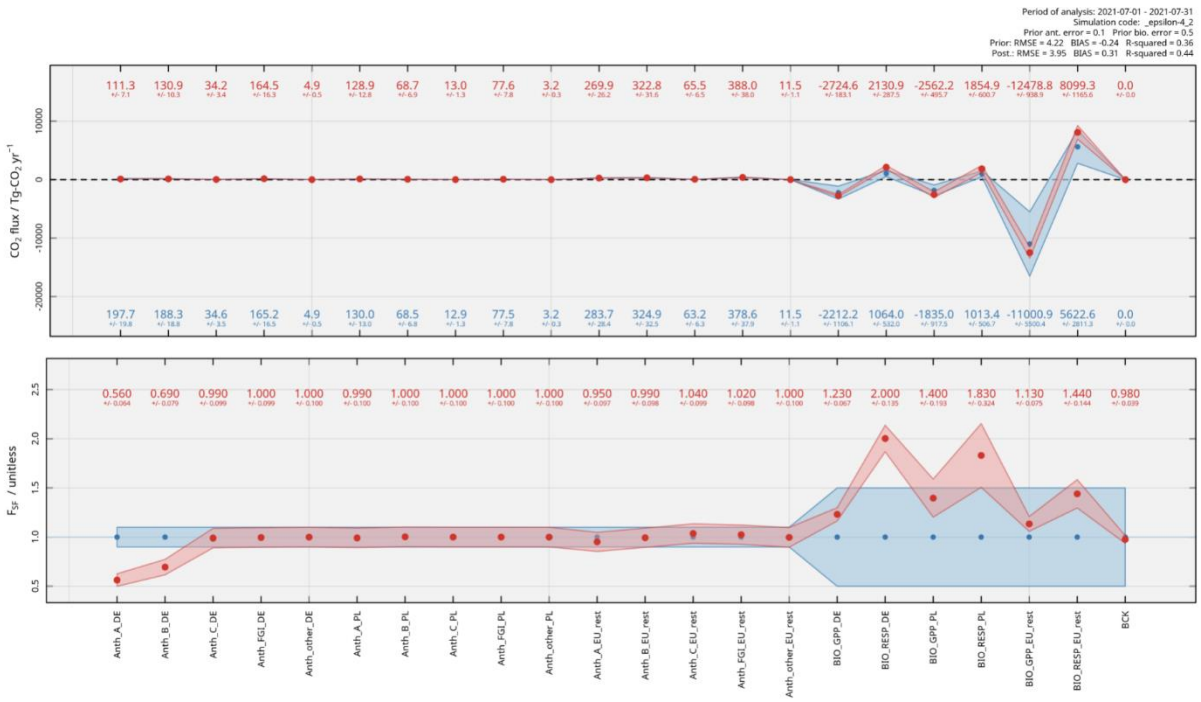


Figure 36: Results for the full state space, July 2021. Blue: a priori. Red: a-posteriori. Top panel: absolute flux value for each state vector element. Emissions for BCK have no physical meaning and are set to 0. Bottom panel: adjustment of scaling factors.

Table 11: Results of full state space inversion. Fluxes in Tg-CO$_2$ yr^{-1}.

Period	Region	Prior anthr	Post anth	Prior bio	Post bio
February 2018					
	PL	432 ± 96	435 ± 97	232 ± 164	174 ± 96
	DE	985 ± 220	950 ± 199	277 ± 195	341 ± 163
	Rest	1659 ± 371	1647 ± 363	1356 ± 959	1428 ± 710
July 2018					
	PL	287 ± 64	289 ± 65	-578 ± 409	-527 ± 172
	DE	662 ± 148	634 ± 140	-751 ± 531	-216 ± 40
	Rest	1104 ± 247	1090 ± 242	-4995 ± 3532	-4379 ± 711
February 2021					
	PL	441 ± 99	451 ± 100	243 ± 171	268 ± 144
	DE	924 ± 207	787 ± 149	313 ± 221	329 ± 81
	Rest	1629 ± 364	1469 ± 317	1448 ± 1024	940 ± 221
July 2021					
	PL	292 ± 65	291 ± 65	-822 ± 581	-707 ± 267
	DE	591 ± 132	446 ± 89	-1148 ± 811	-593 ± 89
	rest	1062 ± 237	1057 ± 232	-5378 ± -3803	-4379 ± 711

Overall, the results seem to suggest that the system remains under-constrained in this setup, and at least in one case a dipole effect (where fluxes of correlated source categories cause compensating, unreasonable flux adjustments) was apparent – namely in the simultaneous decrease of emissions by power generation (GNFR_A) and increase in industrial and other stationary combustion (GNFR_B and GNFR_C, respectively) in Germany in winter 2018. We therefore elected to aggregate the anthropogenic fluxes, running the inversion framework at reduced complexity.

3.5.2.2 Reduced state space

In the reduced state space, which would be representative of a more “classic” inversion configuration, we considered only the fluxes from aggregated anthropogenic and aggregated biogenic fluxes, effectively reducing the state space from 22 members to 10. Figure 37 shows the optimised fluxes calculated for Feb 2021 (c.f. Figure 37), and Table 12 presents the complete set of results.

Qualitative results remain unchanged as compared to the full state space, meaning that in all subperiods and all subregions the sign (as well as the order of magnitude) of flux adjustment remain unchanged, supporting the conclusion that the system is able to successfully constrain the fluxes using the mole fractions in a consistent manner.

For Poland the ability of the model to constrain anthropogenic emissions remains severely limited. Small increases in emissions for all studied periods (1.8%, 4.5%, 7.5%, 5.5% for Feb

2018, Jul 2018, Feb 2021 and Jul 2021) are not accompanied by uncertainty reductions at all, further emphasizing the lack of data necessary to accurately employ the inversion system.

For Germany, we do observe anthropogenic flux adjustments and simultaneous uncertainty reductions, of, chronologically: -11%, -2% (Feb and Jul 2018) and -35% and -24% (Feb and Jul 2021). Based on the results for the full state space we hypothesise that significant reductions in 2021 likely have to do with improper transport of the anthropogenic tracers.

Biogenic fluxes are optimized with much higher uncertainty reduction rates. Winter fluxes are estimated to be larger in 2021 for both countries, by 64% for Poland and 20% for Germany. Again, a reduction of CO2 release is observed in Poland for winter 2018 (to 168 ± 42 Tg-CO_2 yr^{-1}) with simultaneous increase of the flux in Germany (436 ± 86 Tg-CO_2 yr^{-1}) and the rest of the analysed domain, an effect not observed in 2021 (PL: 276 ± 72 Tg-CO_2 yr^{-1}, DE: 526 ± 62 Tg-CO_2 yr^{-1}).

The inversion also shows that net ecosystem fluxes were more negative (i.e. taking up more carbon dioxide) in 2021 than in 2018, with CO_2 net flux diminishing from -672 ± 95 Tg-CO_2 yr^{-1} to -806 ± 139 Tg-CO_2 yr^{-1} in Poland (20% increase absolute value), and -481 ± 37 Tg-CO_2 yr^{-1} to -890 ± 48 Tg-CO_2 yr^{-1} in Germany (85% more uptake). Same relative rates for a-priori fluxes were larger for Poland (42%) and smaller for Germany (53%).

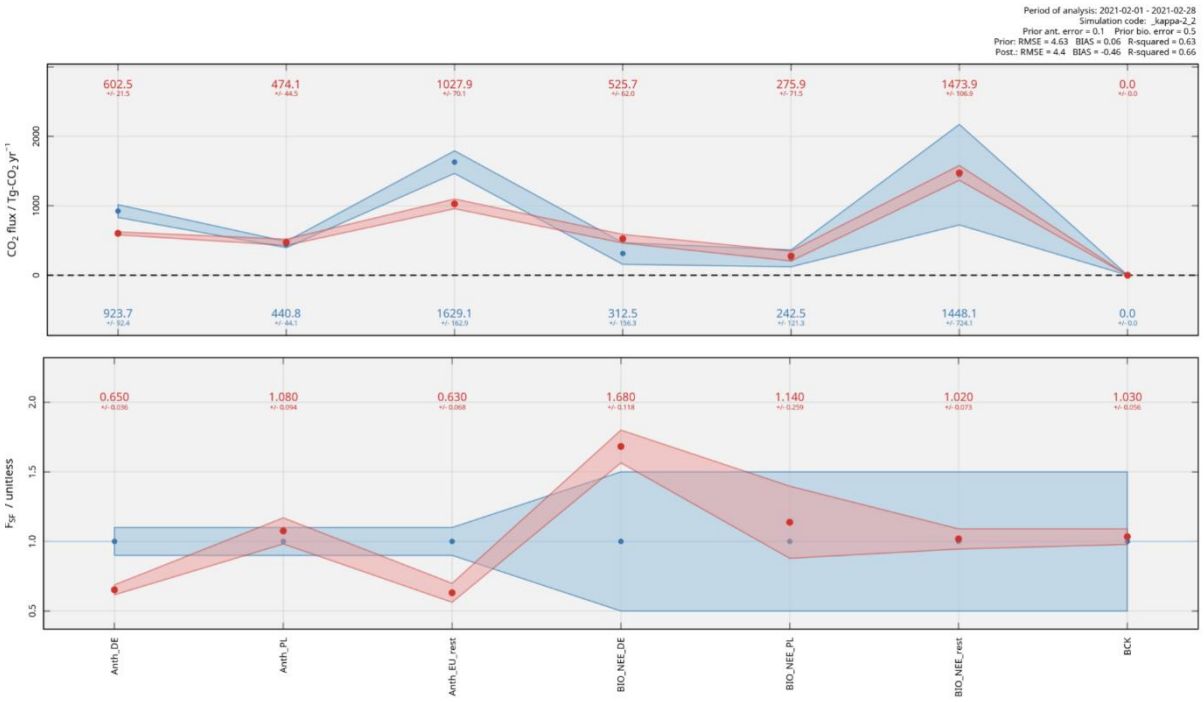


Figure 37: Results for reduced state space, Feb 2021. Blue: a priori. Red: a-posteriori. Top panel: absolute flux value for each state vector element. Emissions for BCK have no physical meaning and are set to 0. Bottom panel: adjustment of scaling factors.

Table 12: Results of reduced state space inversion. Fluxes in Tg-CO_2 yr^{-1}.

Period	Region	Prior anthr	Post anth	Prior bio	Post bio
February 2018					
	PL	432 ± 43	440 ± 40	232 ± 116	168 ± 42
	DE	985 ± 98	879 ± 54	276 ± 138	436 ± 86

	rest	1659 ± 166	1574 ± 135	1356 ± 678	1546 ± 228
July 2018					
	PL	287 ± 29	300 ± 29	-578 ± 289	-672 ± 95
	DE	662 ± 66	676 ± 49	-751 ± 376	-481 ± 37
	rest	1104 ± 110	1173 ± 104	-4995 ± 2498	-5596 ± 397
February 2021					
	PL	441 ± 44	474 ± 44	242 ± 121	276 ± 72
	DE	924 ± 92	602 ± 22	312 ± 156	526 ± 62
	rest	1629 ± 163	1028 ± 70	1448 ± 724	1474 ± 107
July 2021					
	PL	292 ± 29	308 ± 30	-822 ± 411	-806 ± 139
	DE	591 ± 59	450 ± 20	-1148 ± 574	-890 ± 48
	rest	1062 ± 106	1131 ± 88	-5378 ± 2689	- 5131 ± 289

To conclude, it should be noted that the system presented here has been deployed for the first time and many of the questions raised in the presented run cannot be easily answered without further and deeper analysis. While the reductions of anthropogenic fluxes in Germany for 2018 at least seem to be corroborated by simulations from other groups (e.g. Sect. 3.3), the numbers cannot be directly compared due to limitations in temporal coverage. This is also true in the case of biogenic fluxes.

3.5.2.3 Summary

The results clearly demonstrate the importance of increased data availability to constrain anthropogenic fluxes over middle-sized countries like Germany and Poland. It is apparent from the results that without significant expansion of observational coverage, the increase of the accuracy of anthropogenic emissions cannot be obtained using currently available data. Without a dense surface observation network like ICOS, and no perspective to develop them quickly, satellite observations appear to be the only quick solution available to remedy the situation in Poland, but also in other countries in Eastern Europe - and beyond.

3.6 LUMIA for CO_2 inversions from in situ observations (LUND)

3.6.1 Model description

3.6.1.1 Transport model

LUMIA inversions rely on the FLEXPART model (version 10.4). FLEXPART is a Lagrangian particle dispersion model: it computes the dispersion (backward in time, in our case) of virtual air "particles", based on meteorological fields from mesoscale weather models and on internal parametrizations for turbulent convection. Our FLEXPART simulations were driven by hourly ECMWF ERA5 data, at a resolution of 0.25°.

In our setup, FLEXPART is used to compute "footprints" of the observations, i.e. vectors storing the sensitivity of each observation y to surface fluxes f. The model estimates of the CO_2 concentrations are then obtained with:

$$y^i{}_{mod} = y^i{}_{bg} + \sum_{cat} K^i f^{cat}$$

with $K^i = \partial y^i / \partial f$ the footprint corresponding to the observation y^i. The footprints are calculated at a horizontal resolution of 0.25°, for a regional domain ranging from 15°W, 33°N to 35°E, 73°N, and for a period of two weeks prior to each observation. The background term $y^{i_{bg}}$ results from the influence of the boundary condition, defined in space by the edges of the domain, and in time by the 14 days length of the footprint.

For calculating each footprint, the trajectory of 10000 particles was simulated in FLEXPART, released at the observation time and coordinates. For in-situ observations, the particles' release is distributed continuously over the integration time of the observation, i.e. an hour, for most sites. The particles are transported until they reach the edge of the domain, for a maximum period of 14 days. The sensitivity of the observation to the fluxes (i.e. the footprint) is inferred from the aggregated residence time of the particles in surface grid cells (defined as the layer below 100 m a.g.l). The final position of the particles (latitude, longitude, time, height above ground level) and the local pressure and orography are also stored, and used to interpolate the background concentration from the CAMS CO_2 concentration product.

3.6.1.2 Prior fluxes

Following the task protocol, fluxes in five categories were considered:

- Anthropogenic emissions: TNO product
- Biogenic CO2 fluxes: VPRM model
- Ocean: MERCATOR-OCEAN product
- Biomass burning: GFAS
- Lateral fluxes: F. Chevallier (LSCE) product (v3)

All fluxes were regridded on a 0.25° hourly resolution to be used in the inversion. When increasing the temporal resolution of the data (ocean and lateral fluxes products), the low-resolution fluxes were distributed homogeneously over the higher resolution time steps (i.e. no interpolation).

3.6.1.3 Assimilated observations

We assimilate observations from the ICOS ObsPack of ground-based CO2 observations (obspack_co2_466_GLOBALVIEWplus_v8.0_2023_03_30, https://doi.org/10.18160/CEC4-CAGK). All sites that had observations for the year 2018 were selected. When observations were available from several sampling heights, the highest one was systematically selected. An overview of the data used is provided in Table 13 below.

The observations to assimilate were further selected in time:

- For sites with ground altitude below 1000 m a.m.s.l, data between 13:00 and 18:00 local time were used. For these sites, the FLEXPART particle release (i.e. the model sampling) is done at the height above ground of the sampling height.
- For high-altitude sites, data between 00:00 and 5:00 were used. For these sites, the FLEXPART particle release is done at the altitude above sea level of the observations.

For each observation site, an estimate of the model representation error is calculated, based on the quality of the fit of the model to the short-term variability of the data. The procedure is the following:

1. An initial estimate of the concentrations is calculated based on the prior fluxes

2. The observations and the a priori concentrations are detrended by subtraction of their weekly moving average

3. The standard deviation of these detrended model-data mismatches is used as estimate for the model error.

The rationale behind this approach is that short-term variability of the modelled concentrations is largely a result of the modelled atmospheric dynamic, whereas long-term variability can result from errors in the prior fluxes, which the inversion should be able to resolve.

The resulting model error estimates are reported in Table 13 below, for each site. The observation uncertainty is then combined with the measurement error set to the value reported by the data providers (with a minimum of 0.3 ppm).

Table 13: Observation sites used in the inversions. Column 6 shows the site-specific model representation error used in the inversions.

	Name	latitude (˚N)	longitude (˚E)	surface altitude (m a.m.s.l)	sampling_height (m a.g.l)	model error (ppm)	dataset_name
BIK	Bialystok	53.23	23.01	183	300	04.01	co2_bik_tower-insitu_45_allvalid-300magl
BIS	Biscarrosse	44.38	-1.23	73	47	3.77	co2_bis_surface-insitu_11_allvalid
BRM	Beromunster	47.19	8.18	797	212	5.74	co2_brm_tower-insitu_49_allvalid-212magl
BSD	Bilsdale	54.36	-1.15	380	248	3.94	co2_bsd_tower-insitu_160_allvalid-248magl
CBW	Cabauw	51.97	4.93	0	207	5.97	co2_cbw_tower-insitu_445_allvalid-207magl
CMN	Monte Cimone	44.19	10.7	2165	8	03.04	co2_cmn_surface-insitu_443_allvalid
CRP	Carnsore Point	52.18	-6.37	9	14	3.24	co2_crp_surface-insitu_137_allvalid
ERS	Ersa	42.97	9.38	533	40	3.52	co2_ers_surface-insitu_11_allvalid
FKL	Finokalia	35.34	25.67	250	15	2.58	co2_fkl_surface-insitu_11_allvalid
GAT	Gartow	53.07	11.44	70	341	4.15	co2_gat_tower-insitu_147_allvalid-341magl
HEI	Heidelberg	49.42	8.68	113	30	9.23	co2_hei_surface-insitu_22_allvalid
HPB	Hohenpeissenberg	47.8	11.02	934	131	5.1	co2_hpb_tower-insitu_147_allvalid-131magl
HTM	Hyltemossa	56.1	13.42	115	150	3.27	co2_htm_tower-insitu_424_allvalid-150magl
HUN	Hegyhátsál háttérszennyettség-mérő állomás	46.96	16.65	248	115	5.34	co2_hun_tower-insitu_35_allvalid-115magl
IPR	Ispra	45.81	8.64	210	100	9.1	co2_ipr_tower-insitu_103_allvalid-100magl
JFJ	Jungfraujoch	46.55	7.99	3580	5	3.13	co2_jfj_surface-insitu_49_allvalid
KAS	Kasprowy Wierch	49.23	19.98	1987	7	05.05	co2_kas_surface-insitu_53_allvalid
KIT	Karlsruhe	49.09	8.42	110	200	8.67	co2_kit_tower-insitu_39_allvalid-200magl
KRE	Křešín u Pacova	49.57	15.08	534	250	4.14	co2_kre_tower-insitu_441_allvalid-250magl
LHW	Laegern-Hochwacht	47.48	8.4	840	32	5.59	co2_lhw_surface-insitu_5_allvalid

LIN	Lindenberg	52.17	14.12	73	98	09.01	co2_lin_tower-insitu_147_allvalid-98magl
LMP	Lampedusa	35.52	12.63	45	8	2.27	co2_lmp_surface-insitu_28_allvalid
LUT	Lutjewad	53.4	6.35	1	60	5.62	co2_lut_surface-insitu_467_allvalid
MHD	Mace Head	53.33	-9.9	5	24	2.67	co2_mhd_surface-insitu_11_allvalid
MLH	Malin Head	55.36	-7.33	22	47	3	co2_mlh_surface-insitu_468_allvalid
NOR	Norunda	60.09	17.48	46	100	3.38	co2_nor_tower-insitu_424_allvalid-100magl
OHP	Observatoire de Haute Provence	43.93	5.71	650	100	3.32	co2_ohp_tower-insitu_468_allvalid-100magl
OPE	Observatoire pérenne de l'environnement	48.56	5.5	390	120	5.53	co2_ope_tower-insitu_469_allvalid-120magl
PAL	Pallas	67.97	24.12	565	12	2.43	co2_pal_surface-insitu_30_allvalid
PDM	Pic du Midi	42.94	0.14	2877	28	2.51	co2_pdm_surface-insitu_11_allvalid
PRS	Plateau Rosa	45.93	7.7	3480	10	2.6	co2_prs_surface-insitu_21_allvalid
PUI	Puijo	62.91	27.65	232	84	4.35	co2_pui_tower-insitu_471_allvalid-84magl
PUY	Puy de Dôme	45.77	2.97	1465	10	3.83	co2_puy_surface-insitu_11_allvalid
RGL	Ridge Hill	52	-2.54	199	90	3.7	co2_rgl_tower-insitu_160_allvalid-90magl
SAC	Saclay	48.72	2.14	160	100	7.15	co2_sac_tower-insitu_11_allvalid-100magl
SMR	Hyytiälä	61.85	24.29	181	125	03.09	co2_smr_tower-insitu_421_allvalid-125magl
SVB	Svartberget	64.26	19.77	269	150	2.89	co2_svb_tower-insitu_440_allvalid-150magl
TAC	Tacolneston	52.52	1.14	56	185	4.87	co2_tac_tower-insitu_160_allvalid-185magl
TOH	Torfhaus	51.81	10.54	801	147	5.51	co2_toh_tower-insitu_147_allvalid-147magl
TRN	Trainou	47.96	2.11	131	180	4.95	co2_trn_tower-insitu_11_allvalid-180magl
UTO	Utö - Baltic sea	59.78	21.37	8	57	3.47	co2_uto_surface-insitu_30_allvalid
WAO	Weybourne	52.95	1.12	31	10	5.76	co2_wao_surface-insitu_13_allvalid
ZSF	Zugspitze	47.42	10.98	2666	3	3.37	co2_zsf_surface-insitu_25_allvalid

3.6.1.4 State vector

The inversions solve for daily offsets to the prior emissions, at the native resolution of the fluxes (0.25°):

$$f_{apos}^{cat} = f_{apri}^{cat} + TX^{apos}$$

with *X* the state vector, reshaped as a $(n_{opt}^{t}, n_{opt}^{p})$ matrix, and *T* is a $(n_{mod}^{t}, n_{opt}^{t})$ matrix, where $T(t^{mod}, t^{opt})$ is the fraction of the optimised interval t^{opt} that overlaps with the model time step t^{mod}. The size of the state vector is therefore $n_{opt}^{t} \times n_{cat}^{p} \times n_{cat}$, with:

- n_{opt}^{t} the number of time steps in the optimization (i.e. 365 for a one-year inversion),
- n_{cat}^{p} the number of grid cells for the category *cat* (grid cells where the emissions are not defined, e.g. biosphere fluxes in the ocean, are not included in the state vector).

- n_{cat} the number of optimised categories (1 or 2, depending on whether anthropogenic emissions are prescribed or optimised).

The prior error-covariance matrix is constructed with:

- exponential decay of the spatial correlations, with correlation length of 500 km for the biosphere fluxes, and 200 km for the anthropogenic fluxes (in the main experiments).
- exponential decay of the temporal correlations, with correlation lengths of 30 days for the biosphere fluxes, and either 7 days for the anthropogenic fluxes.
- variances are set proportional to the sum of the absolute value of the prior fluxes in each optimization interval. In other words, since the fluxes are transported hourly, and optimised daily, the variances are constructed from the sum of the absolute value of the hourly fluxes (for each grid cell and category).
- A category-specific scaling factor is applied to the variances, to ensure a total annual uncertainty of 0.5 PgC for biosphere fluxes, and 0.07 PgC (i.e. ~5% of the annual total) for anthropogenic emissions, at the continental scale. These values are in line with the reported differences in bottom-up inventories for Europe.

3.6.1.5 Period of study

The inversion covers the entire year of 2018, plus a spin-up period of one year at the beginning, and a spin-down period of one month at the end (i.e. it ranges from 1st December 2018 to 1st February 2019, but only data for 2018 are interpreted and included in the deliverable).

3.6.2 Results

Two main inversions were performed (plus additional sensitivity tests):

- In the **bio** inversion, only the biogenic emissions (VPRM prior) are optimised. This corresponds to our pre-existing LUMIA setup.
- In the **biofos** inversion, both the anthropogenic and biogenic emissions are adjusted by the inversion

3.6.2.1 Domain-scale results

Aggregated over the entire European domain, the **bio** inversion points to a reduction of the carbon uptake by the terrestrial ecosystems, from a net value of -0.83 PgC/year in the prior VPRM fluxes to a posterior flux of -0.57 PgC/year. The largest part of that flux correction occurs during the growth season, from April to the end of August. During the winter months, the inversion points to a reduction of the (positive) C flux, small in absolute terms (-2 TgC/day) but important in relative terms (~50% of the prior flux). This result is globally in line with previous results obtained with the LUMIA system (e.g. McGrath et al. 2023).

The **biofos** inversion, which also adjusts the anthropogenic emissions, leads to a relatively similar correction to the prior biosphere fluxes, but it sticks more to the prior during the winter. Over the year, this integrates to a near neutral net flux (-0.05 PgC/year). This is compensated by anthropogenic emissions, with a very strong (and clearly unrealistic) reduction of the flux estimate, from +1.39 PgC/year in the prior, to 0.86 PgC/year in the posterior.

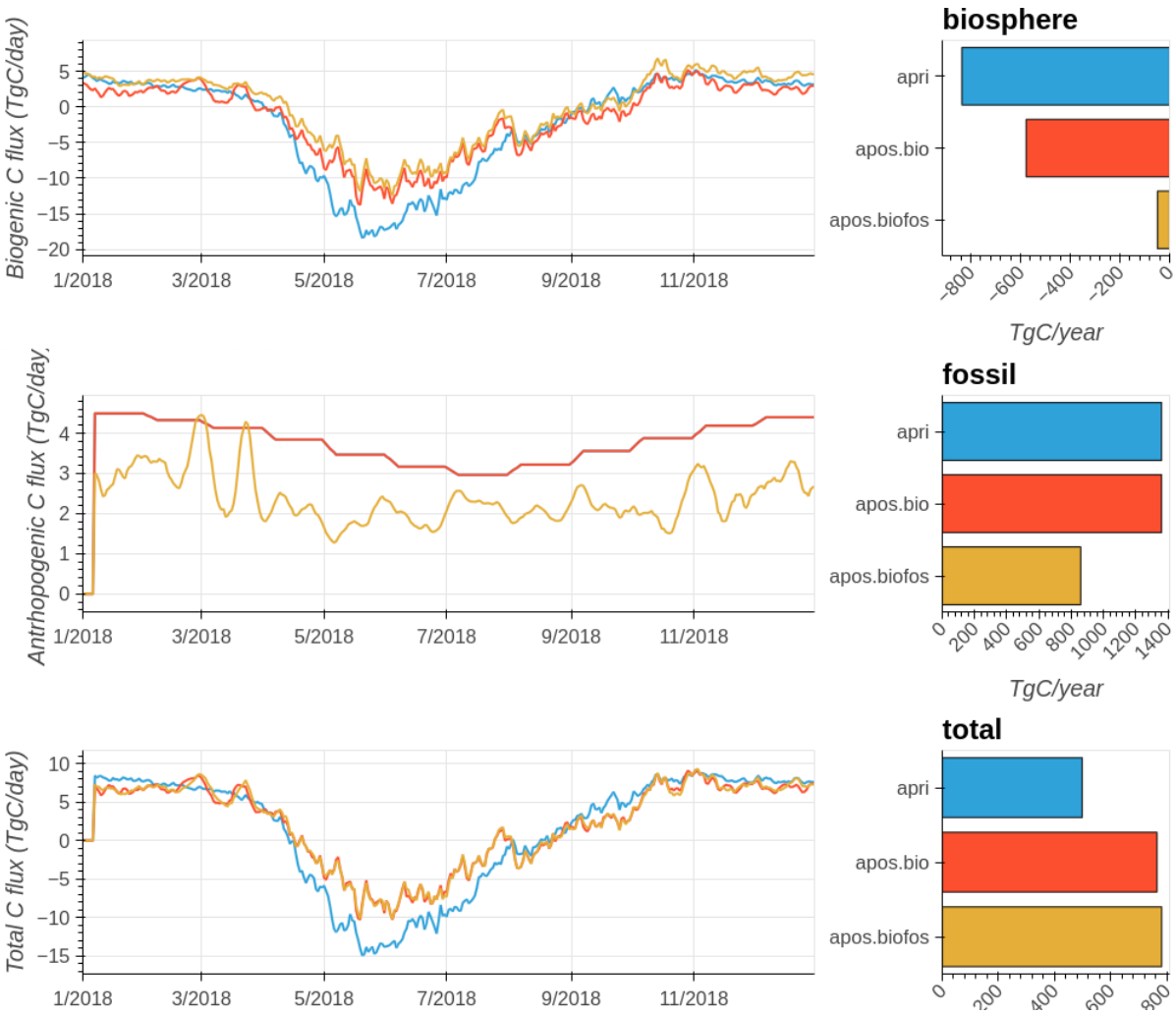


Figure 38: Prior (blue) and posterior fluxes, at the daily and annual scale, aggregated over the entire European domain.

The third row in Figure 38 shows the total of the biosphere and anthropogenic emissions: they are almost identical in **bio** and **biofos**: the two inversions differ essentially in how they distribute the flux correction across the two flux categories. Since anthropogenic and biosphere emissions are largely co-located (or at least, the observation network is not fine enough to be sensitive to differences in location), the **biofos** inversion relies mainly on the (prescribed) prior flux uncertainties to distribute the flux adjustments between the biosphere and anthropogenic categories. The prior uncertainties are shown in Figure 39. Although, overall, the uncertainty on the *biosphere* category is the largest (0.5 PgC/year vs. 0.07 PgC/year for the *anthropogenic* category), the uncertainty on anthropogenic emissions actually dominates in a band ranging from Southern UK to Western Poland. This also happens to be where the ICOS network is the densest, and explains why the **biofos** inversion seems to adjust in priority the anthropogenic emissions.

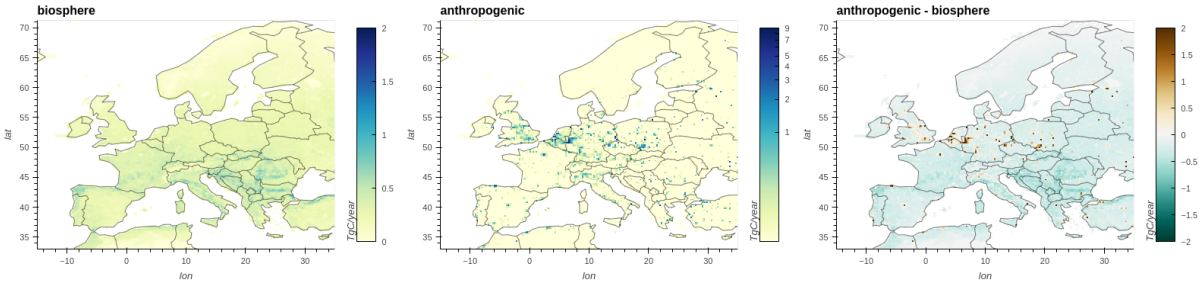


Figure 39: Prior uncertainties of the biosphere (left) and anthropogenic (middle) flux categories. The plot on the right shows the difference between the two, highlighting the areas dominated by uncertainties in anthropogenic emissions (brown) and those where biogenic emission uncertainty dominates (green).

3.6.2.2 Spatial distribution of the flux adjustments

Gridded maps of the prior fluxes and of the flux adjustments obtained in **bio** and **biofos** are shown in Figure 40. Essentially, it confirms the analysis of fluxes at the domain-scale: allowing the inversion to separately adjust the biogenic and anthropogenic emissions does not lead to significant differences in the map of total flux adjustments (lower-middle vs. lower-right plots), but leads to a reduction of the adjustment in the biosphere emission category. However, since the optimised emissions of the **biofos** emissions appear quite unrealistic, this category-specific result should not be interpreted further. The **bio** inversion seems to point to an overestimation of the C uptake in Southern Europe by the VPRM prior, and to an underestimation in North-East Europe.

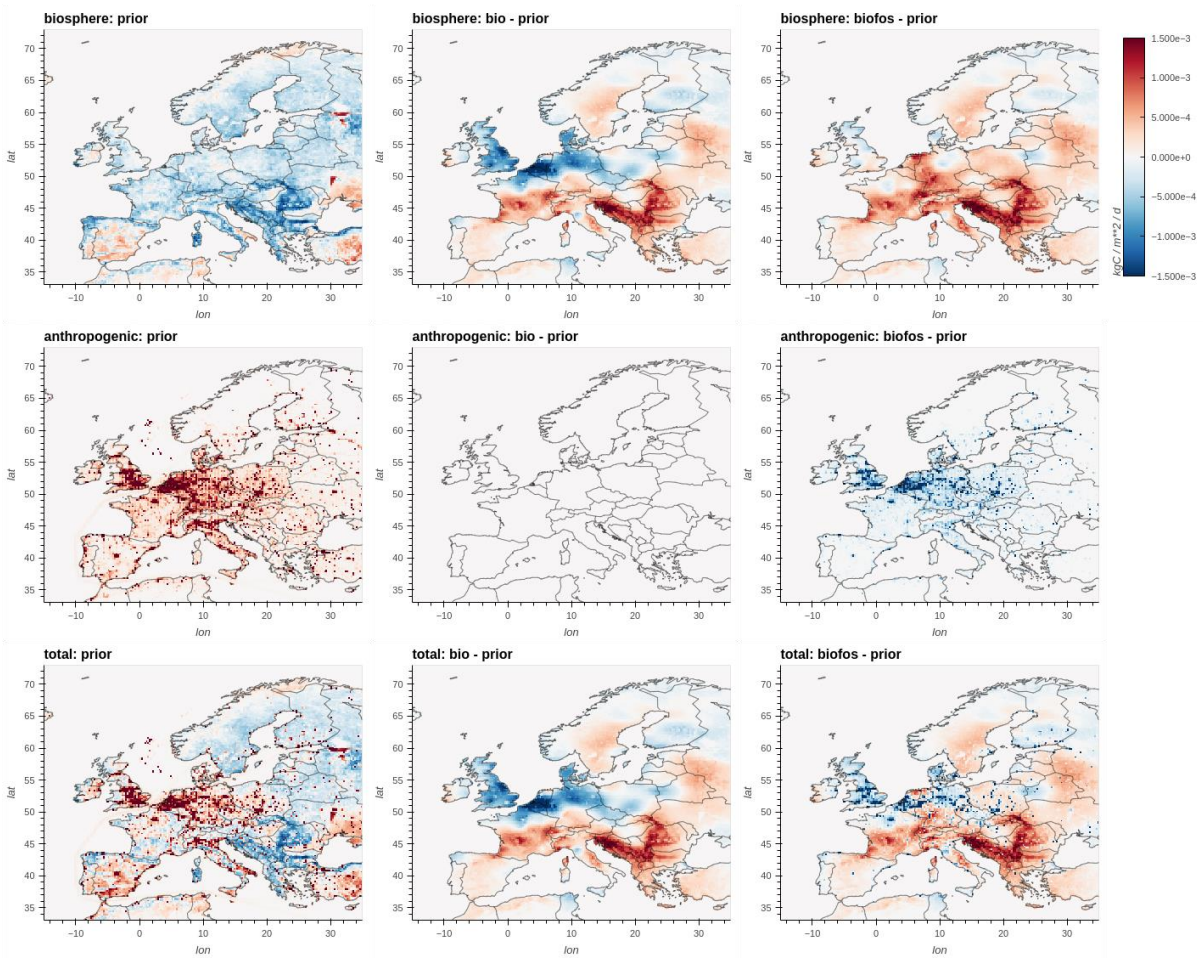


Figure 40: Prior fluxes (left) and optimized flux adjustments in the bio (middle) and biofos (right) inversions, for the biosphere (top) and anthropogenic (bottom) fluxes. The lower row shows the sum of anthropogenic and natural fluxes / flux adjustments.

3.6.2.3 Country-scale results

Given the poor performance of the **biofos** inversion, we did not interpret country-scale results much further. They are however included as part of the deliverable data, and summarised in Figure 41, for the countries with the largest prior fluxes. The daily country totals are also provided as one of the deliverable datasets (along with the gridded fluxes). The aggregation at this scale confirms that the total flux adjustment is more robust than the category-specific flux corrections.

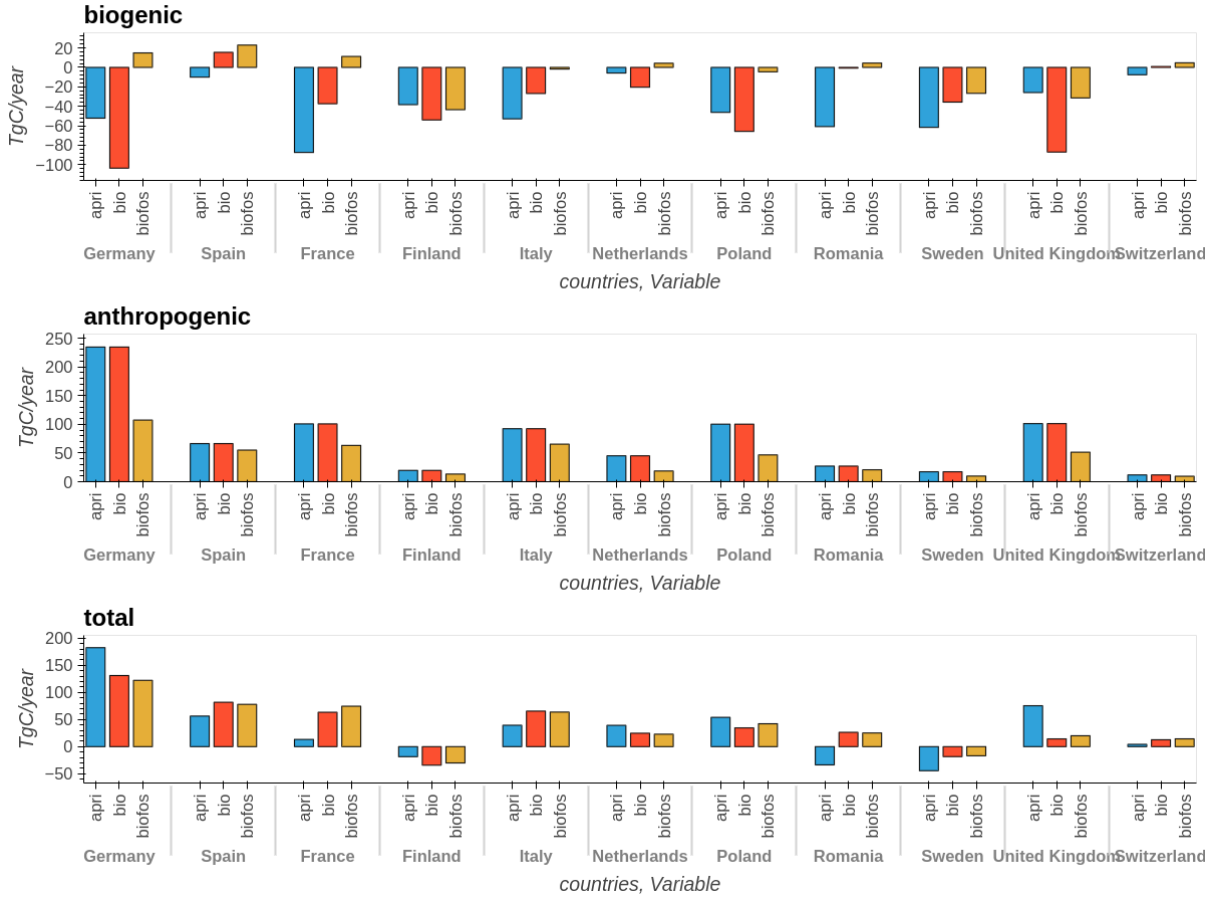


Figure 41: Prior and posterior fluxes for the biosphere (top) and anthropogenic (middle) fluxes, and the sum of the two (bottom).

3.6.2.4 Fit to the observations

The fit to observations is a classical diagnostic of atmospheric inversions. It fulfils essentially three purposes:

- verifying that the basic model settings are correct. In our case, we focus in particular on the quality of the background concentration interpolation, as it is a new development;
- verifying that the data selection is adequate, given the transport model performances;
- verifying that the inversions lead to an improvement in the fit to observations, and that these improvements are in line with the expectations (e.g., the **biofos** inversion has more degrees of freedom than the **bio** one, so it should lead to slightly better posterior fit to the data).

In order to compare the performance of the inversions, we computed, for each site the prior and posterior χ^2 fit to the assimilated data (which is favoured to e.g. the RMSE, as it accounts for observation uncertainties), as well as the mean biases (Figure 42).

The prior χ^2 are between 1 and 2. Both inversions are able to reduce the χ^2 at all sites, but perform less well at Ispra (IPR) and also, to some extent, at high-altitude sites (JFJ, KAS, PRS, CMN, PDM). Ispra is known to be a complex site to represent, and high-altitude sites are occasionally very sensitive to background concentrations, which are prescribed, and the model fit there is therefore more difficult to improve. The **biofos** inversion consistently leads to lower χ^2 values than the **bio** inversion, which is expected as it has additional degrees of freedom to fit the data.

The fact that the prior χ^2 are already so close to 1 could suggest the prior fit to the data is already close to optimal: either the observation uncertainties are in fact underestimated (which would raise the prior χ^2, and lead to posterior χ^2 closer to 1), or the inversions are, in fact, overfitting the data. Both the prior and observation uncertainties were set to reasonable values, therefore this overfitting could be solved by either reducing strongly the number of degrees of freedom of the inversions (e.g., by imposing much stronger covariances between the grid cells), or by increasing drastically not just the number of observations, but the observation coverage (e.g., by using satellite data).

The posterior mean biases are below 1 ppm at each station, except for Heidelberg (HEI), Ispra (IPR) and Lutjewad (LUT). In the case of HEI and LUT, the relatively poor posterior fit is due to the fact that these two sites have high observation uncertainties (Table 13), therefore they have a reduced weight in the inversion. A sensitivity test has been performed excluding the sites with the lower χ^2 reduction, but did not significantly improve the separation between anthropogenic and biogenic fluxes.

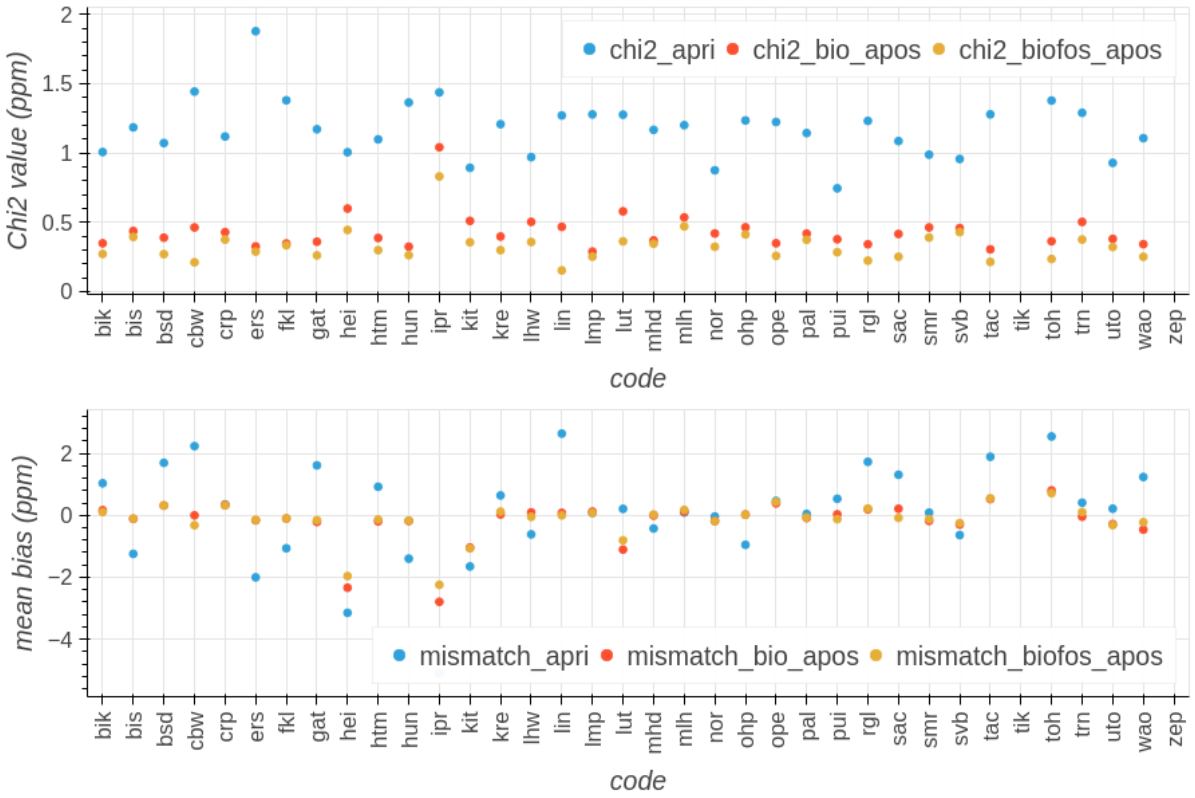


Figure 42: Prior and posterior χ^2 fit to the data (top) and biases (bottom) at each site

3.6.2.5 Background concentrations

In these inversions, the background concentrations were interpolated from the CAMS CO_2 inversions (FT18r1), using FLEXPART to transport the background concentration from the transport model domain edges (defined in space, for all the observations, and in time, for each observation) to the observation coordinates. Previous LUMIA inversions had relied on in-house global TM5-4DVAR CO_2 inversions, assimilating in-situ data distributed through the NOAA GlobalViewPlus obspacks. Significant differences exist between the two datasets of background concentrations, with a mean bias of 0.48 ppm, and a RMSE of 2.21 ppm. This difference is likely to have a significant impact on the results, however, it is impossible, at this stage, to assess the relative qualities of each background concentration dataset. It is technically possible to let the inversions adjust the background concentrations, however, we

choose not to do this at this stage, as it would make the results even more difficult to interpret. The impact of the boundary condition (and associated uncertainties) can instead be assessed through sensitivity experiments.

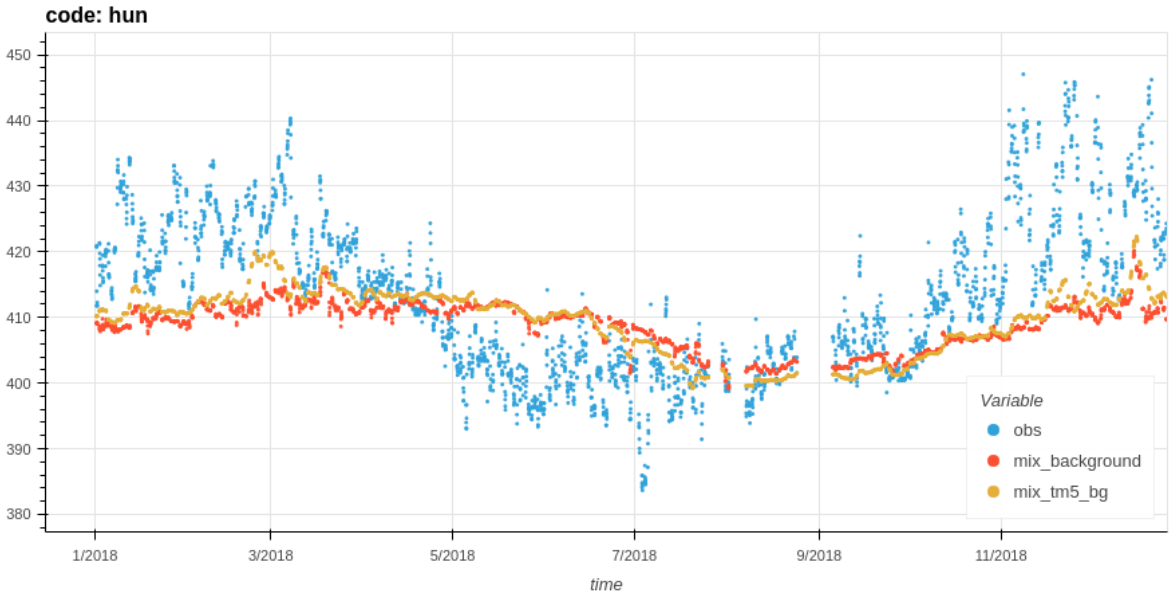


Figure 43: example of background concentrations, from CAMS (red, used in the inversions) and from TM5-4DVAR (yellow, not used) at Hegyatsal, Hungary.

3.6.2.6 Discussion

Many sensitivity tests were performed in order to try improving the results. In particular, we tested:

- increasing and reducing the observation uncertainties
- performing inversions with reduced observation site selections (excluding the sites that were the hardest to fit) or, on the contrary, with a larger quantity of observations (extended hour range for the assimilation periods)
- adjusting the covariance lengths (several configurations were tested, including much shorter spatial covariance lengths (down to 100 km), shorter or longer temporal covariance lengths (from 7 days to 30 days), for both flux categories)
- optimising separately GPP and respiration (in which case, the inversion had three categories: GPP, respiration and anthropogenic)
- reducing the uncertainties on the anthropogenic emissions prior to a near negligible value (0.00007 PgC, i.e. a factor 10 000 lower than in **biofos**)

Out of these, only the last experiment was able to avoid an unrealistic reduction of the anthropogenic emissions. The aim of this test was purely to verify that the results in **biofos** were not due to a malfunctioning of the inversion system, such a low uncertainty value for the anthropogenic emissions is not in line with the real uncertainties. It proves that more realistic country-scale results are achievable through a better specification of uncertainties. However, this displaces the problem from estimating the fluxes to estimating the prior uncertainties: the values used for both the prior biogenic and anthropogenic uncertainties are in line with those used in literature and in recent intercomparison experiments (Scholze et al., 2019, Monteil et al., 2020; Thompson et al., 2020; McGrath et al., 2023), and those uncertainties would need to be drastically modified in order to achieve a realistic category distribution of the posterior flux adjustments. In contrast, the total flux adjustment (anthropogenic + biogenic in our **biofos** inversion) is much less dependent on the prior uncertainty settings, which gives us a reasonable confidence that it can be interpreted scientifically.

The other pathway towards discriminating anthropogenic and biogenic fluxes would be the use of additional informational constraints. Those could come from satellite observations

(which LUMIA is currently unable to assimilate), or from the assimilation of co-emitted tracers, e.g. radiocarbon ($^{14}CO_2$), which was recently implemented in LUMIA (https://egusphere.copernicus.org/preprints/2023/egusphere-2023-2215/).

Finally, allowing the inversion to optimise the background concentrations would be a logical step, given the differences between the two background concentration datasets at our disposal, and it would likely reduce the tendency of the inversion to adjust the anthropogenic and biogenic emissions, but there is a risk that the amplitude of the adjustment in optimised fluxes would become highly dependent on the uncertainty attributed to the background concentrations.

3.7 TRACE for OSSEs with CO_2 inversions from pseudo in situ observations (LUND)

3.7.1 Model description

3.7.1.1 Transport model

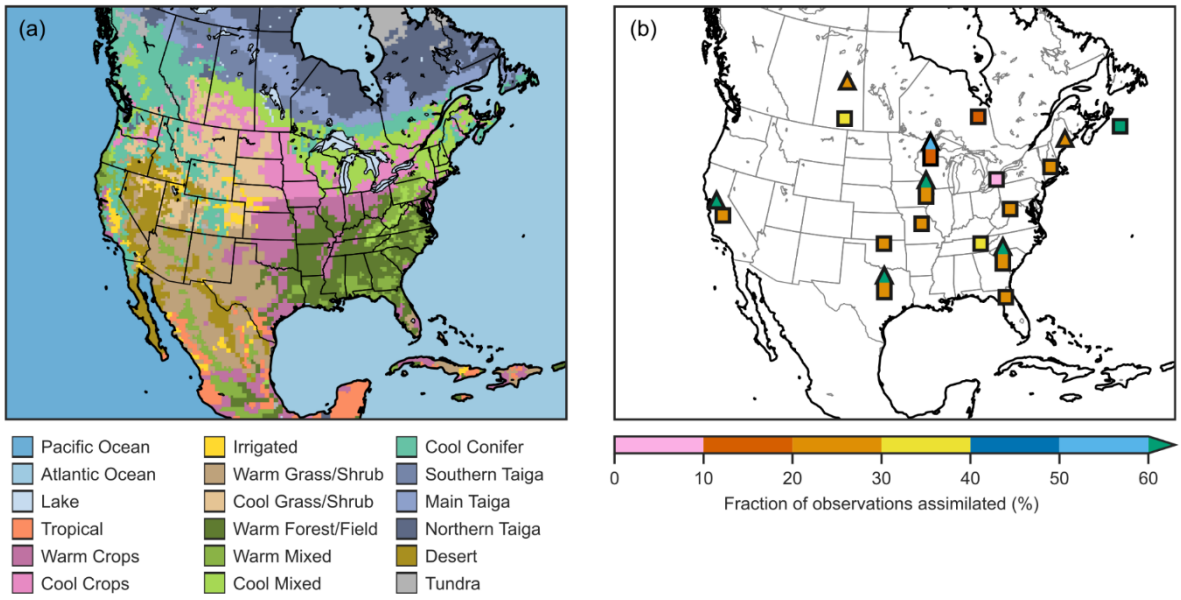


Figure 44: (a) Model domain and ecoregions for the inversion experiments with the TRACE system. (b) Location of ground-based CO_2 observations. Triangles denote tall towers (≥100 m height above the ground) and squares denote short towers. A combined triangle and square denotes a tower with multiple intake heights, and the tower location is then indicated by the lower square. The colours show the fractions of observations assimilated, which were determined based on when the observations were determined to be representative of well-mixed conditions.

The TRACE Regional Atmosphere–Carbon Ensemble (TRACE) system uses the online WRF model to dynamically downscale meteorological conditions and simulate the transport of CO_2 (and potentially other species). A description of WRF is given in Section 3.11.4.1. TRACE is currently configured to use WRF version 3.6.1. The modeling of atmospheric CO_2 transport is carried out by the WRF-Chem module, which has been modified to enable large ensembles of tracer simulations within a single WRF run, making it computationally efficient to perform ensemble runs with, e.g., perturbed flux fields.

Here, WRF was set up to run for a single domain covering a large portion of North America (Figure 44a) at a horizontal resolution of 27 km and 60 vertical levels spanning up to 50 hPa, with denser vertical levels closer to the surface and increasing distances between levels higher up in the atmosphere. The time step was set to 60 seconds. Physical parameterizations include a positive definite sixth-order diffusion scheme, the Noah land-surface model, the

Mellor-Yamada Nakanishi and Niino Level 2.5 boundary layer scheme, and the Kain-Fritsch convective scheme.

For the experiments reported here, we used meteorological initial and boundary conditions from the ERA-Interim reanalysis (81 km resolution and 60 vertical levels), and CO_2 mole fractions from the CarbonTracker CT2019B dataset (3° × 2° longitude-latitude and 25 vertical levels). TRACE has also been updated to run using the ERA5 reanalysis and CO_2 from CAMS, however, these inversions have not been finalized yet. We first carried out observing system simulation experiments (OSSEs) to rigorously test the system, which are also reported in Chen et al. (2023).

3.7.1.2 Prior fluxes and assimilated observations

The prior fluxes were obtained from the downscaled 3-hourly prior fluxes in CT2019B (1° × 1° horizontal resolution). The prior biogenic fluxes come from a simulation with the Carnegie-Ames Stanford Approach (CASA) model called CASA GFED 4.1s, and the oceanic fluxes from the Takahashi et al. (2009) climatology. The results here focus on these two flux components, as they generally have the largest influence on the CO_2 mole fraction variations in the domain. A test experiment was performed using prior anthropogenic emissions from the Miller fossil fuel emission inventory, but the ground-based CO_2 observation network used in the OSSEs did not show much sensitivity to the anthropogenic emissions, thus these results are not included.

In OSSEs, a reference run is taken as the “truth” and synthetic observations can then be derived from this reference run. A big advantage of OSSEs is that the truth is known, making it possible to evaluate the inversion results directly against the true fluxes. Here, the true fluxes were taken from different datasets than the prior fluxes, namely the CASA GFED_CMS simulation for biogenic fluxes, and Ocean Inversion Fluxes for oceanic fluxes, both provided by CT2019B. All fluxes were interpolated to the WRF grid using a first-order conservative regridding scheme.

The observations were derived based on the location of ground-based in situ tower observations within the domain, mainly from NOAA and Environment Canada, see Figure 44b. Random Gaussian noise was added to the synthetic observations with a mean of 0 and standard deviation of 1 ppm to simulate observation errors.

3.7.1.3 State vector

The state vector in TRACE consists of both atmospheric CO_2 mole fractions and, through state augmentation, parameters controlling CO_2 fluxes at every model grid point (and all vertical levels for CO_2 mole fractions), similar to the ICON inversion system (Section 3.4). Different flux parameters are applied for the different flux components (anthropogenic emissions, biogenic fluxes, and oceanic fluxes). TRACE can also include meteorological fields in the state vector, which makes it possible to optimize meteorological conditions and atmospheric transport by assimilating conventional weather observations, but this functionality was not tested here. The main purposes of coupling the atmospheric and carbon flux states are to (1) make it computationally efficient to assimilate a large number of observations, including column-integrated satellite observations; and (2) add the possibility to account for correlated atmospheric transport errors in the simulated CO_2 mole fractions (see Chen et al., 2019).

CO_2 mole fractions and flux parameters are jointly optimized in the inversion through ensemble-based simultaneous state and parameter estimation (ESSPE). The ESSPE

approach in TRACE is based on a square-root formulation of the ensemble Kalman Filter (EnKF). In the original Kalman filter, the state vector is updated according to:

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}\left(\mathbf{y}^o - \mathbf{H}\left(\mathbf{x}^b\right)\right),$$

where $\mathbf{x}^a$ and $\mathbf{x}^b$ are the posterior and prior state vectors, respectively, $\mathbf{y}^o$ is a vector of observations, **H** is the linearized observation operator, and **K** is the Kalman gain:

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^\top\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^\top + \mathbf{R}\right)^{-1},$$

where $\mathbf{P}^b$ and **R** are the error covariance matrices for the prior and observations, respectively. The posterior error covariance matrix $\mathbf{P}^a$ is then given by:

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b,$$

where **I** is the identity matrix. After the assimilation step, the error covariance matrix is propagated forward in time using the dynamic model.

The Kalman filter is computationally prohibitive to apply for high-dimensional systems due to the need to propagate the full covariance matrix. The EnKF solves this by estimating the error covariance matrix with an ensemble representation:

$$\mathbf{P}^b = \frac{1}{N-1}\mathbf{X}'\mathbf{X}'^\top,$$

where *N* is the number of ensemble members and **X**' is a matrix consisting of ensemble perturbations (deviations from the ensemble mean) in its columns. The experiments here used *N*=40 members with perturbed flux parameters.

Instead of computing the full Kalman gain matrix, which involves inverting a *p* x *p* matrix (where *p* is the number of observations), the square-root filter solves the update equation by assimilating observations sequentially. This assumes that there are no error correlations between observations, i.e., that **R** is diagonal. In this case, the matrix inversion becomes a scalar division. The ensemble mean is updated according to the original update equation, while the ensemble perturbations **x**' (with superscript *a* for posterior and *b* for prior) are also updated in each iteration to account for the reduced uncertainty from assimilating previous observations:

$$\mathbf{x}'^a = \mathbf{x}'^b - \tilde{\mathbf{K}}\mathbf{H}\mathbf{x}'^b,$$

where $\tilde{\mathbf{K}}$ is the reduced Kalman gain matrix:

$$\tilde{\mathbf{K}} = \alpha\mathbf{K}, \quad \alpha = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^\top + \mathbf{R}}}\right)^{-1}.$$

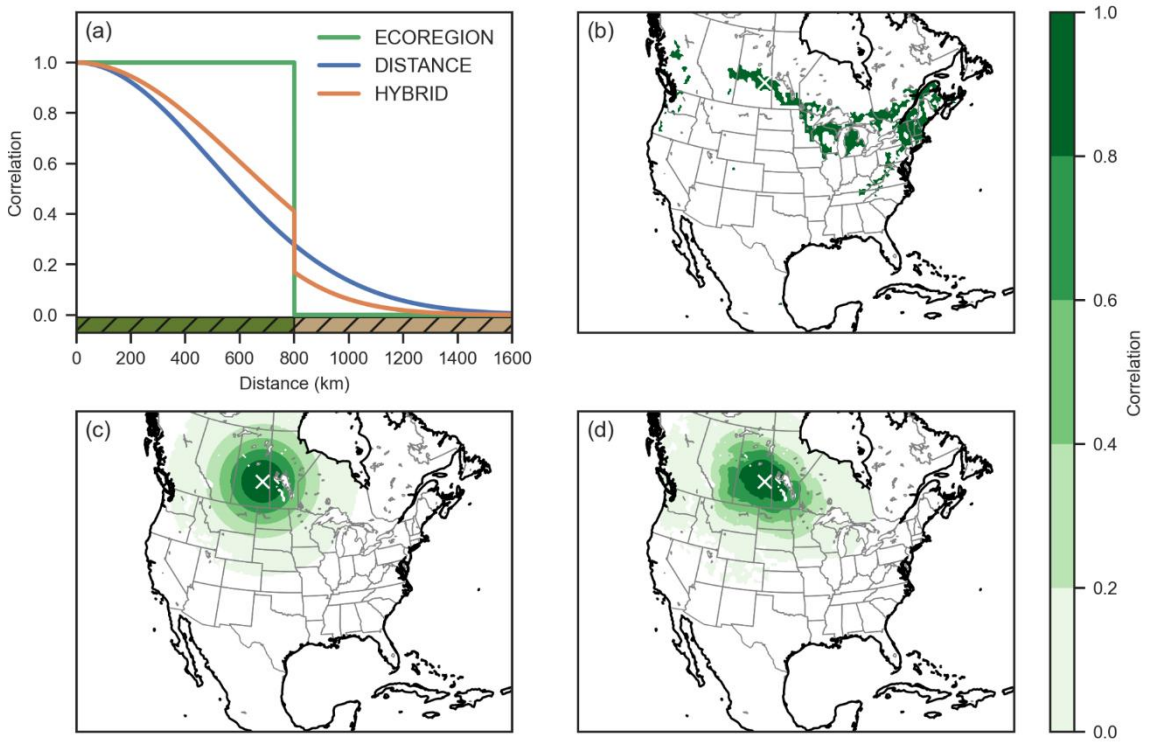


Figure 45: (a) Variations in correlations between flux parameter errors with distance and between ecoregions for different correlation functions. The green and brown shading at the bottom indicate two different ecoregions. (b–d) Examples of spatial flux error correlations between the grid point indicated by the white cross and all other grid points for the (b) ECOREGION, (c) DISTANCE, and (d) HYBRID correlation functions.

In the forecast step, the atmospheric state, including CO_2 mole fractions, are propagated forward using the WRF model. The flux parameters are assumed to be constant scaling factors for the prior fluxes (40% uncertainty for biogenic fluxes and 80% uncertainty for oceanic fluxes). The dynamic model for the flux parameters is a persistence model. Because we do not know the true prior error covariances for the flux parameters, three experiments were carried out with different spatial correlations in flux parameter errors: (1) Flux parameter errors that are 100% correlated within ecoregions and no correlation between ecoregions, denoted as ECOREGION. This formulation is similar to parameters that vary with Plant Functional Types in vegetation models. (2) An error correlation that decays with distance according to a Gaussian function with a length scale of L=500 km, denoted DISTANCE. This is similar to what is used in many modern inversion systems. (3) A hybrid approach that takes into account both distance and ecoregions, denoted HYBRID. Figure 45 shows the error correlation functions and examples of how they can vary spatially.

A horizontal localization scheme was applied to the *in situ* CO_2 observations following a Gaspari and Cohn fifth-order polynomial with a radius of influence of 5,400 km. The posterior ensemble spread was relaxed to the prior spread by 20% to avoid filter divergence. Additionally, the posterior ensemble spread in the flux parameters was relaxed to 10% of the prescribed initial parameter perturbations to ensure there is a spread in the flux parameters, as there is no inherent error growth in the persistence model. Further details about the TRACE data assimilation system can be found in Chen et al. (2023).

3.7.1.4 Period of study

The TRACE experiments here focus on North America in July 2016, which is different from the other inversion systems in this report. Given that these experiments are OSSEs, the results should not be compared with real-data experiments, as the results depend on our assumed truth, and do not necessarily reflect reality. The purpose of these experiments is to evaluate if the new ESSPE approach is appropriate for CO_2 inversions and if there are any issues due to, e.g., mass balance and short assimilation windows.

3.7.2 Results

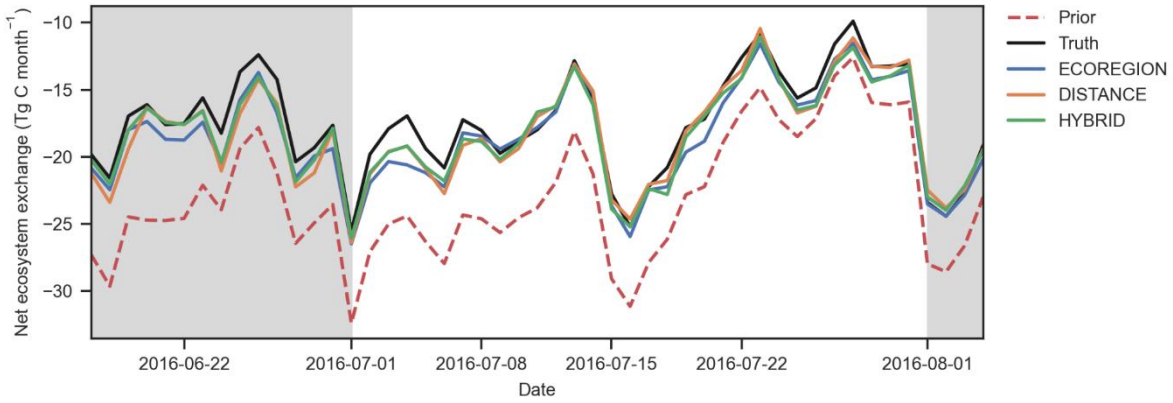


Figure 46: Domain-integrated daily net CO_2 fluxes for the true, prior fluxes, and posterior fluxes. The different inversion experiments used different prior flux parameter error correlations. The white background highlights the study period.

Figure 46 shows a time series of the net CO_2 fluxes in the domain for the study period. The inversion experiments were spun up for 2 weeks without assimilating observations to establish the links between perturbed CO_2 flux parameters and atmospheric CO_2 mole fractions, and then 2 additional weeks with assimilation. The first 4 weeks were then discarded as spin-up. A 1-week assimilation window was used, which meant that the experiments ended on 5 August 2016. The days in August were similarly discarded to focus on the month of July.

The posterior fluxes in all inversion experiments show a closer fit to the true fluxes compared with the prior fluxes, which overestimated the carbon sink over North America. There are some days when the posterior fluxes still had a too strong sink, for example in the beginning of July, but overall the inversions were able to constrain the time evolution of the domain-wide net fluxes well.

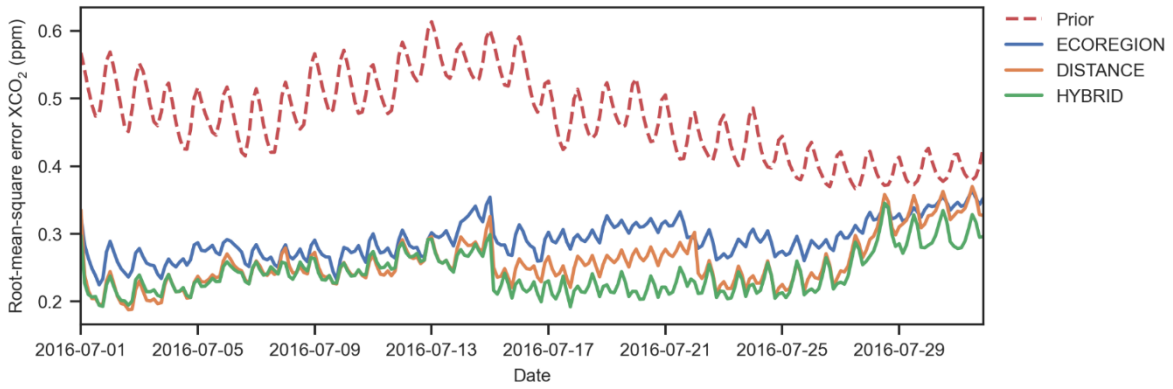


Figure 47: Domain-integrated root-mean-square error in column-integrated dry air mole fraction of CO_2 (XCO_2) in the simulations run using the prior and posterior CO_2 fluxes.

A common worry about jointly optimizing CO_2 concentration and fluxes, also known as dual-state inversions, is that the posterior CO_2 fluxes and/or the concentrations turn out to be unrealistic. In our OSSEs we can directly evaluate both the posterior fluxes and concentrations over the whole domain against the truth. Figure 47 shows the time evolution of the root-mean-square error (RMSE) in column-integrated CO_2 (XCO_2) over the whole domain (excluding a margin of 20 grid points around the edge) in the simulations using prior and posterior fluxes. By assimilating the ground-based CO_2 observations and optimizing the CO_2 concentration and fluxes, the RMSE in domain-integrated XCO_2 decreased from around 0.47 ppm to 0.24–0.29 ppm, with the largest RMSE decreases in the HYBRID experiment and smallest in ECOREGION. Thus, the joint optimization was able to reduce the RMSE in simulated CO_2 concentrations in the regional domain by about 50%.

To evaluate the spatial patterns of the fluxes, Figure 48 shows the monthly mean net CO_2 fluxes over the whole domain for the prior, true, and posterior fluxes. Compared with the true fluxes, the prior underestimates the carbon sink in the northwestern part of the domain and around the Corn Belt in the United States, and simulates a too strong carbon sink in most other parts of the domain, especially in the southeastern United States. The posterior fluxes from the different experiments all show realistic patterns and are able to overall reduce the mismatch to the true fluxes. The ECOREGION experiment shows some large deviations from the truth—some even larger than the mismatch between the prior and truth—especially in the southern part of the Corn Belt, which can be attributed to incorrect upscaling of the flux corrections through the ecoregion-shaped flux parameter error correlations. The differences between the truth and posterior fluxes from the DISTANCE and HYBRID inversions tend to be smaller in both magnitude and spatial scale. All inversions were able to increase the carbon sink around the Corn Belt where the observation coverage was relatively dense.

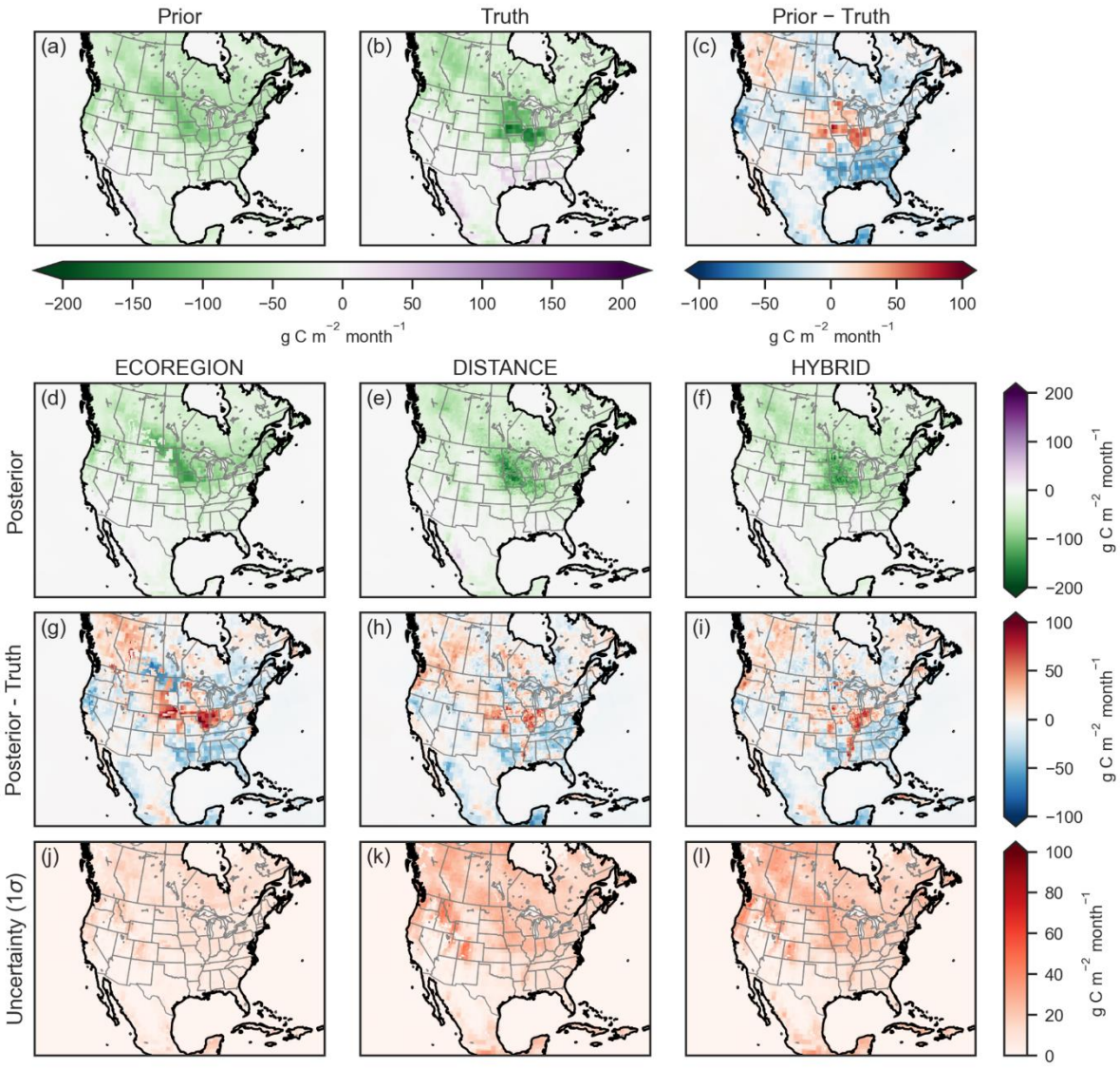


Figure 48: Spatial distribution of monthly averaged net CO_2 fluxes for the study period. (a) Prior fluxes, (b) true fluxes, and (c) differences between the prior and true fluxes. (d–f) Posterior fluxes, (g–i) differences between the posterior and true fluxes, and (j–l) estimated uncertainties (1σ) for the posterior fluxes for the ECOREGION (first column), DISTANCE (second column), and HYBRID (third column) experiments.

The ensemble-based TRACE system produces not only posterior CO_2 fluxes after assimilating atmospheric observations, but also uncertainty estimates for these fluxes. The last row of Figure 48 shows the estimated posterior uncertainties (one ensemble standard deviation) at the grid point scale, which can be compared with the differences between posterior and true fluxes in the previous row. The ECOREGION inversion appears to generally underestimate the uncertainties, likely due to the strong constraint provided by the ecoregion-specific flux parameters (100% correlated flux parameter errors within ecoregions), which in this case does not reflect the true flux parameter errors. In contrast, the DISTANCE and HYBRID experiments produced posterior uncertainties that are generally consistent with the true errors, except for the southern part of the United States, where the uncertainties were slightly underestimated.

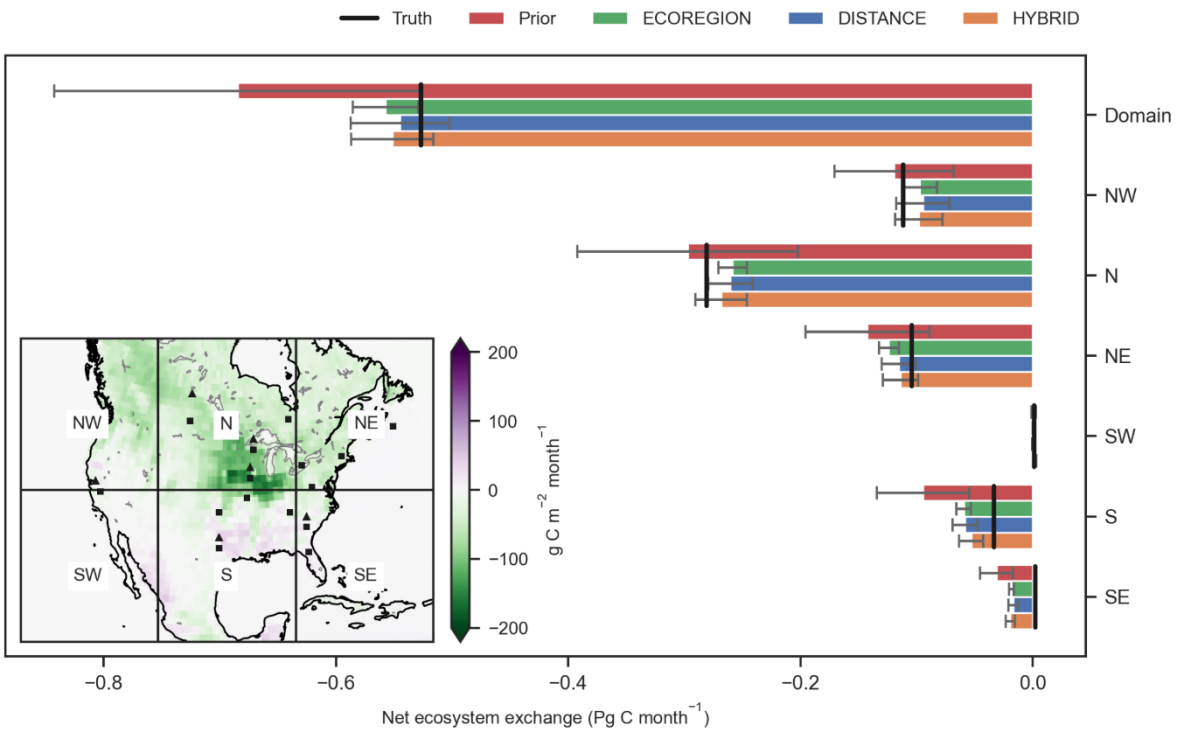


Figure 49: Domain-integrated net CO_2 flux and regional net CO_2 fluxes for the subregions shown in the inset: northwestern (NW), northern (N), northeastern (NE), southwestern (SW), southern (S), and southeastern (SE) part of the domain. The true values for the fluxes are shown as black vertical lines. The error bars represent the flux uncertainty derived from the ensemble spread (2σ). The shading in the inset shows the true fluxes, and the squares and triangles display the location of the ground-based observations.

Figure 49 summarizes the results in terms of regional carbon budgets for the month of July 2016. The domain was divided into six roughly equally-sized subregions, and the net CO_2 fluxes were calculated for each subregion for the true, prior, and posterior fluxes. In terms of the domain-wide CO_2 emissions and uptake, all inversions drastically improved the budget by reducing the overestimated carbon sink in the prior fluxes, consistent with what Figure 46 showed. Moreover, uncertainties on the posterior fluxes are about 4 times smaller than the prior flux uncertainties, reflecting an increased confidence in the flux estimates after assimilating the CO_2 concentration observations, consistent with the true errors.

When focusing on the carbon budgets for the subregions, the picture becomes a bit less clear. Here, the prior fluxes are close to the truth in some subregions, especially in the northwestern (NW) and northern (N) part of the domain, while the inversions underestimate the sinks. Further investigations show that the degradation in the posterior fluxes in the NW region occur because the inversions were able to partially reduce the overestimated CO_2 sink in northern California and southern Oregon (compare Figure 48c and Figure 48g-i), while the underestimated carbon sink in the northwesternmost part of the domain where observations are sparse remained in the posterior fluxes. Thus, although the inversions improved the fluxes at the grid point scale, these improvements led to an overall degradation of the net fluxes integrated over the NW subregion. One should therefore be careful when evaluating regional carbon budgets, as the results can be sensitive to how aggregation is done (i.e., how the regions are defined). With that said, when considering all regional carbon budgets, for example by summing the absolute differences between the prior/posterior and true fluxes, the inversions all led to smaller total absolute errors in regional budgets: 103.3 Tg C month^{-1} for ECOREGION, 92.6 Tg C month^{-1} for DISTANCE, and 77.5 Tg C month^{-1} for HYBRID, compared with 157.7 Tg C month^{-1} for the prior fluxes. The best-performing HYBRID inversion

thus improved upon the prior fluxes by more than 50% in terms of regional budgets for the subregions shown in Figure 49. At the continental scale, all inversions show robust error reductions compared with the prior, with 80.6%, 88.5% and 84.4% relative error reductions for the ECOREGION, DISTANCE, and HYBRID experiments, respectively.

Consistent with previous results, the DISTANCE and HYBRID inversions produced uncertainty estimates on region carbon budgets that generally match the true errors, except for in the southern (S) and southeastern (SE) subregions, where the prior uncertainties also underestimated the true errors. The ECOREGION inversion generally produced too small uncertainties on the posterior fluxes. Inversions using such formulations for the flux (parameter) error correlations should therefore consider additional inflation methods to avoid underestimating the flux errors.

To conclude, it is important to remember that the OSSEs here are idealistic and do not include all error sources. The results reported here should therefore be considered optimistic and do not reflect the expected performance from real-data inversions. Nevertheless, the OSSEs show that the dual-state CO_2 flux and concentration optimization in TRACE using the ESSPE approach is a promising framework to further develop novel CO_2 inversion algorithms.The OSSEs can gradually be made more realistic by including additional error sources, such as atmospheric transport errors. One strength of the TRACE system is that atmospheric transport errors can be simulated using the online weather model by perturbing meteorological conditions and then assimilating weather observations (see Chen et al. 2019). Additional errors, for example due to uncertain CO_2 lateral boundary conditions, can be included in a straightforward manner in the ensemble system by introducing appropriate perturbations to the ensemble members. Next steps include extending the TRACE system to satellite observations and increasing the model resolution, which hopefully increases the sensitivity to fossil fuel emissions. The system is already capable of assimilating XCO_2 observations (not shown), and further experiments will focus on making the XCO_2 observations more realistic, for example to mimic future measurements from CO2M.

3.8 LOTOS-EUROS for CH_4 inversions from in situ observations (TNO)

The LOTOS-EUROS inversion system is based on a 4D-Var technique and aims at estimating gridded emission fields at European scale. The system inherits the framework used for the CAMS global CH_4 emission inversion that is used in combination with the global TM5 model. As the system has been developed mainly within this project, the current application is limited to CH_4 only, although tests have been conducted including CO_2 too.

3.8.1 Model description

3.8.1.1 Transport model

The LOTOS-EUROS model (Manders et al., 2017) is a regional chemistry transport model (CTM) driven by offline meteorological data. The main application is in full chemistry simulations including trace gases and aerosols related to air quality and deposition, but the model could also be configured to simulate concentrations of greenhouse gases.

In this study, the model domain is defined following the “Protocol for the intercomparison of national CH_4 emissions” that was developed under T5.6 (Scholze *et al.,* 2023), hereafter named *Inversion Protocol*. A cartesian longitude-latitude grid over Europe is used with a rectangular domain over [15°W,35°E] longitude and [35°N,70°N] latitude. The grid resolution is 0.5°x0.5° which leads to a total of 100 x 70 grid cells, with size of about 25 km x 50 km at mid-European latitudes. In the vertical 12 layers are used that cover the troposphere up to about 200 hPa; the layers are a coarsening of the ECMWF meteorological model layers, with the lowest layer about 20 m thick.

The meteorological data are obtained from ECMWF from the operational daily forecasts over 0-12 hours; surface fields are obtained at hourly resolution, while 3D fields are retrieved at 3 hourly resolution.

In this study, the model is configured to simulate CH_4 only. The only processes required are emissions (described below), transport (advection and vertical mixing), and chemical sink. For the sinks the model reads concentration fields of the hydroxyl radical (OH) that has been saved from a previously performed simulation with full chemistry configuration (all trace gases and aerosols related to air quality). In the inversion framework, boundary conditions are set to zero since from the observations the contribution from the global background will be removed.

For use in inversion systems, the model could read time series of 2D parameter fields to perturb the model data, in this case the emission fields. Specific for 4D-var methods the model could run in adjoint mode to compute the sensitivity of observation simulations towards changes in the parameter fields.

3.8.1.2 Prior fluxes and assimilated observations

The prior fluxes have been provided by the T5.6 *Inversion Protocol*. These include fossil fuel, agricultural, and waste treatment emissions from EDGAR v6.0, peatlands and soil uptake from JSBACH-HIMMELI, biomass burning from GFED, and other smaller sources.

Observations of CH_4 have been taken from the *Inversion Protocol* too. The data consists of a set that should only be used for validation only, and a set from which (selected) observations could be analysed. Only observations from non-elevated locations (below 1500 m above sea level) are used, and only if the local-standard time (longitude based) is within [12:00,16:00]. This ensures that the observations used in the inversion are sampled from well-mixed air volumes that are not strongly influenced by local sources. Observation errors are taken from the values in the data set (if provided), with a minimum of 3 ppb. At the moment this is the only contribution to the *observation representation error* that is taken into account; contributions from the spatial, vertical, and temporal sampling from the model are not included yet.

The inversion system will try to optimise the local emission only. Therefore, the contribution of the global background has been subtracted from the observations. These *baseline* concentrations were provided to the *Inversion Protocol* by the CAMS global CH_4 emission inversion system following the (Rödenbeck, 2009) approach. An example of these *baselines* is shown in Figure 50. In this way, boundary conditions for the model could be set to zero as the contribution from the global background is taken into account via the baselines.

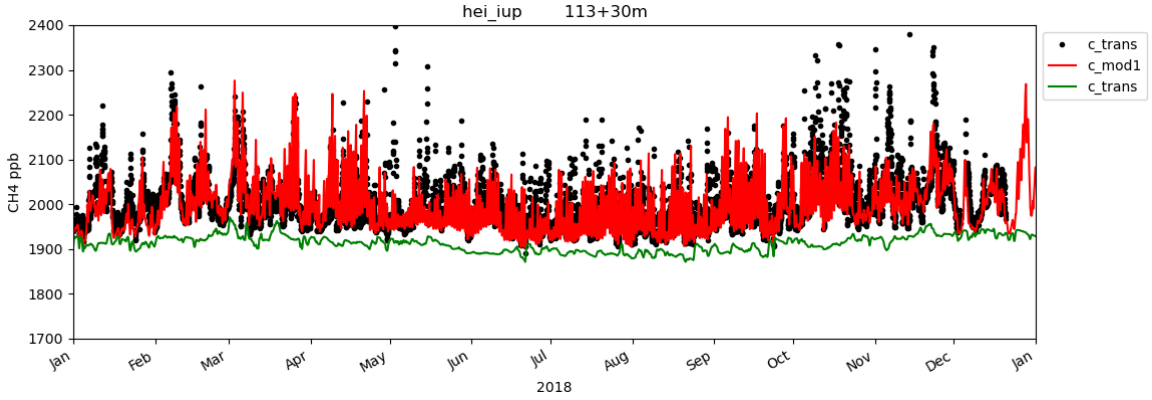


Figure 50: Example of CH_4 baseline concentrations (green) for observations (black) in Heidelberg (Germany, Institut für Umweltphysik). The red line shows the posterior simulation by the TM5 global inversion system that was used to compute the baselines.

3.8.1.3 State vector

The state vector **x** for the inversion consist of emission deviation factors *x* per grid cell and month that are applied to the emissions following:

$$E(i,j,t) = E^b(i,j,t)\,[\,1 + x(i,j,t)\,]$$

where E^b denotes the prior emissions, *E* denotes the actual emissions, *i,j* denotes the grid cell, and *t* denotes the time (month). The prior estimate $\mathbf{x^b}$ is zero. The 4D-Var inversion system optimises **x** by minimising the cost function:

$$J(x) = 1/2\,(x - x^b)^T B^{-1}(x - x^b) + 1/2\,(H(x) - y)^T R^{-1}(H(x) - y)$$

Here, **B** is the background covariance that describes how the state elements are allowed to vary from their prior estimate. In the current application, the elements of *x* are given a standard deviation of 0.5, equivalent to a relative uncertainty of 50% in the emissions. The spatial correlation between the emissions factors is parameterized with a Gaussian decay with a length scale *L*=100 km, and the temporal correlation follows an exponential decay with a time scale of τ=1 month. The observations ***y*** contain the above described observation selection with the baseline concentrations removed. Observation operator ***H(x)*** simulates the observations for a given state vector using the LOTOS-EUROS simulation model. The observation representation error covariance ***R*** contains the above described observation uncertainty, with no correlations between observations.

The cost function is minimised in an iterative procedure, where each iteration consists of one forward and one adjoint simulation with the model. Since the model includes a single tracer only, a full minimization could be performed in a reasonable time period of about a week.

3.8.1.4 Period of study

The current study applies the inversion system to 2018 only.

3.8.2 Results

For a first impression of the inversion results, Figure 51 shows maps of the CH_4 emissions over 2018. The left panel shows the *a priori* emissions, with the high emission regions over Netherlands, northern Italy, and southern Poland. The right panel shows the analysis increment obtained from the inversion, derived by combining the optimal state *x* with the *a prori* emissions. The markers in the later panel show the locations of the sites from which observations are used in the inversion. The analysis increments are roughly only non-zero if an observation site is within about 300-500 km distance. An exception is formed by the emissions of Poland, which are increased by the inversion although not directly observed by an observation site. Higher emissions are in general obtained for the whole of central and eastern Europe. In contrast, emissions are decreased in for example the Netherlands and Belgium, northern Italy, the Paris region, and the central U.K. An isolated increase of emissions is also estimated for the southern North Sea.

It should be kept in mind that the structure of the emission changes is also defined by the assumed spatial length scale in the background covariance; a different length scale might lead to different patterns in regions that are not strongly constrained by the observations.

The largest relative changes are seen for northern Scandinavia, where the already low emissions (mainly from wetlands) are further reduced. For interpretation of the emission changes, a selection of concentration time series at observation sites is shown in Figure 53, where the baselines concentrations obtained from a global inversion were subtracted. The emissions in northern Scandinavia are mainly constrained by site Pallas in Finland, for which the concentration time series is shown as the fifth panel. The lower emissions around Pallas bring simulations closer to the observations, especially in summer.

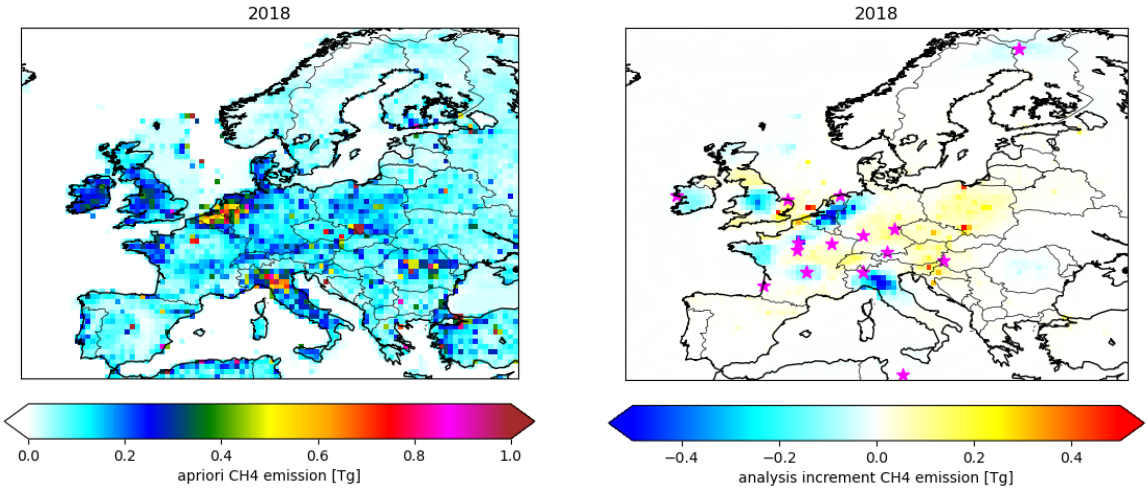


Figure 51: Map of *a priori* CH_4 emissions for 2018 (left) and analysis increment from inversion (right). The magenta markers in the right panel show the locations of the observation sites that were included in the inversion.

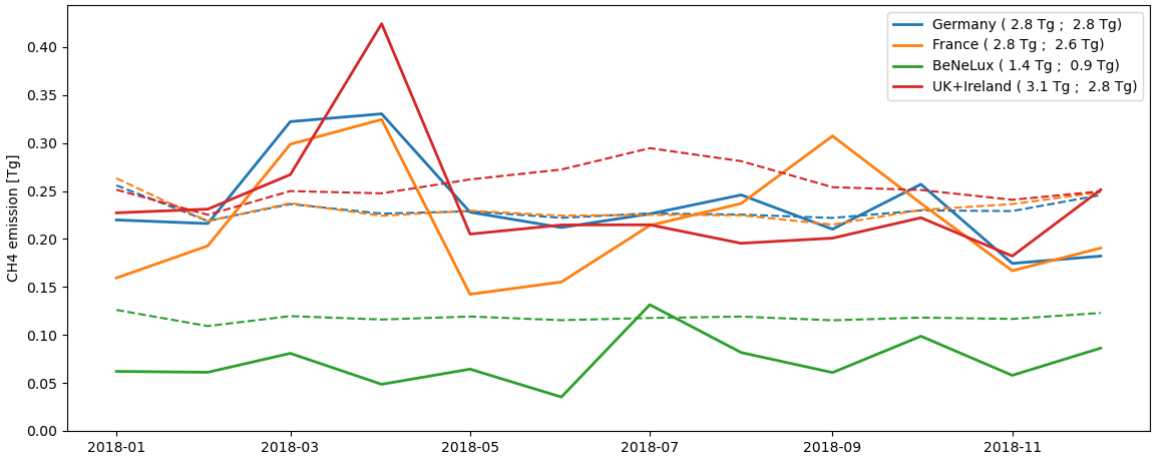


Figure 52: Time series of estimated emissions for selected regions in Europe. Dashed lines show *a priori* estimates, solid lines show posterior estimates. Total emissions over the year are shown in the legend as "prior ; posterior".

Figure 52 shows the time series of emissions aggregated over four selected (groups of) countries. The selection is the same as what was used in (Bergamaschi *et al.*, 2022) that evaluated the results of a high-resolution European CH_4 inversion too. The legend with the region names also shows the *prior* and *posterior* totals over the year.

- For Germany, the estimated total *posterior* CH_4 emissions are 2.8 Tg/year which is equal to the *prior* estimate. However, within the year the changes are substantial, with

increased emissions in spring (March-April) and decreased emissions in early winter (January and November-December). The spatial pattern of the emission increments in the individual months is rather similar to the yearly pattern shown above, with lower emissions in north-west, and higher emissions in the south and east of the country. The higher emissions are seen in the whole of central Europe, and are mainly enforced by the observations in Hegyhátsál (first panel in the figure with concentrations) that require an emission increase for a better simulation, especially in the winter months.

- For France, a small net decrease from 2.8 to 2.6 Tg/year is estimated for the yearly total CH_4 emission. Within the year, similar to Germany, an increase in emissions is estimated for early spring and a decrease for winter but also for early summer. The lower emissions are mainly due to decreased emissions from the Paris region, Bretagne, and central France; higher emissions are estimated for the east of the country. The lower emissions around Paris are mainly imposed to obtain lower concentrations in nearby site Saclay, for which the time series are shown in the second panel of the figure below.
- The inversion estimates substantially lower emissions for the BeNeLux countries. The yearly total emissions have decreased from 1.4 Tg/year to 0.9 Tg/year. The decrease is rather stable over the year, with the exception of July for which a small increase is estimated. The emissions in the region are mainly constrained by site Lutjewad in the north of the Netherlands, for which concentration time series are shown in the third panel of the figure below. The time series show that especially outside the summer the simulations are improved if emissions are decreased in the region.
- Also for the U.K.+Ireland the net emissions are lower, although less strong than for BeNeLux. Especially in Ireland and the central U.K. emissions are decreased, for example to obtain better simulations in Mace Head (fourth panel in the figure below). The decrease of emissions is strongest in summer and autumn, in March a net increase is seen.

The estimated emissions have been compared with the values shown in Figure 2 of Bergamaschi *et al.*, (2022), denoted as (B2022) in the remainder. In that study, the estimated *posterior* emission estimates are more constant and do not show the strong increase in March-April that is seen here. A main difference is also that for the BeNeLux countries a net increase of emissions was estimated instead of the decrease found here. However, in (B2022) also observations from Cabauw (NLD) were taken into account which lead to an overall decrease of emissions around that site, and rather strong changes over small distances. The resolution of the simulation model might have a strong influence therefore. For the other regions the differences between (B2022) and this study are smaller: for Germany and the U.K.+Ireland a similar change in yearly total emissions is estimated; for France, (B2022) estimated a small increase in emissions while this study found a small decrease. The differences between (B2022) and this study shows that differences between inversion systems could be large, and require further experiments to determine the sensitivity to the simulation model, resolutions, and uncertainty specifications.

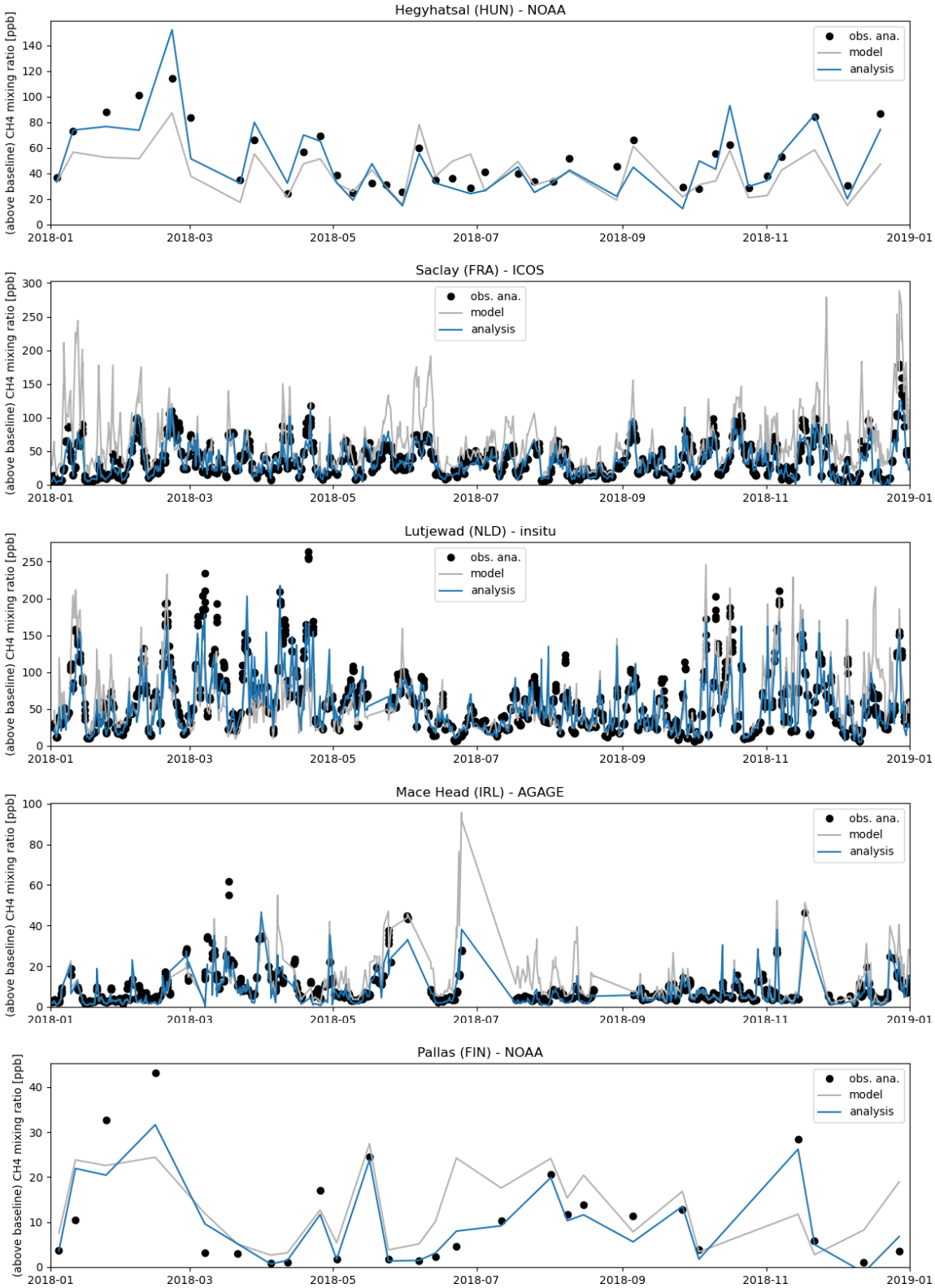


Figure 53: Time series of observed and simulated observations at selected stations, where the baseline concentrations computed by a global inversion have been removed. Only samples that are used in the inversion are included in the time series, thus excluding samples outside [12:00,16:00] local time.

3.9 GEOS-Chem for CO2 inversions from in situ and satellite observations (UEDIN)

3.9.1 Model description

Here, we describe the measurements we use to infer CO_2 fluxes across the UK and mainland Europe; the GEOS-Chem atmospheric chemistry transport model that describes the relationship between *a priori* inventories, atmospheric chemistry and transport, and the observed atmospheric concentrations of CO_2; and the ensemble Kalman Filter that is used to infer CO_2 fluxes from *a priori* knowledge and the measurements.

3.9.1.1 Transport model

The forward model describes the relationship between *a priori* flux estimates of CO_2 and CO and the atmospheric observations. We use the GEOS-Chem atmospheric chemistry transport model to relate surface fluxes of CO_2 and CO to 4-D atmospheric concentrations. We then sample these concentration fields at the time and location of measurements. In the case of satellite observations, we also use the scene-dependent averaging kernel to describe the instrument's vertical sensitivity to changes in CO_2 and CO. Resulting sampled model atmospheric values can then be compared with observations:

y = H . x (Eqation UE1)

where **y** denotes the observation vector, and **x** denotes the state vector that includes our *a priori* CO_2 and CO flux estimates.

We use the GEOS-Chem version 12.5.2 atmospheric chemistry and transport model which we run at 0.25° x 0.3125° resolution for a nested European domain (15°W to 35°E longitude and 34°N to 66°N latitude, Figure 54) with 47 vertical levels. GEOS-Chem is driven by meteorological reanalyses fields from the NASA Global Modelling and Assimilation Office (GMAO) Global Circulation Model.

Our *a priori* flux estimates include all sources contributing to observed atmospheric CO_2 and CO. Equation UE2 shows the sources for CO_2 including combustion emissions ($CO_2^{Combust}$), non-combustion fluxes (CO_2^{Bio}), and background CO_2 that is transported to and from our domain (CO_2^{Trans}). Atmospheric CO sources include combustion emissions ($CO^{Combust}$), transport (CO^{Trans}), and production of CO through oxidation (CO^{Chem}), as shown in equation UE3.

$$CO_2 = CO_2^{Trans} + CO_2^{Combust} + CO_2^{Bio} \quad \text{(Equation UE2)}$$

$$CO = CO^{Trans} + CO^{Combust} + CO^{Chem} \quad \text{(Equation UE3)}$$

For our nested domain, we use boundary conditions for CO_2 (CO_2^{Trans}) from the CAMS inversion-optimized global greenhouse gas analysis with assimilation of *in situ* observations (Chevallier, 2020). Our boundary conditions for CO (CO^{Trans}) are from the CAMS global reanalysis (Inness et al., 2019). We use the CAMS fields at their provided temporal resolution (3-hourly) and re-grid to the GEOS-Chem horizontal spatial resolution of 2° x 2.5°. Because the vertical resolution of GEOS-Chem does not align with CAMS, we translate the CAMS native vertical resolution to our 47 model layers using linear interpolation of logarithmic pressure values. We fill in the species concentrations at the lowest or highest pressure level in CAMS for the top or surface of the atmosphere, respectively, when the GEOS-Chem pressure levels go beyond the bounds of CAMS.

We treat the relationship between surface fluxes and concentrations (equation UE1) as linear (e.g., a doubling of emissions leads to a doubling of the atmospheric signal). To linearize the CO simulation, we use offline chemistry terms to represent the chemical production of CO (CO^{Chem}). CO is primarily produced by oxidation of methane and non-methane volatile organic compounds by the hydroxyl radical (OH), so we generate the production terms using offline 3-D loss fields of OH generated from a previous GEOS-Chem full-chemistry simulation (Fisher et al., 2017).

3.9.1.2 Prior fluxes and assimilated observations

Prior fluxes

For our 2018-2021 a priori fluxes, we use a combination of regional and global inventories (Figure 54). Combustion emissions for both species ($CO_2^{Combust}$ and $CO^{Combust}$) are from the TNO GHGco v3.0 emission inventory at 0.1° x 0.1° resolution (Super et al., 2020; Kuenen et al., 2022) with national totals based on emissions reported in national inventories and extrapolated from 2019 to more recent years. We apply scaling factors provided by TNO to reflect monthly, hourly, and daily patterns in emissions by sector. Our combustion source also includes biomass burning emissions from the GFAS v1.2 inventory (Kaiser et al., 2021). Non-combustion fluxes (CO_2^{Bio}) include ocean fluxes from the NEMO-PISCES model (Lefèvre et al., 2020), lateral carbon fluxes related to crop removal (Deng et al., 2022), and hourly terrestrial biosphere fluxes at 1/120° x 1/60° resolution produced by the VPRM model following methods described by Gerbig (2021) driven by ERA5 meteorology. We include non-combustion anthropogenic emissions from the TNO inventory in our non-combustion fluxes.

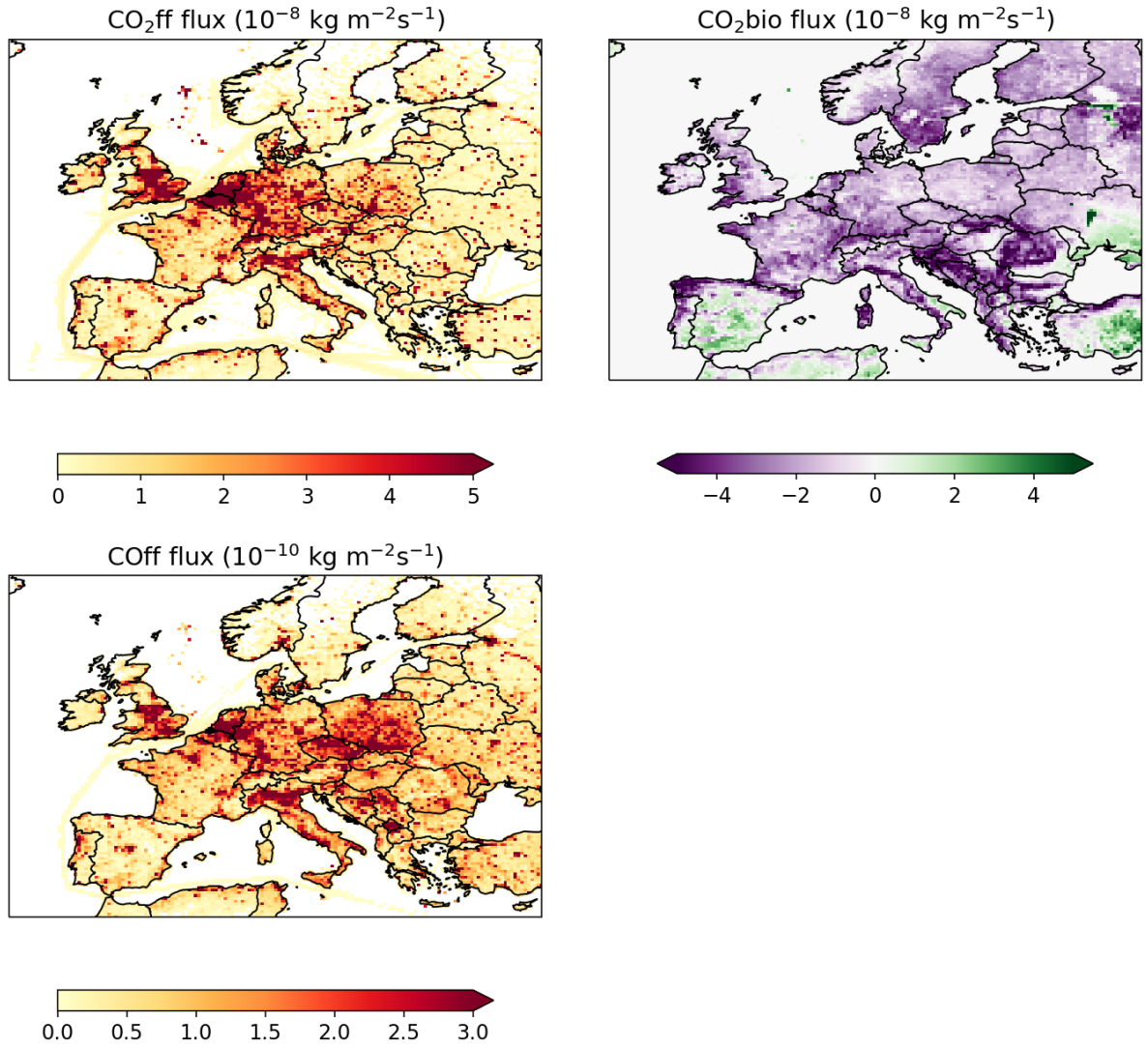


Figure 54: Annual mean emissions for 2018-2021 in the a priori inventories. Combustion emissions (CO_2ff, COff) are from the TNO inventory while biogenic fluxes (CO_2bio) are from the VPRM model.

Satellite and In Situ Observations

For CO_2, we use observations of the atmospheric CO_2 column-averaged dry-air mole fraction (XCO_2) from the OCO-2 satellite, launched in 2014 (Crisp et al., 2017; Eldering et al., 2017). We use OCO-2 ACOS v10r data for 2018-2021 (OCO-2 Science Team et al., 2020; Taylor et al., 2023). For CO, we use XCO observations from TROPOMI, July 2018 - December 2021, aboard the Sentinel-5p satellite, launched in 2017 (Veefkind et al., 2012; for CO retrieval: Vidot et al., 2012; Landgraf et al., 2016). For both satellite products, we filter observations as recommended in the Product User Guide, including a strict quality assurance flag of >0.75 for TROPOMI XCO. We remove glint observations and those over the oceans and re-grid satellite columns and averaging kernels to a 0.25° x 0.3125° spatial grid to match model output (Figure 55). To compare our model output to the satellite observations, we first sampled the model at the time and location of each instrument. We then interpolate our model pressure levels to the satellite pressure levels and apply the scene-dependent retrieval averaging kernel to our 3-D model concentration fields.

We use in situ observations for 2018-2021 (Figure 55). We use the DECC surface measurement network in the UK (Stanley et al., 2018) and the ICOS measurement network

for Europe (ICOS R., 2022), including drought adjusted observations for 2018 (Ramonet et al., 2020). We retain in situ observations collected between 0900 and 1800 local time – to avoid instances when the tall tower inlets sit above a shallow boundary layer – and then time-average to 3-hourly intervals to match our GEOS model meteorology. All in situ sites have CO_2 observations, but some sites are missing CO observations. We additionally remove observations when the atmosphere is not well-mixed, diagnosed where the standard deviation of CO_2 concentrations across the lowest five vertical model levels is smaller than 0.3 ppm.

Figure 55 also shows European sites from the Total Carbon Column Observing Network (TCCON). Five sites are within our domain, including Bremen (Germany), Karlsruhe (Germany), Nicosia (Cyprus), Orléans (France), and Paris (France). We use the TCCON observations as an independent comparison for our inversion results.

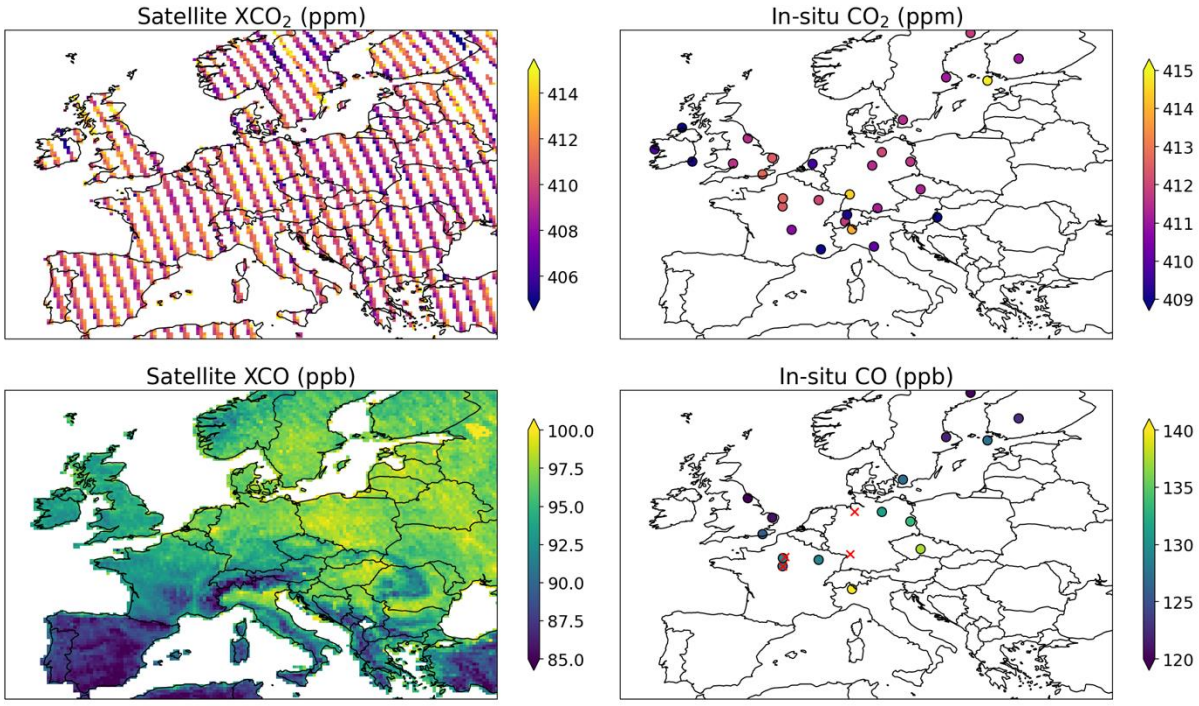


Figure 55: Annual mean CO_2 and CO observed by satellite and in situ networks across Europe for 2018-2021. Satellite observations of XCO_2 and XCO are from OCO-2 and TROPOMI, respectively, and in situ observations are from the DECC and ICOS networks. The red X points in the *in situ* CO plot show the locations of the five TCCON sites we used to evaluate our inversions. The observations are filtered as stated in the text and satellite observations are shown at 0.25° x 0.3125° resolution. TROPOMI observations only include observations after July 2018.

3.9.1.3 State vector

For our inversion, we use the Ensemble Kalman Filter (EnKF) approach as discussed in detail by others (e.g., Peters et al., 2005; Hunt et al., 2007; Feng et al., 2009; Liu et al., 2016). We specifically follow the methods derived by Hunt et al. (2007) and summarized by Liu et al. (2016) for the Local Ensemble Transform Kalman Filter (LETKF).

We solve the inversion in ensemble space rather than for the state vector elements. For each state vector element, we have an ensemble of potential scale factors that follow our prescribed error statistics. For each assimilation time-period (over which we ingest observations), we solve for the mean *a posteriori* state vector that represents the mean of our N ensemble members (where we use N = 100):

For CO_2, we use an *a priori* model error of 1.5 ppm for the satellite inversion (Feng et al., 2017) and 3 ppm for the *in situ* inversion (within the range of Monteil et al., 2020 and White et al., 2019). For CO, we use an *a priori* model error of 15 and 20 ppb for the satellite and *in situ* inversions, respectively (Northern Hemisphere CO column and surface mole fraction model-observation differences from Bukosa et al., 2023). We use the observation errors as provided for the satellite or *in situ* network, averaged to the model resolution. We generate the off-diagonal error covariances based on the spatial and temporal proximity of observations following an exponential decay with spatial and temporal length scales of 100 km and 4 hours, respectively.

We use an assimilation window of two weeks and a lag window of one month, accounting for the impact of historical emissions on our assimilation period. This means that the state vector for each time-step includes scale factors for the assimilation window and lag window. We perform our inversion sequentially, using the *a posteriori* scale factors for a given assimilation window to update the *a priori* scale factors for the next lag window over the same date range. To avoid unrealistically small prior uncertainties, we apply a 10% error inflation when we update the *a priori* state vector.

The benefit of the LETKF is that we can localize the inversion so that each state vector element is only influenced by a subset of observations. For our inversions using *in situ* observations, we localize by distance so that each state vector element that represents a grid-scale variable is only influenced by observations within a 1000 km range.

We test different approaches to investigate the usefulness of satellite observations for evaluating CO_2 combustion emissions. The approaches vary in the observations that are used and the representation of error covariances for our *a priori* estimates. For each type of inversion, we compare our satellite inversion results to comparable inversions using *in situ* observations.

In the inversions, instead of solving for CO_2 or CO fluxes, we solve for scale factors that scale up or scale down the source terms from equations UE2 and UE3. We first assume that our *a priori* scale factors are all equal to one. We solve for *a posteriori* scale factors that, when applied to our source terms, will result in modelled atmospheric CO_2 or CO concentrations in better agreement with observations.

For our first approach, we perform a CO_2-only inversion that assimilates CO_2 observations. Our state vector includes scale factors for the sources of equation UE2. Each scaling factor applies to a non-combustion or combustion grid cell at 0.5° x 0.625° resolution. For the transport scale factors, each element applies to CO_2 transported from the North, South, East, or West boundary.

In our second approach, we perform a joint CO_2:CO inversion that assimilates both CO_2 and CO observations. For the joint inversion, we assume there is 100% correlation for the CO_2 and CO combustion emission errors. This means any adjustment made by our inversion to the CO_2 combustion scale factors will also apply to the CO scale factors and vice versa. We can then use a common combustion scaling term for both species in our state vector. Our state vector also includes scale factors for transport of each species, and for CO we include a vector with two scale factors for the chemistry terms.

For our first two approaches, we assume a prior uncertainty of 20% (relative standard deviation) for the combustion scale factors. We use an *a priori* uncertainty of 50% for the non-combustion scale factors, and 5% for the atmospheric transport and chemistry scale factors. For our non-combustion and combustion scale factors, we generate error covariances for nearby grid cells that exponentially decays with increasing distance.

We acknowledge that the assumption of 100% error correlation for CO_2 and CO combustion emissions is likely to be a gross overestimate. For example, we may underestimate CO emissions due to an underestimate of incomplete combustion activities, and this will not

translate to the same underestimate in CO_2. So, for our third approach, we test this assumption by solving for the CO_2 and CO combustion scaling terms separately.

We call this our TNO approach because we use estimates of the uncertainties in the TNO emission inventory to create our error covariance matrix (Super et al., 2023). We increase the provided uncertainties by a factor of 3 to make them more comparable with our other simulations. This results in a mean grid-scale CO_2 combustion uncertainty of 18%, though there is greater variability in grid cell uncertainties compared to our other approaches. We expect higher correlation between CO_2 and CO gridded emissions in regions where the same spatial product is used to distribute emissions for both species (e.g., road network maps) and that spatial product has high uncertainties.

3.9.1.4 Period of study

We report fluxes from 2018 to 2021, inclusively.

3.9.2 Results

First, we describe the comparison between our *a priori* and *a posteriori* model simulations against observations. We then report our *a posteriori* CO_2 fluxes for Europe and its constituent countries and the UK.

3.9.2.1 Inversion performance

Our *a priori* CO_2 emissions are already consistent with data from the five relevant TCCON sites (locations shown in Figure 55; R= 0.87), and *in situ* (R=0.76) and satellite (R=0.84) observations. The model has small, positive relative mean bias compared to TCCON (0.7%) and a very small bias compared to *in situ* and satellite observations (0.2%). The satellite inversions show improvement for the model-satellite fit (R=0.92-0.95), as expected, and the model-*in situ* fit (R=0.80-0.82). Similarly, the *in situ* inversions improve model-*in situ* fit (R=0.83-0.84) and to a lesser extent the model-satellite fit (R=0.85-0.87).

In general, including CO and TNO uncertainty estimates improves the model-observation fit and reduces the mean bias. For example, the satellite joint CO_2:CO (R=0.93) and TNO (R=0.92) inversions show the greatest improvement in fit with TCCON. The one exception is that the mean bias compared to TCCON is slightly larger with CO (0.3-0.5%) compared to CO_2-only (0.2-0.4%). The TCCON bias is seasonal with the *a priori* model showing no bias in July-August and a positive bias of 1-4 ppm for the rest of the year. The *in situ* inversions reduce the mean bias for March-June by 1 ppm, and this improvement lines up with a reduction in the biosphere sink for these inversions (discussed later).

We also assess inversion performance by the degree of uncertainty reduction for the *a posteriori* CO_2 combustion emissions. Table 14 shows *a posteriori* uncertainties for our domain-scale CO_2 combustion emissions. The reductions in relative uncertainty achieved at the domain scale for all inversions are small (6-12%) with the CO_2-only and TNO satellite inversions showing no reduction. The TNO inversions show smaller reductions in uncertainty (0-6%) compared to the joint inversions (8-12%), but they also start with a lower *a priori* uncertainty at 1.6% (relative standard deviation; RSD) compared to 2.4% for non-TNO *a priori* uncertainties.

Table 14: Average domain CO_2 combustion emissions for 2018-2021

	Mean (Gt a^{-1})		RSD [a] (%)
	Emission	Change [a]	
A priori[b]	4.9		2.4
TNO *a priori*[b]	4.9		1.6
Satellite			
CO_2-only	4.9	-	2.4
Joint CO_2:CO[c]	4.6	▼	2.1
TNO CO_2:CO[c]	4.8	-	1.6
In-situ			
CO_2-only	4.8	▼	2.2
Joint CO_2:CO	5.0	▲	2.2
TNO CO_2:CO	4.9	-	1.5

[a]The arrows indicate the change of the mean from the *a priori*. Blue-downward pointing arrows show a decrease, red-upward show an increase, and grey dashes show no change. RSD stands for relative standard deviation.

[b]The *a priori* uncertainty labelled as '*A priori*' is for the CO_2-only and joint inversions, so we also include the *a priori* uncertainty for the TNO inversion.

[c]The Joint and TNO satellite inversions only include July 2018 - December 2021. The *a priori* combustion emission for this period is 4.8 Gt a^{-1} so we show no change for the TNO *a posteriori* emissions.

At the national-scale, we see the greatest uncertainty reduction in CO_2 combustion emissions for the top 10 emitting countries when satellite CO observations or *in situ* CO_2 measurements are included and the non-TNO uncertainties are used (not shown). The average uncertainty reductions for the joint satellite and CO_2-only *in situ* inversions are 11% and 9%, respectively. This is not surprising given the greater number of observations provided by these two platforms and increased sensitivity to surface fluxes compared to OCO-2. Including *in situ* CO observations in the inversion does not improve the national-scale uncertainty reduction. Because we use lower *a priori* uncertainties in the TNO inversion (national-scale 2-10% RSD) compared to the other inversions (6-14% RSD), fewer countries have reduced uncertainties

for the TNO inversion, though *a posteriori* uncertainties are reduced in the Netherlands (2%) for both in-situ and satellite compared to *a priori* uncertainties (3%).

3.9.2.2 European-scale Emission Estimates

Table 14 shows our mean European combustion emissions for 2018-2021. The inversions show a small decrease or no change from the *a priori* emissions (4.9 Gt a^{-1}), except for the joint satellite and *in situ* inversions that show a larger decrease (4.6 Gt a^{-1}) and an increase (5.0 Gt a^{-1}) from the *a priori*, respectively. Figure 56 shows that the joint satellite inversion decreases combustion emissions year-round for all years with the greatest decreases in winter. The TNO satellite/*in situ* and CO$_2$-only *in situ* inversions also show decreases in the winter and early spring (Figure 56 and Figure 57), providing more confidence in this scaling down of emissions.

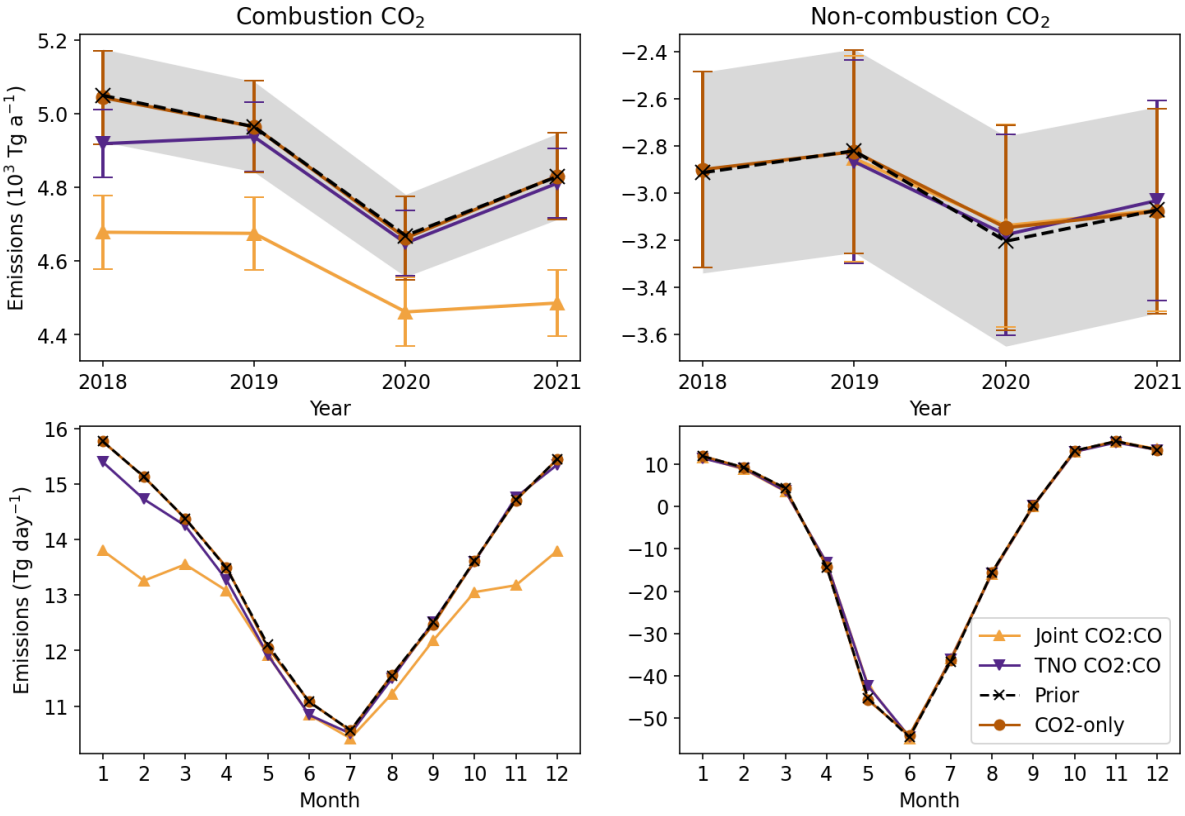


Figure 56: Annual and monthly mean European CO$_2$ combustion and non-combustion emissions inferred from satellite inversions for 2018-2021. The top row shows annual mean CO$_2$ flux estimates by inversion type, with errors bars showing the 1-σ errors except for the a priori errors which are shown as a shaded region. The bottom row shows monthly mean fluxes for 2018-2021. The TNO and joint inversions only include July 2018-December 2021 for combustion and 2019-2021 for non-combustion.

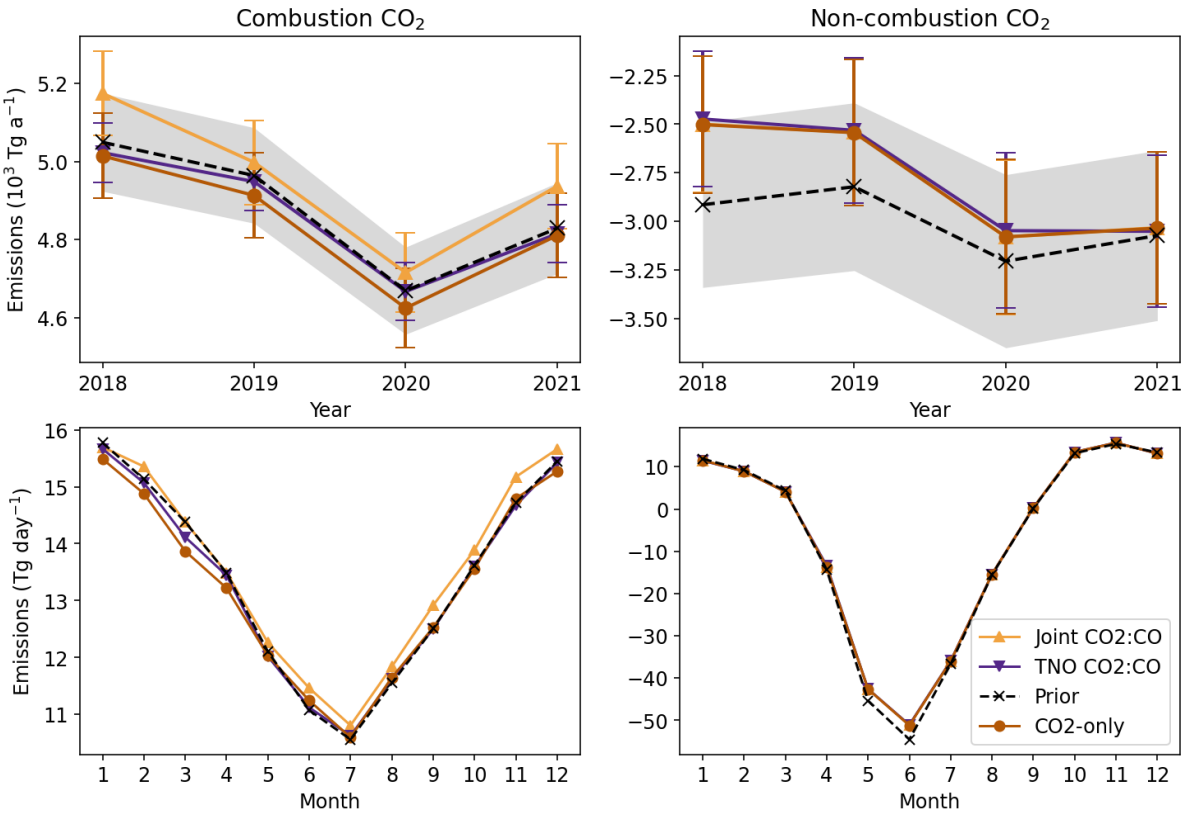


Figure 57: Same as Figure 56 for in situ inversions.

In contrast, the joint *in situ* inversion shows an increase for all months and all years (Figure 57). This pattern is not reflected in our other inversion approaches and is likely, in part, due to an underestimation of the fine-scale variability in CO measured at some *in situ* sites, leading to an over-correction upward of combustion emissions. For example, we find that removing a single site close to an urban region in northern Italy (Ispra ICOS site) reverses the sign of scaling in the region from an increase to a decrease. The disagreement between satellite and *in situ* CO_2:CO inversions is less pronounced for the TNO inversions because the separation of CO_2 and CO in our state vector prevents the CO underestimates from heavily influencing the CO_2 combustion emissions.

Figure 56 shows a slight (1%) decrease in mean *a priori* combustion emissions from 2018 to 2021, and all satellite and *in situ* inversion results show a similar trend (Figure 56 and Figure 57). The mean *a priori* non-combustion CO_2 sink shows a slight increase (1%) for 2018-2021, and the inversion results show a similar (satellite; Figure 56) or greater increase (*in situ*; Figure 57) in the CO_2 sink. Figure 57 shows the monthly mean non-combustion CO_2 sink is weakened for the *in situ* inversions, mostly in summer, whereas Figure 56 shows almost no change in the sink for the satellite inversions, indicating that the CO_2 *in situ* observations are needed for constraining biogenic flux estimates.

The differences between *a posteriori* and *a priori* annual emissions for all inversions except the Joint satellite inversion are not statistically significant and remain within the 1-σ uncertainties of the *a priori* estimate. The inter-annual trends are also smaller in magnitude than the *a posteriori* uncertainties, making it difficult to assess if CO_2 combustion emissions in Europe have decreased from 2018 to 2021.

3.9.2.3 National-scale Emission Estimates

Figure 58 shows national CO_2 combustion emissions for the top 10 emitting countries in our European domain. Germany is the highest emitter with an *a priori* emission of 821 Tg a^{-1}. Most inversions show a decrease in Germany's emissions (717-806 Tg a^{-1}) except for the *in situ* joint inversion which shows an increase (830 Tg a^{-1}) and the CO_2-only inversion which shows little change from the *a priori* estimate (819 Tg a^{-1}). The other top emitting countries, including Poland, the UK, France, Italy, Spain, Belgium, the Czech Republic, the Netherlands, and Romania, show emissions decreases for the satellite joint (3-17%) and TNO (0-4%) inversions. The *in situ* CO_2-only and TNO inversions generally show only small changes (<1%) in national emissions except for a 4% national emission decrease in the Netherlands and Belgium for the CO_2-only inversion.

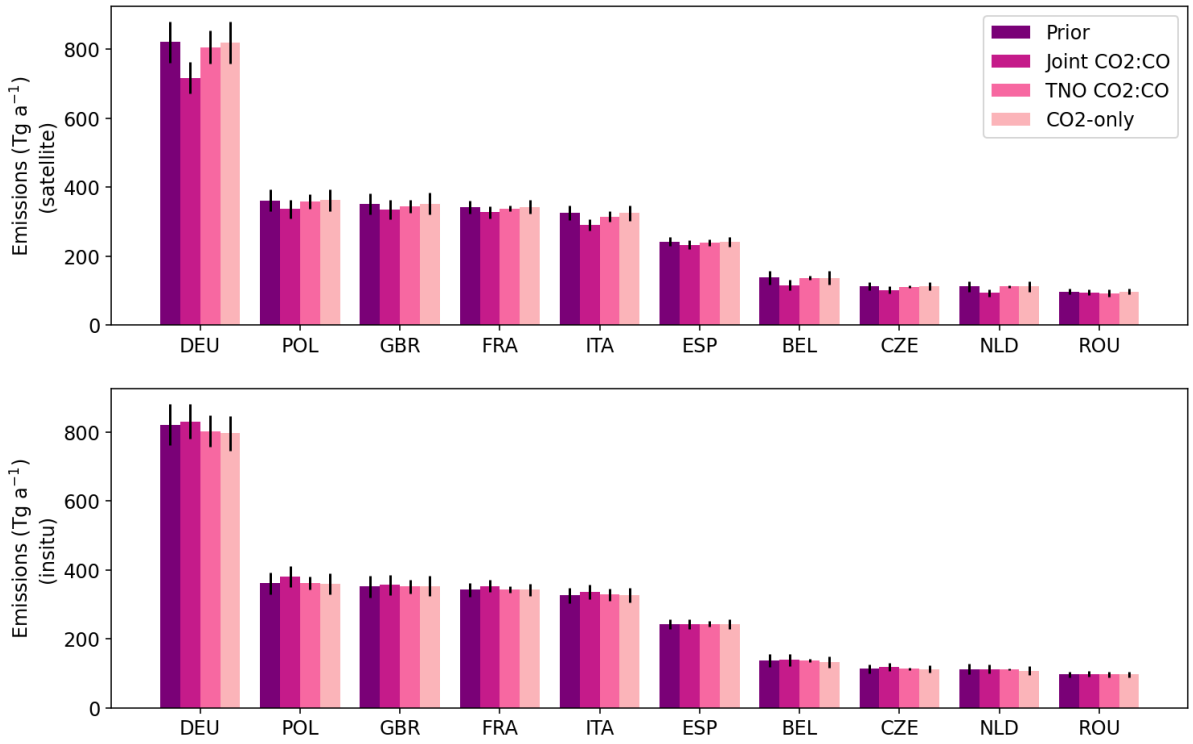


Figure 58: Annual mean *a priori* and *a posteriori* CO_2 combustion emissions by country for satellite (top) and in situ (bottom) inversions. We show the top 10 emitting countries in our European domain with emissions averaged over 2018-2021. The TNO and joint satellite inversion averages do not include dates prior to July 2018.

The joint inversions show the largest changes in national emissions but in opposite directions. In contrast, the TNO inversions show smaller changes from the *a priori* (in part, due to the lower *a priori* uncertainties) and better agreement, including agreement in Germany where there is greater divergence from the *a priori* estimate (2% decrease for both TNO inversions).

Despite the national-scale disagreements for some inversions, we find regional corrections to combustion emissions are consistent for all inversions. Figure 59 shows that the densely-populated North Rhine-Westphalia region in western Germany shows a decrease in CO_2 combustion emissions for all inversions. The TNO and CO_2-only inversions show mixed corrections in Poland with TNO inversions showing the best agreement. Most inversions, including both TNO inversions, show an increase in emissions near Milan and Vienna, but

over other major cities like Paris, Madrid, and London there is less agreement in the sign and magnitude of the emissions changes.

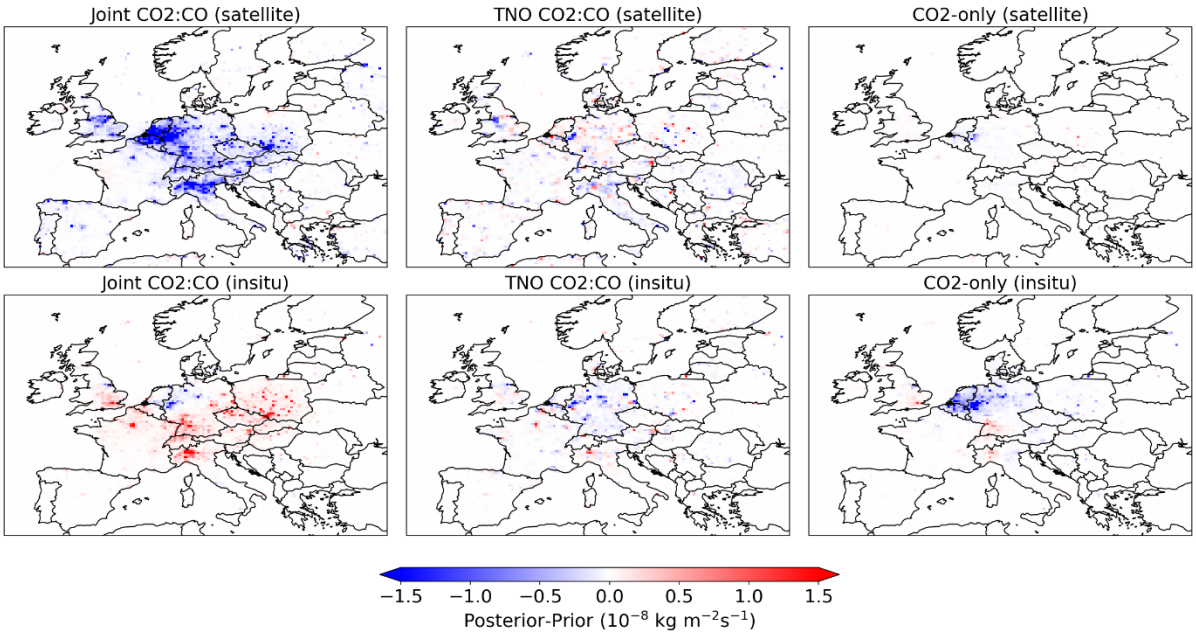


Figure 59: Annual mean CO_2 combustion emissions difference (*a posteriori* minus *a priori*) for satellite (top row) and in situ (bottom row) inversions, 2018-2021, shown at the native model resolution of 0.25° x 0.3125°. The TNO and joint satellite inversion averages do not include dates prior to July 2018.

The differences in the joint inversions are due to contrasting corrections to CO emissions that carry over into the CO_2 emissions. We find that the *in situ* joint inversion shows decreases for high-emitting regions in Europe for winter and spring, but this is mostly offset by large emission increases in summer and fall (not shown). In contrast, the satellite joint inversion shows decreases for all seasons. For the TNO inversion, there is less disagreement between the seasonal emissions corrections for CO_2, but there are disagreements in CO corrections. We also find that CO corrections for the TNO inversion generally occur at the national-scale (not shown) and we know there is low error correlation between the two species at the national-scale (Super et al., 2023), so it is not surprising that these corrections do not carry over to CO_2.

Figure 60 shows national non-combustion emissions for the countries in Figure 58. All countries show a net sink with France having the largest net sink. The *in situ* inversions tend to decrease (lessen) the CO_2 sink for all countries and reduce uncertainties. Figure 61 shows the spatial pattern in the flux changes is consistent for all *in situ* inversions. In contrast, the national CO_2 non-combustion fluxes show little change from the *a priori* for the satellite inversions, highlighting the importance of *in situ* CO_2 observations for constraining biogenic flux estimates. For all inversions, the CO_2 sink in northern Germany is strengthened (more negative fluxes) and weakened in southern Germany and Switzerland, though there are conflicting corrections in surrounding regions such as France and northern Italy. These disagreements may be due to the differing observing capacities with satellites having seasonal limitations due to snow and clouds. We find low *a posteriori* error correlations between national-scale combustion and non-combustion fluxes (mostly $R < 0.1$, except for Germany $R = 0.2$ negatively correlated), indicating that the disagreement in *in situ* and satellite *a posteriori* non-combustion fluxes will not carry over into combustion emission estimates.

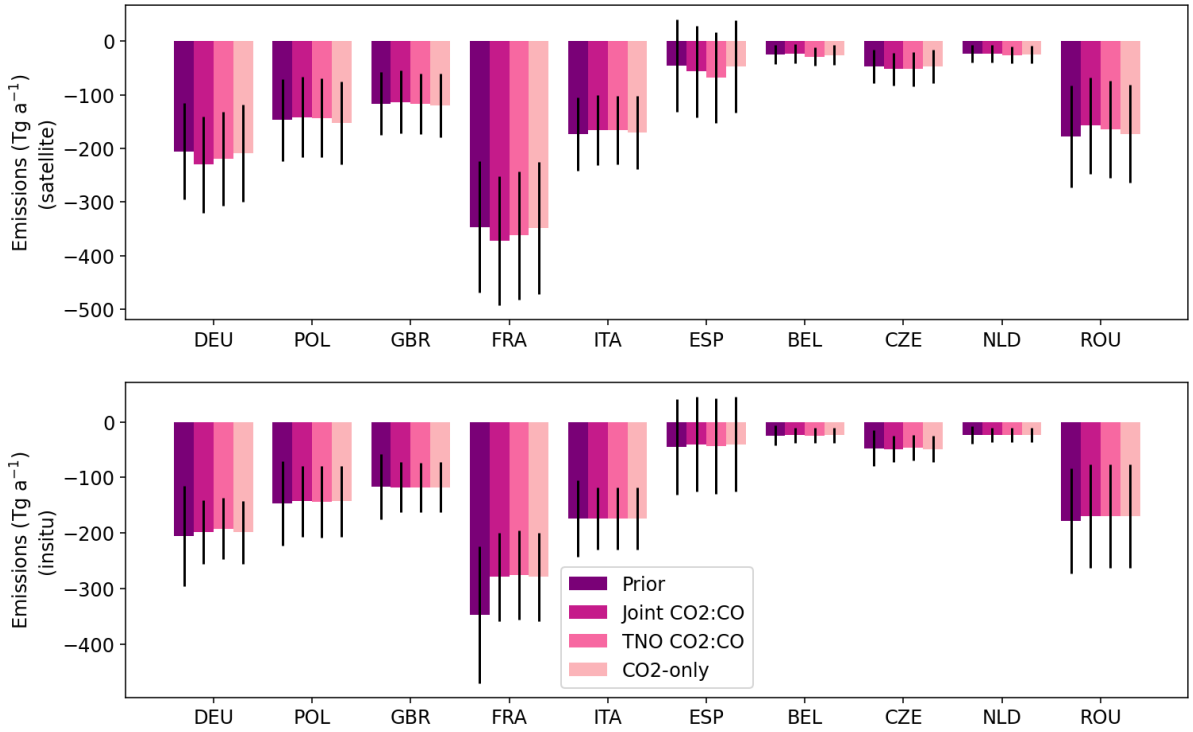


Figure 60: As Figure 58 but for non-combustion CO_2 fluxes estimates. The TNO and joint satellite inversion averages do not include 2018.

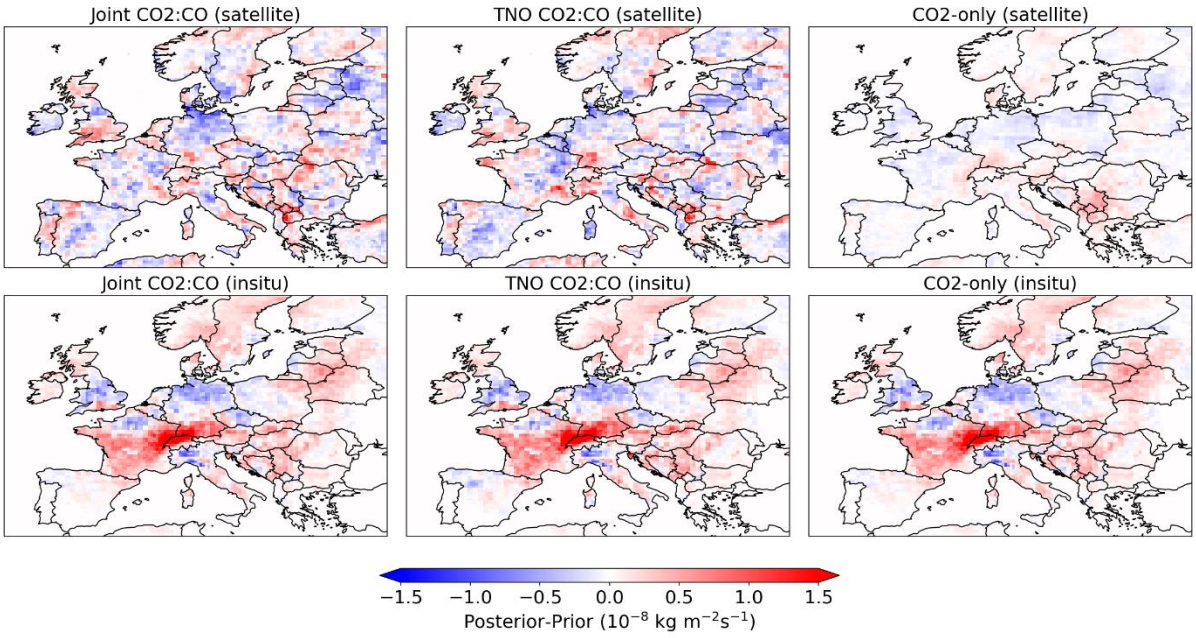


Figure 61: As Figure 59 but for non-combustion CO_2 flux estimates. The TNO and joint satellite inversion averages do not include 2018.

3.9.3 Concluding remarks

We find that using CO_2 satellite observations from OCO-2 alone cannot reproduce *a posteriori* European CO_2 fluxes inferred from the European *in situ* CO_2 measurement network. The

satellite observations (CO_2-only) do not show significant combustion emissions changes from our *a priori* estimates, whereas when we use *in situ* CO_2 or CO_2 and CO satellite observations we see greater divergence from the *a priori* emissions. All our inversions indicate that CO_2 combustion emissions for regions of Germany are overestimated in winter, and most inversions show this overestimate extends to other countries in Europe. We also find that the *in situ* inversions show a smaller summertime European CO_2 sink which is not shown for the satellite inversions.

We find that using CO observations and TNO error estimates leads to better agreement between satellite and *in situ* inversions and the best model-observation fit. The one exception we find is that the model bias compared to TCCON does not improve when we include CO. This is likely because the bias is in part due to biogenic fluxes, and only including *in situ* CO_2 observations provides a useful constraint for biospheric flux estimates. In general, these improvements in model-observation fit are small and we do not see significant reduction in uncertainties compared to our *a priori* estimate.

The use of CO as a tracer for combustion in a CO_2 inversion system is promising for improving our ability to track CO_2 combustion emissions, though we find that our interpretation of results is highly dependent on the assumptions of *a priori* error correlation between CO and CO_2. The *in situ* network is still essential for constraining non-combustion fluxes, but we find low correlation between combustion and non-combustion fluxes so the differences in biogenic flux estimates does not prevent the estimation of combustion emissions at the national scale using satellite observations.

We find that the existing observational networks are not able to significantly reduce the errors for our European or national emission estimates to the extent necessary for distinguishing inter-annual emission trends that represent only a few percent of total emissions. Our results also show that including co-emitted species in the inversion improves our ability to estimate combustion emissions of CO_2, suggesting that the increase in observational capacity for CO_2 and co-emitted trace gases promised by the Copernicus CO_2 Monitoring (CO2M) satellite mission may improve our ability to constrain regional and national emission estimates.

Despite the sensitivity of our *a posteriori* emission estimates to the choice of *a priori* CO_2 and CO uncertainties, we find that the existing observing system is not able to clearly distinguish between inversion results, with the joint and TNO satellite inversions performing similarly. This highlights the need for not only further satellite observing capacity but also improved ground-based networks for evaluating satellites and the usefulness of including co-emitted species observations.

3.10 CIF-FLEXPART for CH_4 inversions from in situ observations (FMI)

3.10.1 Model description

We utilize the Community Inversion Framework's (CIF) 4-dimensional variational optimization approach (4DVAR) as the backbone of our inversion system (Berchet et al. 2021). To evaluate source receptor sensitivity and estimate footprints, we use the Lagrangian Particle Dispersion Model (LPDM), FLEXible PARTicle dispersion (FLEXPART) model, driven by ECMWF's ERA5 reanalysis data.

3.10.1.1 Transport model

Our inversion system uses FLEXPART v10.4 as its atmospheric transport model, which is a widely-used open-source LPDM. FLEXPART is employed to replicate the dispersion and turbulent mixing of gases and aerosols in the atmosphere (Stohl et al. in 1995; Pisso et al. in 2019). This model relies on meteorological data, either reanalyses or forecasts, for its operation. In this study, we employ meteorological input data from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hittmeir et al. 2018). The ECMWF data are

available at hourly intervals, span 137 vertical layers, and features a horizontal resolution of ~0.25° x 0.25°, though here a resolution of 1° × 1° is used.

To prepare the ECMWF ERA5 reanalysis data for use, we utilize the Flex-Extract toolbox (Tipka et al. 2020). We then run the FLEXPART model to generate source-receptor sensitivity matrices. For each observation event, we conduct a 10-day backward transport simulation using the FLEXPART model to produce surface flux footprints. These footprints are saved with a resolution of 0.2° × 0.2° for the nested region and 1° × 1° for the global domain, all of which are stored at hourly intervals. For atmospheric greenhouse gases like CH_4 which have longer atmospheric lifetime, the footprints need to be coupled with background concentration. Thus, we use concentration from CAMS global inversion-optimised greenhouse gas (https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-greenhouse-gas-inversion?tab=overview).

We have defined the region boundary of the nested domain to align with JSBACH, which is responsible for calculating the CH_4 biogeochemical flux. The region is specified as: longitude ranging from 12° W to 37° E and latitude ranging from 35° N to 73° N. This precise geographic area has been chosen to ensure that the data and calculations conducted by JSBACH accurately represent the CH_4 biogeochemical processes within this specific geographical scope.

3.10.1.2 Prior fluxes and assimilated observations

To estimate methane emissions from various sources and sinks, we incorporate monthly data from major categories: anthropogenic, fire, biospheric, as well as climatological estimates for ocean, geological, and termite emissions. These datasets are aggregated to provide comprehensive methane emission estimates. Our monthly prior flux data for anthropogenic emissions are sourced from the EDGAR v6.0 GHG inventory emissions by excluding agricultural waste burning, (Crippa et al., 2021), accessed on October 15, 2023, at https://edgar.jrc.ec.europa.eu/index.php/dataset_ghg60. Monthly prior emission estimates for fire emissions are obtained from GFEDv4.1, which encompasses biomass burning, including agricultural waste burning. Global biospheric emission estimates are derived from the JSBACH ecosystem model products, which replaced LPX-Bern DYPTOP v1.4 for the Europe domain. These biospheric emissions result from aggregating emissions from peat, inundated soil, and net mineral soil emissions, as calculated by JSBACH. For the ocean prior flux, we utilize climatological estimates of ocean emissions from the work of T. Weber (2019). Geological prior emissions are derived from the research of Etiope et al. (2019). Lastly, termite emissions, as a prior flux, are based on the findings of Ito and Inatomi (2012).

We incorporated observational constraints by leveraging data from 43 in-situ measurements sourced from Obspack Europe CH_4 time series, specifically from obspack_ch4_466_GLOBALVIEWplus_v8.0_2023-04-26 downloaded from (https://meta.icos-cp.eu/objects/wIrU4_bb2C74Al01I3d9WyzB, last accessed on Oct 17, 2023).

In situations where multiple inlet heights were available, as exemplified by the Cabauw station with inlets at 27, 67, 127, and 207 meters above ground level, we opted to assimilate data solely from the highest inlet. This approach was taken to ensure that the assimilated data was representative of the uppermost atmospheric layer, where it can have a more significant impact on our understanding of methane concentrations and fluxes.

3.10.1.3 State vector

We use 4DVAR data-assimilation method to minimize the cost function that represents the mismatch between model predictions and observed data:

$$J(x) = 1/2 * (x - x_b)^T * B^{-1} * (x - x_b) + 1/2* (y - H(x))^T * R^{-1} * (y - H(x))$$

where J(x) is the cost function to be minimized, x is the state vector representing model variables we seek to optimize, x_b is the prior state or initial guess for the state vector x, B is the background error covariance matrix, representing the uncertainty in the background state. y is the vector of observed data. H(x) is the observation operator, which maps the model state x to the observation space. R is the observation error covariance matrix, representing the uncertainty in the observed data. The 4DVAR technique seeks to identify the state vector x that minimizes the cost function J(x). By minimizing J(x), the data assimilation process aligns the model state with both the underlying physics of the model and the observed data. To initiate this process, when the state vector x equals x_b initially, the local cost function gradient J(x) is computed through the adjoint method [Errico, 1997]. Subsequently, this gradient is employed in the determination of an updated state vector that leads to a reduced value of J(x). The conjugate gradient algorithm [Lanczos, 1950] is utilized for this purpose. This iterative process continues until the gradient norm falls below a predefined convergence threshold, typically set at 0.1%. We represent emission error correlations, specifically the off-diagonal terms of B, by employing a Gaussian function that considers the spatial and temporal separation between grid cells. The spatial correlation length is set at 200 km, while the temporal correlation length is set at one month.

3.10.1.4 Period of study

3.10.2 Results

Figure 62 displays sample time series of CH_4 measurements obtained from the Cabauw site, accompanied by both the prior and posterior estimates generated through the CIF-FLEXPART inversion. The results clearly illustrate that the posterior estimates exhibit a significantly improved alignment with the observed time series, emphasizing the effectiveness of the data assimilation process.

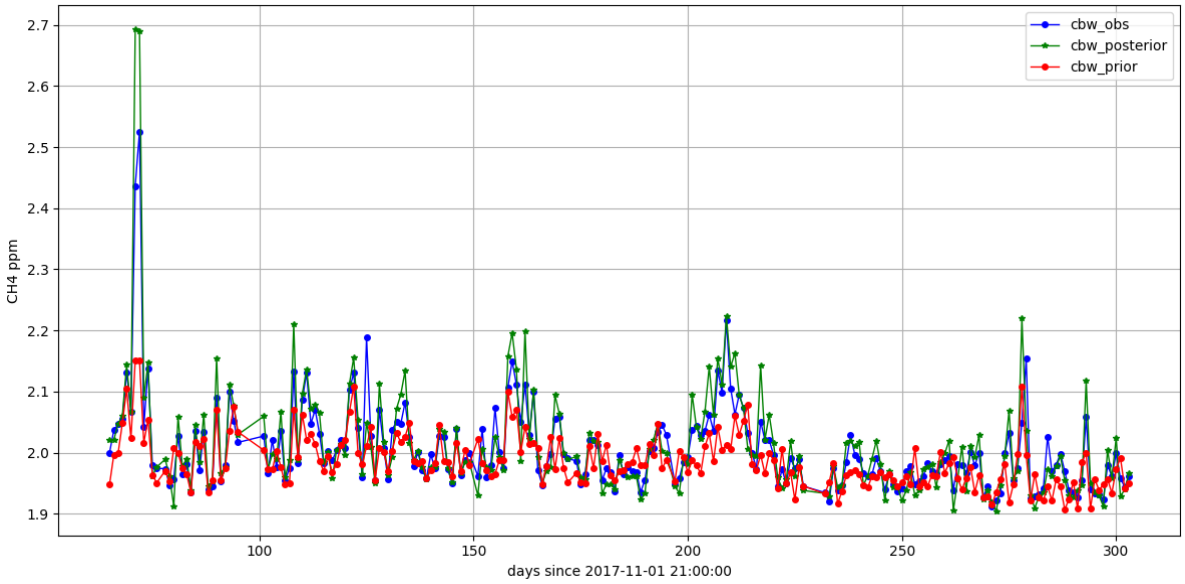


Figure 62: Hourly time series of CH_4 concentration sampled at station Cabauw: assimilated Obspack Measurement (blue), CIF-FLEXPART Inversion results for posterior concentration (green), and prior concentration (red)

In Figure 63, we provide a comprehensive analysis of the root mean square error (RMSE), bias, and correlation metrics for both the prior and posterior estimates associated with each observation integrated into our inversion system. The findings presented in this figure provide compelling evidence of the superior performance of the posterior estimates. They exhibit lower bias and RMSE values while demonstrating a notably higher level of correlation with the observed data. This underscores the substantial enhancements achieved through our data assimilation methodology, resulting in more accurate and reliable estimates of CH_4 emissions.

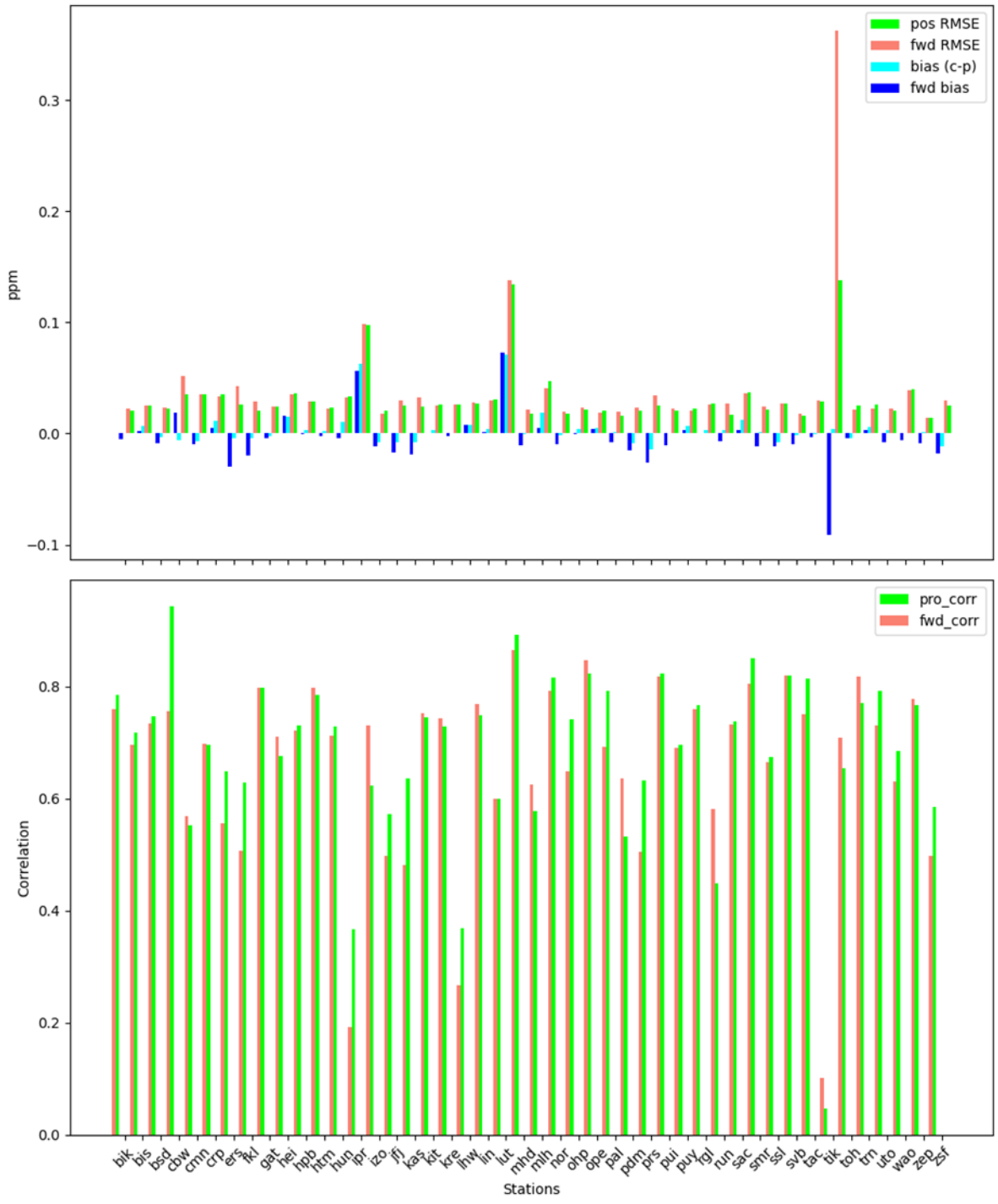


Figure 63: Statistical findings for the assimilated concentration and CIF-FLEXPART posterior and prior: RMSE and Bias (upper panel) and Correlation (lower panel).

In Figure 64, we present a spatial distribution map that depicts both the prior and posterior monthly methane flux at a finely gridded resolution of 0.20° x 0.20°. Additionally, we visualize the percentage change in the posterior flux compared to the prior estimates, which is

calculated as (posterior - prior) * 100 / prior. This figure offers valuable insights into the dynamics of methane flux within the region.

When examining the spatial distribution of methane fluxes, it becomes evident that regions such as the Netherlands, Belgium, Ireland, Italy, and Romania, as well as western France and the southern United Kingdom, exhibit elevated emissions. Meanwhile, in the northernmost parts of Europe, specifically within the Scandinavian countries of Finland and Sweden, the figure highlights distinct monthly variations in methane flux. This variability is attributed to the influence of seasonal factors on biospheric methane emissions in this area.

A comparison between the posterior flux and the prior data clearly reveals significant variations across different geographical regions. Notably, there is a substantial increase in the posterior methane flux over the Netherlands when contrasted with the prior estimates, suggesting a potential rise in methane emissions in this region. Conversely, Northern Italy shows a reduction in methane flux in the posterior data compared to the prior figures, indicating a possible decrease in methane emissions in this part of Italy.

Our inversion analysis provides an estimate of the total methane (CH_4) emissions budget for the EU27 + UK, ranging from 13.44 Tg of CH_4/year in January to a peak estimate of 17.31 Tg of CH_4/yr in July. The eight-month average estimate stands at approximately 15.55 Tg of CH_4/yr, which is consistent with the range of 17.5 ± 2.2 Tg CH4/yr as reported by Petrescu, A-M-R. et al. in 2023. The main contributors to methane emissions in the continent are France, Germany, Italy, Spain, Poland, and the UK. Collectively, these six countries are responsible for emitting approximately 11.12 Tg of CH_4/yr during the eight-month period, which accounts for 72% of the total emissions. For a more detailed breakdown by country and month, refer to Table 15 for the comprehensive monthly country-specific budget.

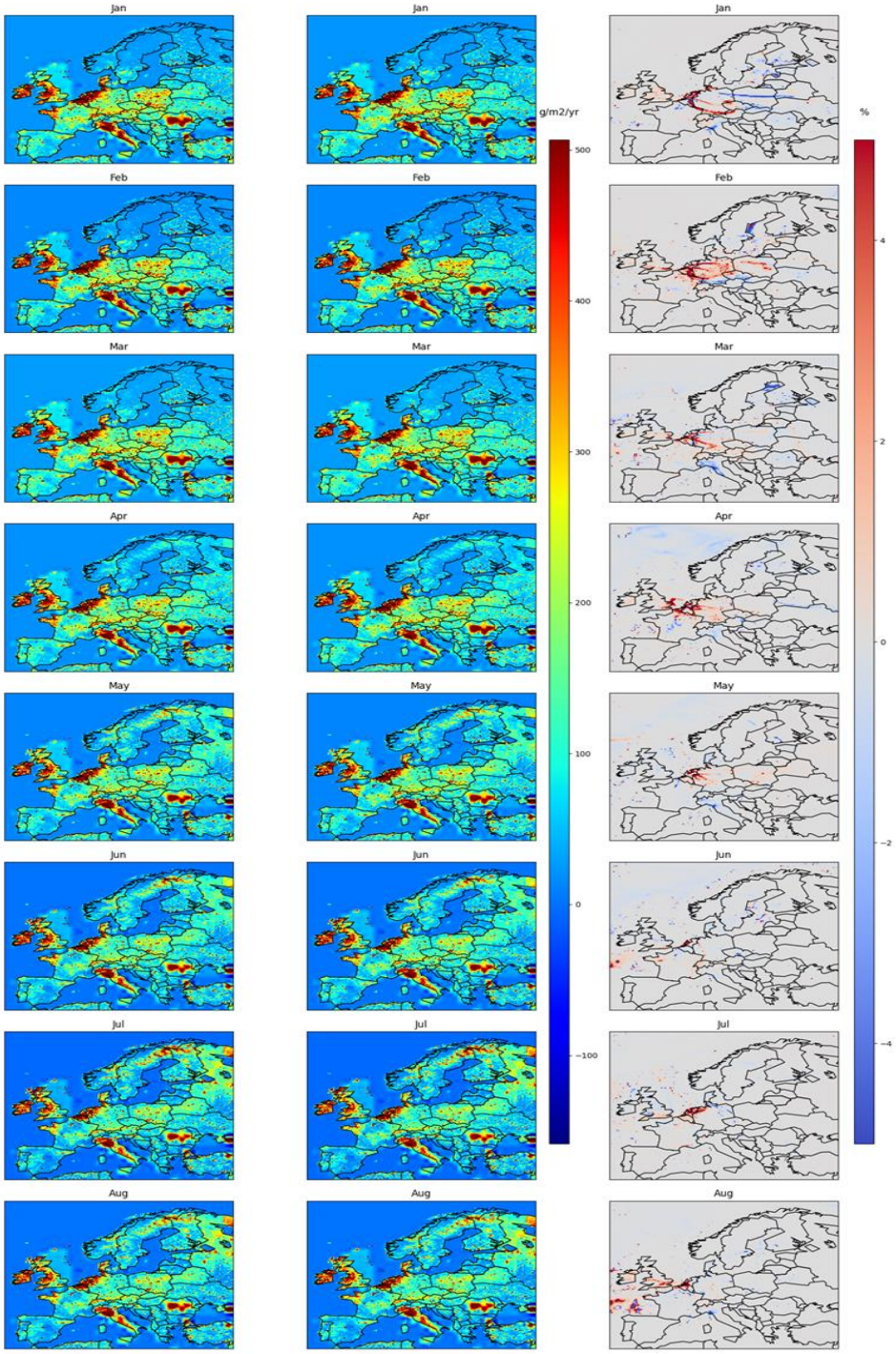


Figure 64: Monthly spatial distribution of CH_4 estimates from CIF-FLEXPART inversion at a higher resolution of 0.2° x 0.2° grid: prior (left panel), posterior (middle panel) and posterior increments computed as (posterior – prior)*100/prior (right panel) .

In addition, Figure 65 presents informative bar plots illustrating the country-specific budgets, allowing for a visual comparison of methane emissions among the various regions. This graphical representation provides a clear overview of how different countries within the EU28

+ UK region contribute to the overall methane budget, highlighting variations and trends in emissions across the months.

Table 15: CIF FLEXPART inversion estimates of country total CH_4 budget for EU28 and UK for the first eight months in the year 2018 in units of $TgCH_4$/year

Country	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Austria	0.1248	0.1060	0.0999	0.1052	0.0963	0.0987	0.0978	0.1011
Belgium	0.0457	0.0537	0.0570	0.0590	0.0512	0.0455	0.0463	0.0493
Bulgaria	0.1118	0.1199	0.1122	0.1023	0.1105	0.1160	0.1084	0.1084
Switzerland	0.0215	0.0229	0.0210	0.0209	0.0203	0.0216	0.0201	0.0213
Cyprus	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
Czechia	0.0770	0.0832	0.1362	0.1184	0.0997	0.0933	0.0989	0.0799
Germany	1.8032	2.7869	2.3440	2.5785	2.2820	2.1769	2.1395	2.0589
Denmark	0.0295	0.0303	0.0262	0.0271	0.0256	0.0271	0.0256	0.0295
Estonia	0.0100	0.0205	0.0152	0.0184	0.0209	0.0223	0.0243	0.0249
Greece	0.1426	0.1559	0.1305	0.0880	0.1356	0.1428	0.1483	0.1302
Spain	1.4245	1.5074	1.4753	1.3987	1.3541	1.4046	1.3237	1.3059
Finland	0.1385	0.7189	0.3618	0.3545	1.3387	1.6475	1.7087	1.8646
France	3.3559	3.4453	3.1961	3.2102	2.8082	3.0031	3.0950	2.9930
Croatia	0.0565	0.0736	0.0667	0.0708	0.0225	0.0350	0.0435	0.0654
Hungary	0.0775	0.0830	0.0754	0.0712	0.0682	0.0737	0.0696	0.0689
Ireland	0.1153	0.1261	0.1129	0.1192	0.1228	0.1355	0.1343	0.1307
Italy	1.5824	1.6362	1.3407	1.4950	1.2297	1.5213	1.5240	1.4848
Lithuania	0.0247	0.0275	0.0228	0.0252	0.0244	0.0267	0.0259	0.0237
Luxembourg	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
Latvia	0.0116	0.0126	0.0093	0.0118	0.0132	0.0145	0.0174	0.0162
Malta	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Netherlands	0.1093	0.0991	0.0879	0.0762	0.0568	0.0551	0.0645	0.0606
Poland	1.4598	1.8040	1.5609	1.7348	1.4031	1.4634	1.4816	1.4087
Portugal	0.0430	0.0435	0.0436	0.0420	0.0401	0.0421	0.0393	0.0395
Romania	0.9016	1.2468	0.9927	1.1261	1.1304	1.1649	1.1384	1.1086
Sweden	0.5343	0.5279	0.4988	0.8100	1.5327	1.8777	2.4667	2.0867
Slovenia	0.0054	0.0058	0.0052	0.0051	0.0048	0.0050	0.0047	0.0049
Slovakia	0.0414	0.0275	0.0491	0.0486	0.0437	0.0462	0.0456	0.0437
United Kingdom	1.1958	1.2804	1.1772	1.2857	1.2347	1.3615	1.4209	1.3674
Total	13.4438	16.0453	14.0191	15.0032	15.2705	16.6222	17.3131	16.6771
Total (FRA. GER.	10.8216	12.4601	11.0943	11.7027	10.3118	10.9308	10.9847	10.6188

ITA.POL.SPA, UK)								
% contribution (FRA. GER. ITA.POL.SPA, UK)	80.4953	77.6557	79.1371	78.0016	67.5277	65.7600	63.4474	63.6729

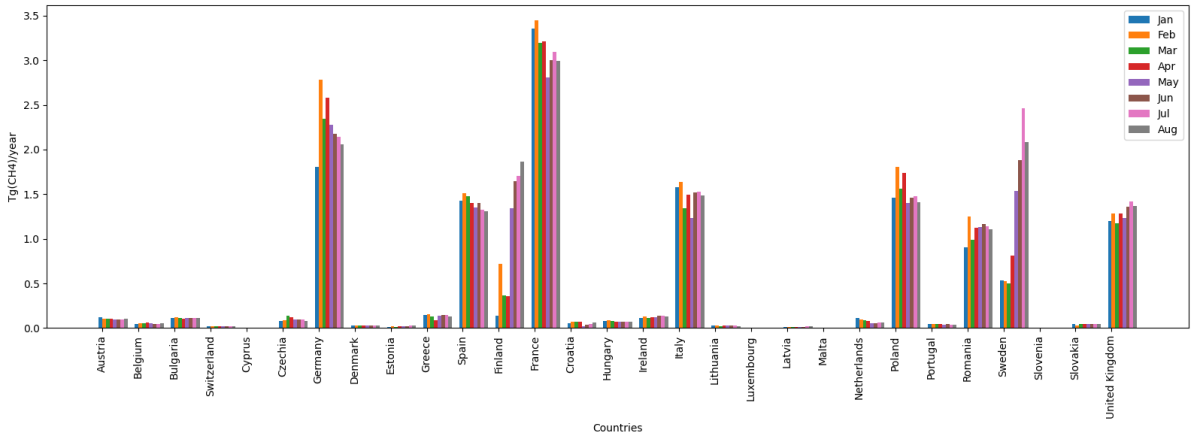


Figure 65: Bar plots illustrating the total CH_4 budget of European countries for the first eight months of 2018, derived from CIF-FLEXPART inversion.

3.11 WRF-CTDAS inversion system development (DLR and VUA)

The inversion system WRF-CTDAS was originally developed in the H2020 projects SCARBO and CHE (see https://che-project.eu/node/239). It is a coupling of the transport model WRF-GHG (Beck et al., 2011) to the Carbon Cycle Data Assimilation Shell (CTDAS; van der Laan-Luijkx et al., 2017). The optimization in CTDAS is carried out using an Ensemble Kalman Smoother (Peters et al., 2005). The work plan included inversions focusing on Germany (DLR) and Europe (VUA), as well the implementation of features for improving performance. Initially it had been foreseen to carry out the inversions focussing on Germany (DLR) with a coupling of WRF-GHG to the Common Inversion Framework, which was carried out in WP5. However, the development with the CIF team was delayed, and instead CTDAS-WRF was used as a fall-back option. We successfully implemented the new features, but due to personnel shortage, the full inversions for Germany (DLR) were not realised according to the protocol. Instead, we demonstrate the inverse model's new capabilities with a numerically small test case. Inversions for Europe have been carried out; however, they were affected by a bug that was discovered too late to rerun the VUA simulations. In the following sections, we first describe the model development (Section 3.11.1), followed by a description of model and results for the test case (DLR) to demonstrate the new capabilities (Sections 3.11.2 and

3.11.3), and model description and results of the European simulation centered on the Netherlands (VUA) that demonstrate the application in a setup adhering to the inversion protocol, albeit suffering from the aforementioned bug (Sections 3.11.4 and 3.11.5).

3.11.1 Inverse model development

At the start of CoCO2, WRF-CTDAS could assimilate column data and optimize small state vectors without correlations, e.g. estimate emissions for a few aggregated regions. This approach is prone to aggregation error. During CoCO2, we added the capability to optimize fluxes at high spatial resolution and to assimilate *in situ* data. The support for *in situ* data was originally written by Ioanna Evangelou (University of Crete) and Nikos Gialesakis (University of Crete and University of Bremen), and within CoCO2, we debugged and integrated it into the current WRF-CTDAS code. To support flux optimization at high spatial resolution, we added two capabilities to WRF-CTDAS: allowing prior flux covariances and replacing the existing localization algorithm with a computationally efficient one. These features were implemented in collaboration with Tzu-Hsin (David) Ho (MPI-BGC, Jena). Localization is an error prevention method for Ensemble Kalman Filters: since the state vector and its covariances are represented by a finite ensemble, spurious correlations can cause unphysical optimizations and thus degrade the estimated fluxes. The standard localization method in CTDAS is based on a t-test. We found that this approach is computationally not feasible for large state vectors and implemented an alternative localization method based on the distance between the observation and the state vector element location, drastically reducing computational time.

3.11.2 Model description (DLR test case)

To demonstrate the new capabilities that we integrated into WRF-CTDAS within CoCO2, we ran a test case in which we assimilated data from the ICOS network to estimate fluxes with prior spatial flux correlations. The configuration is described in the following sections.

3.11.2.1 Transport model

Here, we use the transport model WRF-GHG version 4.3.3. The domain of the test case covers portions of southern Europe in 12x15 grid cells with a resolution of 100 km, 39 vertical levels and a time step of 10 minutes. This configuration is designed to allow fast runs of the transport model instead of optimizing accuracy. The model is driven by meteorological fields from ERA5 (Hersbach et al. 2017, 2023a). Atmospheric transport is computed for an ensemble of 150 passive tracers.

3.11.2.2 Prior fluxes and assimilated observations

In this test case, we optimize CO_2 emissions. Prior fluxes are comprised of 3 datasets: anthropogenic emissions, gross primary production (GPP) and respiration. Anthropogenic emissions are taken from the TNO GHGco dataset (Super et al., 2020), with a resolution of 1/60 x 1/120°. GPP and respiration are from a run of the model VPRM, originally produced in the H2020 project CHE (https://www.che-project.eu/node/149), with a resolution of 5 km. These datasets have then been projected onto the WRF domain (Figure 66).

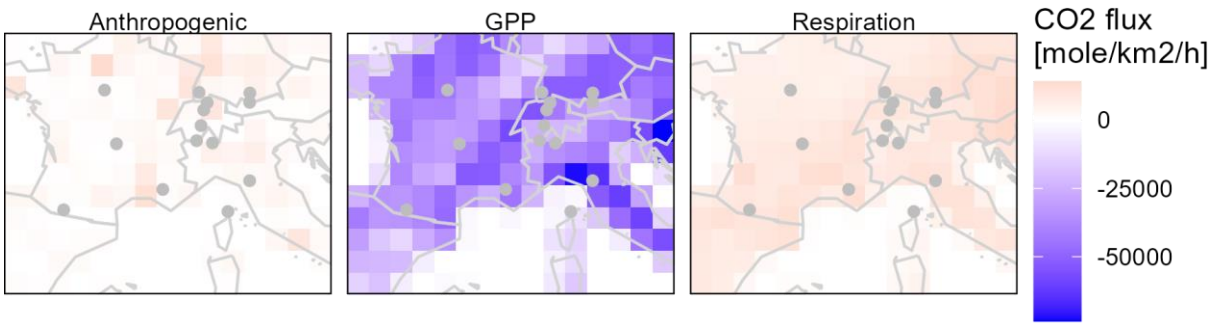


Figure 66: Prior emissions for the WRF-CTDAS test case and locations of stations from which data are assimilated. Shown here are values for 2015-06-01 12 UTC.

Assimilated observations are from the ICOS network at the following locations:

Table 16: ICOS station locations from which data were assimilated

Station code	Station name
BRM	Beromunster
CMN	Monte Cimone
ERS	Ersa
HPB	Hohenpeissenberg
IPR	Ispra
JFJ	Jungfraujoch
LHW	Laegern-Hochwacht
OHP	Observatoire de Haute Provence
PDM	Pic du Midi
PRS	Plateau Rosa
PUY	Puy de Dôme
SSL	Schauinsland
TRN	Trainou
ZSF	Zugspitze

3.11.2.3 State vector

The state vector consists of scaling factors for the three prior flux components at the same extent and resolution as the transport simulation grid, i.e. 12x15 grid cells at 100 km resolution. In addition, one boundary condition offset per corner is optimized, for a total of 544 state vector elements (3*12*15 + 4=544).
Prior uncertainties are 30% for anthropogenic emissions, 50% for GPP and respiration and 5 ppm for boundary conditions. Prior correlations of flux scaling factors are Gaussian with 400

km correlation length and 0.25 for the boundary condition offsets. The temporal resolution is 1 day with an assimilation window of 2 days.

3.11.2.4 Period of study

The test case inversion is run for 2 days, 1-2 June 2015.

3.11.3 Results (DLR test case)

Figure 67 shows prior and posterior fluxes as well as their difference, i.e. the innovation. As expected, the innovation is concentrated around the station locations, as is the uncertainty reduction (Figure 68). The uncertainty reduction of the anthropogenic flux component is much smaller than that of GPP and respiration.

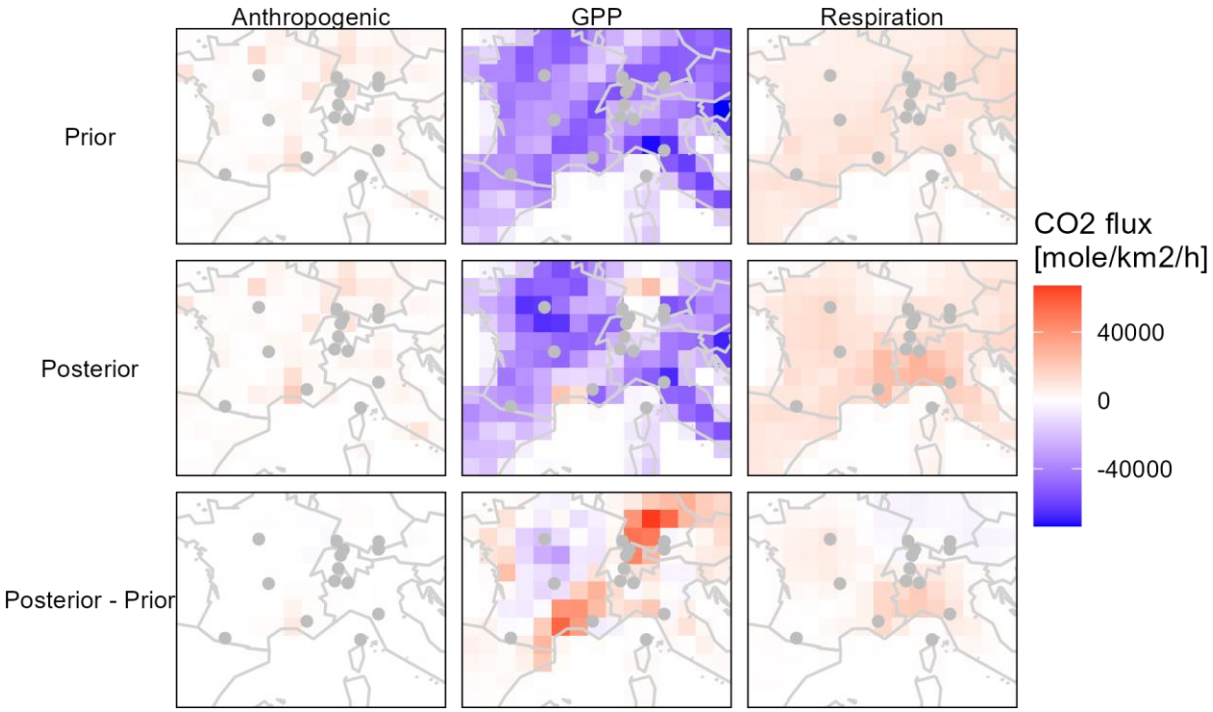


Figure 67: Prior, posterior and innovation fluxes for 2015-06-01 12 UTC.

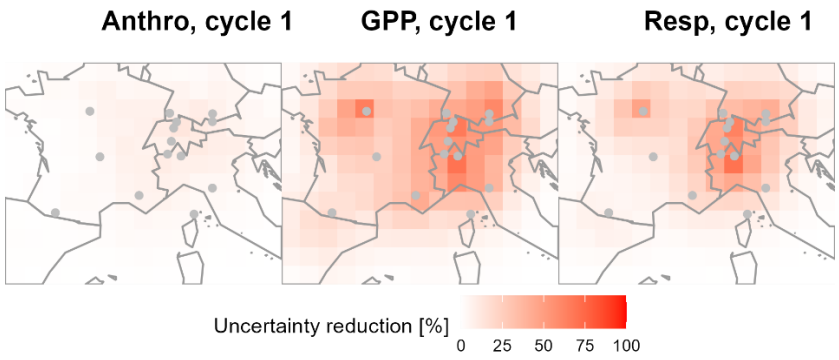


Figure 68: Uncertainty reduction for 2015-06-01

At a late stage of the project, we discovered a bug in the newly implemented spatial localisation code. The bug led to an incorrect mapping of localisation coefficients, which erroneously suppressed some state vector optimizations and erroneously allowed others, potential spurious ones (i.e. far from the stations). In Figure 67 and Figure 68, the bug has been corrected. However, it was too late to rerun the inversions by VUA (Section 3.11.5). Effectively, the error led to only partially optimized flux maps (cf. Figure 68 vs Figure 70). To illustrate the effect of the bug, below are plots of the test case while the bug was still present (Figure 69 and Figure 70). In this case, the uncertainty reductions are not related to the station locations and have a stripe pattern. For pixels whose uncertainty is still reduced, the posterior fluxes and innovation are similar to the ones of the debugged fluxes (cf. Fig. Figure 67 vs Figure 69). However, for pixels whose uncertainty was erroneously not reduced, the posterior fluxes are close to the priors.

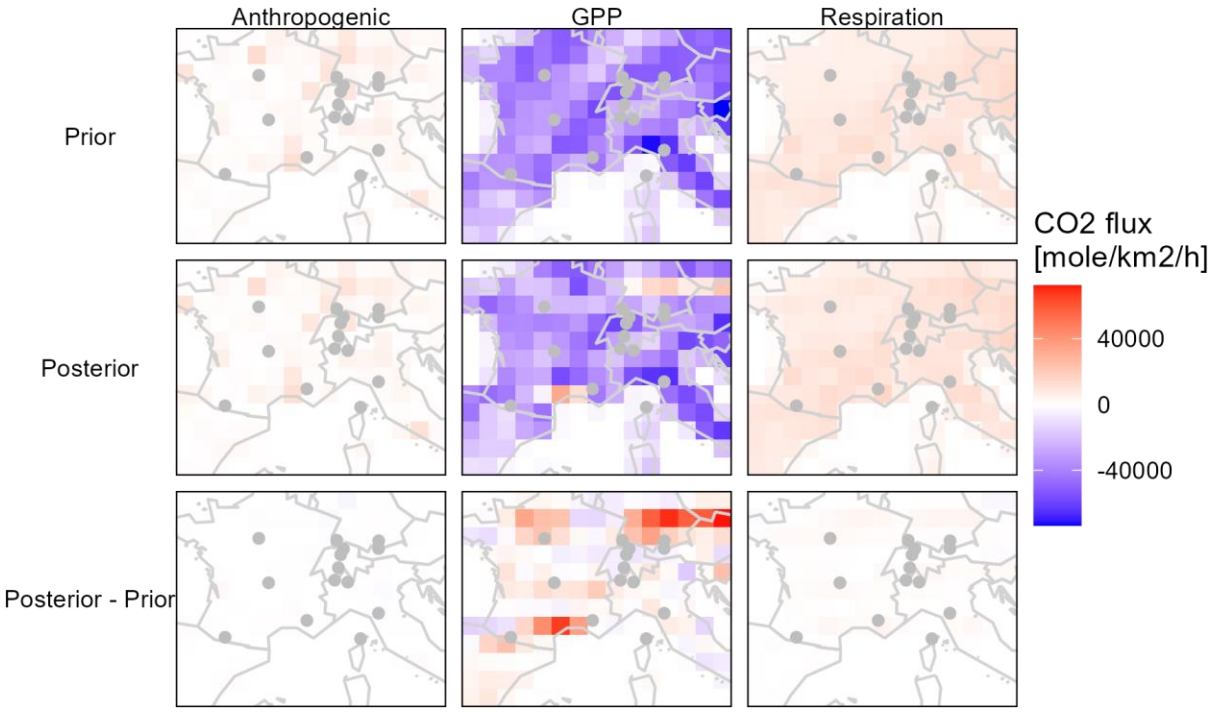


Figure 69: Same as Figure 67, but with the bug that also affected the results in Section 3.11.5

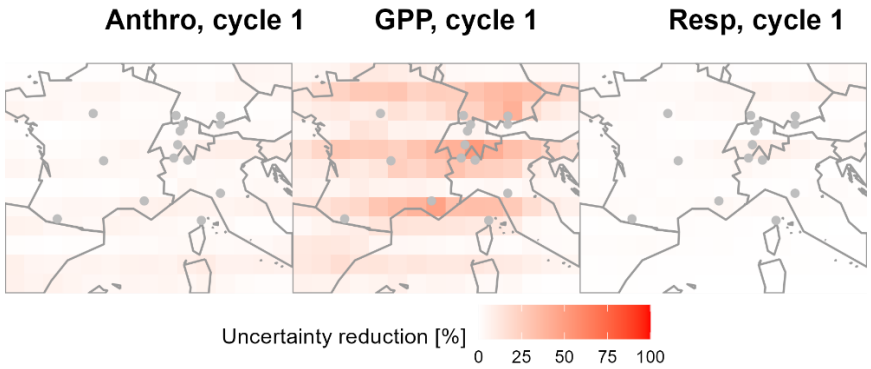


Figure 70: Same as Figure 68, but with the bug that also affected the results in Section 3.11.5

3.11.4 Model description (VUA Europe case)

3.11.4.1 Transport model

WRF v.4.4.1 is used in this study, coupled with chemistry (WRF-Chem) (Skamarock et al. 2019). It is a fully coupled, online meteorological and chemical transport mesoscale model (Grell et al. 2005). The greenhouse gases (GHG) module is used to calculate CO_2 transport (Beck et al., 2011). All other atmospheric chemistry options are switched off, as they are not needed for the long-lived trace gas CO_2, while state of art schemes are used for simulating atmospheric transport. The model is run over Europe using two domains (see Figure 71), where the parent domain (d01) covers almost all of Europe with a horizontal resolution of 30 km x 30 km, and the nested domain (d02), with a horizontal resolution 10 km x 10 km, partially covers Western Europe. The boundary and initial conditions are derived from CAMS for CO_2 and from ECMWF Reanalysis v5 (ERA5) for the meteorological data (Hersbach et al., 2023 a, b). 100 ensemble members and a total of 50 vertical levels are used, with approximately 12 vertical levels below the boundary layer, and spectral nudging for d01 and d02 is applied. WRF-Chem temperature, wind and humidity are nudged at each dynamical step toward the reanalysis and are updated every 3h above the atmospheric boundary layer.

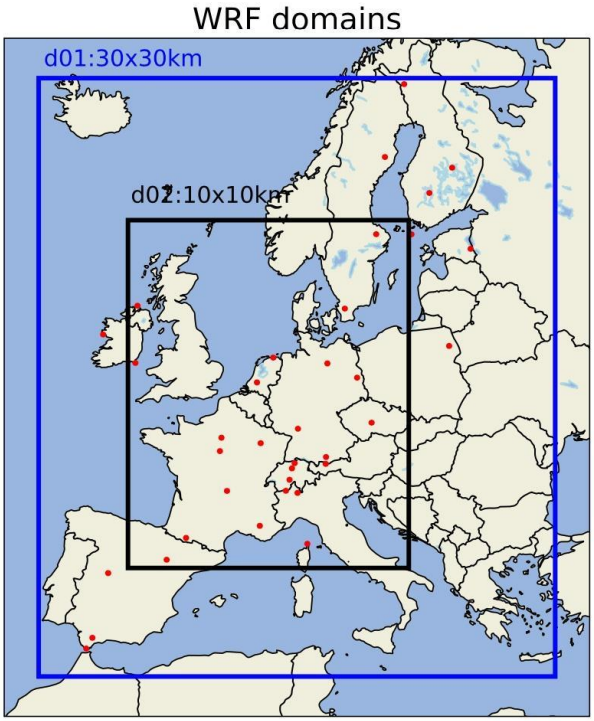


Figure 71: WRF-Chem domains over Europe. The red dots indicate the locations of the in-situ sites used for the inversions.

3.11.4.2 Prior fluxes and assimilated observations

As part of the simulations all available prior CO_2 fluxes from the task protocol are used; namely anthropogenic from TNO, biogenic from VPRM, fire from GFAS, oceanic from Mercator and lateral from LSCE. However, only anthropogenic, biogenic (respiration and gross primary production) and fire CO_2 fluxes are optimised, while the rest of the fluxes fixed in the simulations. The table below (Table 17) provides a summary of the different prior fluxes, as well as with their uncertainties applied in this study.

Table 17: CO_2 prior fluxes and the uncertainties used in this simulation.

Prior Fluxes	Resolution	Prior Uncertainties	Courtesy
Anthropogenic	6x6km	30%	TNO, Hugo Denier van der Gon
Biogenic (respiration & gross primary productivity)	1x1km	50%	VPRM, Julia Marshall
Fire	0.1x0.1	50%	GFAS, Copernicus
Oceanic	0.25x0.25	not optimised	Cyril Germineaud, Mercator
Lateral	8x8km	not optimised	LSCE (CEA/CNRS/UVSQ) - IPSL

Table 18 shows the in-situ observations used in this study and they mostly cover western and northern Europe. Normal stations are filtered for afternoon data and mountain stations for night time data, due to uncertainties on the simulation of the planetary boundary layer height.

The observational data were downloaded from the ICOS Carbon Portal. (Couret and Schmidt, 2023).

3.11.4.1 State vector

The CO_2 fluxes are optimised (see Table 17) using the cost function (J) as it is described in van der Laan-Luijkx et al. 2017.

Table 19 summarises the CTDAS configuration.

In total, 6224 flux parameters are optimised, splitting the domain into 1554 flux regions. Also 8 boundary condition parameters are optimised, with a sigma equal to 6 ppm for boundary condition optimisation. Finally, the correlation for neighbouring boundary condition parameters is set to 0.25.

3.11.4.1 Period of study

The model is run from 15th January to 28th February 2018, covering the full month of February 2018 (one of the two reference months in Task T4.4). The first two weeks are considered as spin-up and are excluded from the analysis.

Table 18: Information about the in-situ ICOS sites used for the inversion. X indicates if a station is a mountain station.

Station Name and inlet height (m)	Mountain station
Beromunster_212	
Bialystok_300	
Cabauw_207	
Carnsore Point_14	
El Estrecho_20	
Ersa_40	
Gartow_341	
Heidelberg_30	
Hohenpeissenberg_131	X
Hyltemossa_150	
Hyytiälä_125	
Ispra_100	
Järvselja_110	
Jungfraujoch_5	X
Karlsruhe_200	
Kresin u Pacova_250	
La Muela_80	
Laegern-Hochwacht_32	
Lindenberg_98	
Lutjewad_60	
Mace Head_24	
Malin Head_47	
Monte Cimone_8	X
Norunda_100	
Observatoire de Haute Provence_100	
Observatoire perenne de l'environnement_120	
Pallas_12	
Pic du Midi_28	X
Plateau Rosa_10	X
Puijo_47	
Puy de Dome_10	X
Saclay_100	
Sierra de Grazalema_20	
Svartberget_150	X
Trainou_180	
Utö - Baltic sea_57	
Zugspitze_3	X

Table 19: CTDAS configuration

Temporal resolution	Assimilation window	Correlation length	State vector resolution	Spatial correlations	Scaling factor
1 week	2 weeks	300km	100km	Yes	Separate for prior fluxes

3.11.5 Results (VUA Europe case)

One inversion is carried out in this study. The result section is split into two subsections. The first subsection presents the results over Europe while in the second the analysis is focused on the western Europe.

3.11.5.1 CO_2 inversion: Europe

Figure 72 shows prior and posterior anthropogenic and biogenic (respiration and gross primary production) CO_2 fluxes over Europe (d01), as well as the differences between posterior and prior for February 2018. The “stripe” patterns shown in Figure 72, for example in the differences between posterior and prior biogenic fluxes, is due to an error in the spatial localization code, meaning that some regions close to the observations are not optimised (for details see Section 3.11.3).

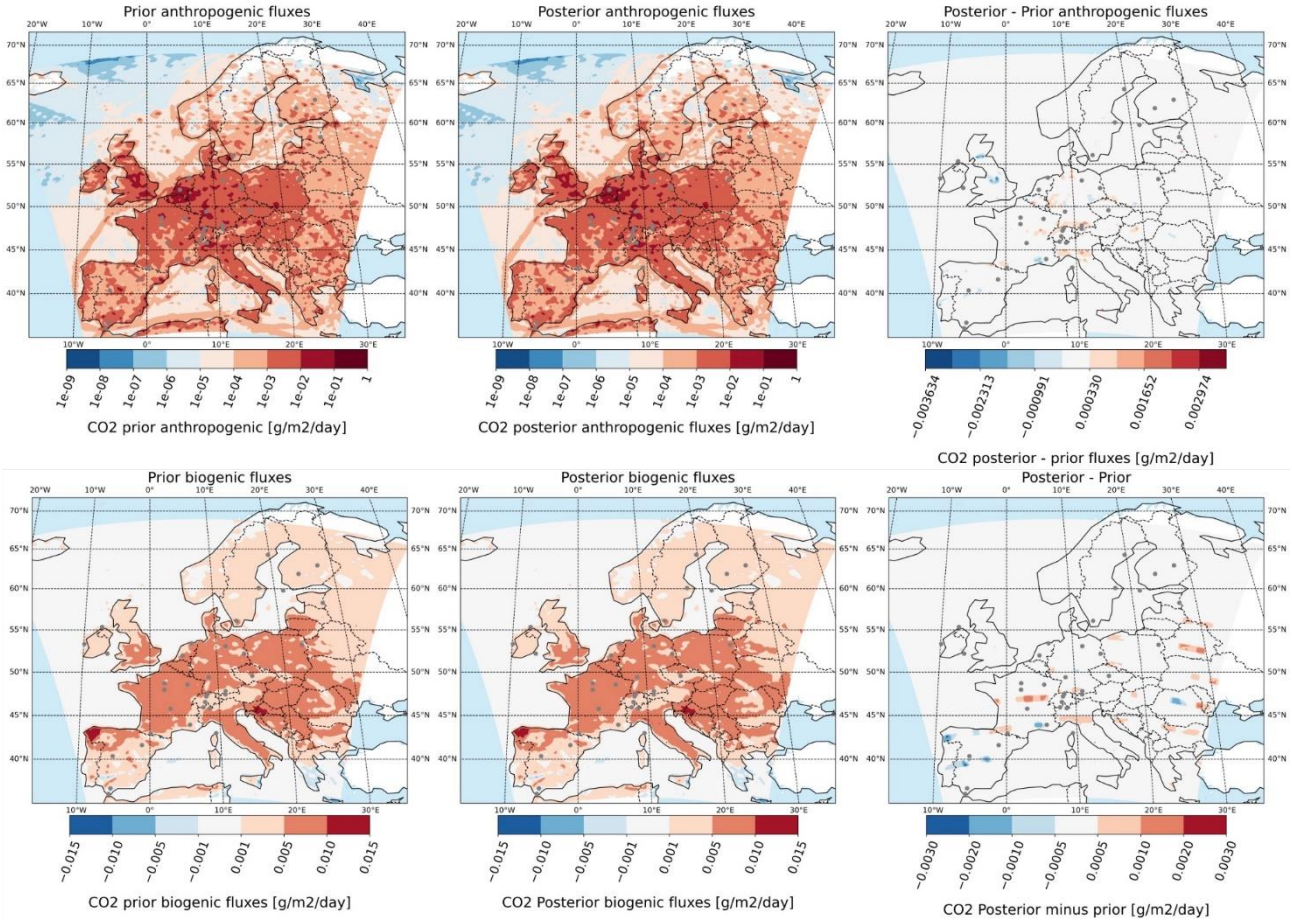


Figure 72: Upper panel: Average CO_2 anthropogenic fluxes. Lower panel: Average CO_2 biogenic fluxes, over Europe (d01) for February 2018. From left to right: Prior, posterior fluxes and the difference between posterior and prior. The grey dots indicate the in-situ observations used for the inversions.

Figure 73 shows the country-scaled prior and posterior anthropogenic and biogenic CO_2 fluxes averaged for February 2018 and for all the countries included in the domain.

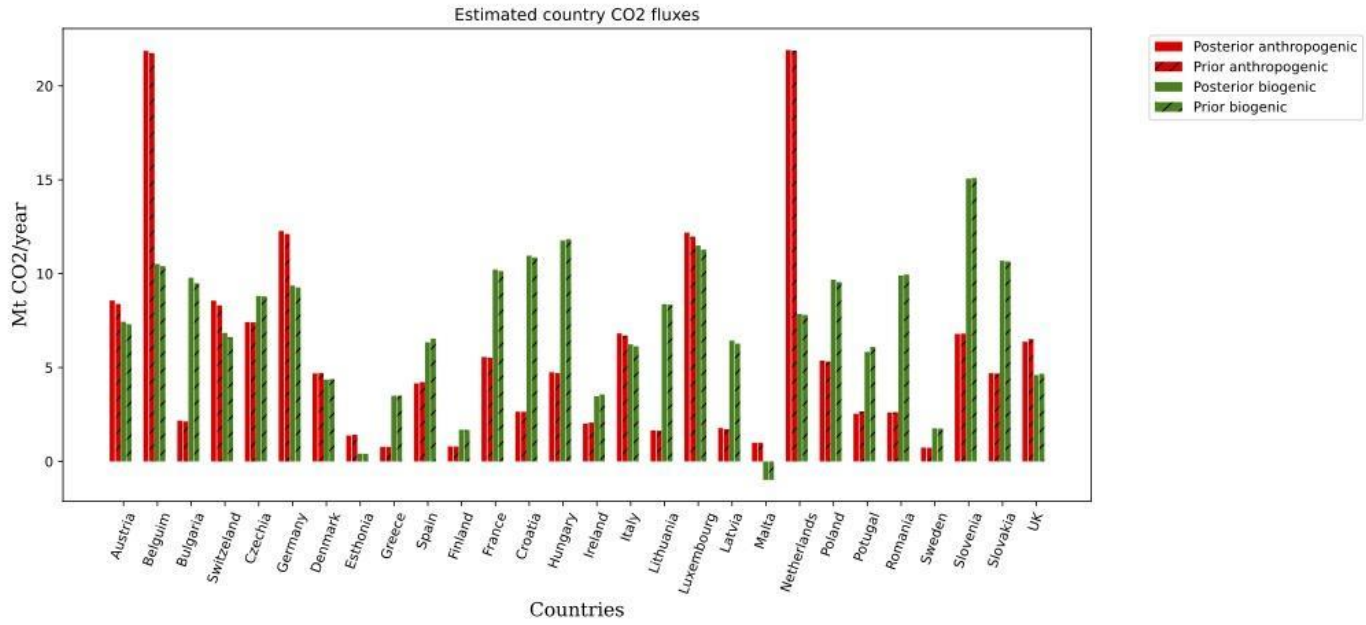


Figure 73: Averaged prior and posterior anthropogenic (in red) and biogenic fluxes (in green) for February 2018 and for the different countries included in the simulation domain (d01).

3.11.5.2 CO_2 inversion: Zoom in Western Europe

The Figure 74 shows prior and posterior anthropogenic and biogenic (respiration and gross primary production) CO_2 fluxes over western Europe (d02), as well as the differences between posterior and prior for February 2018.

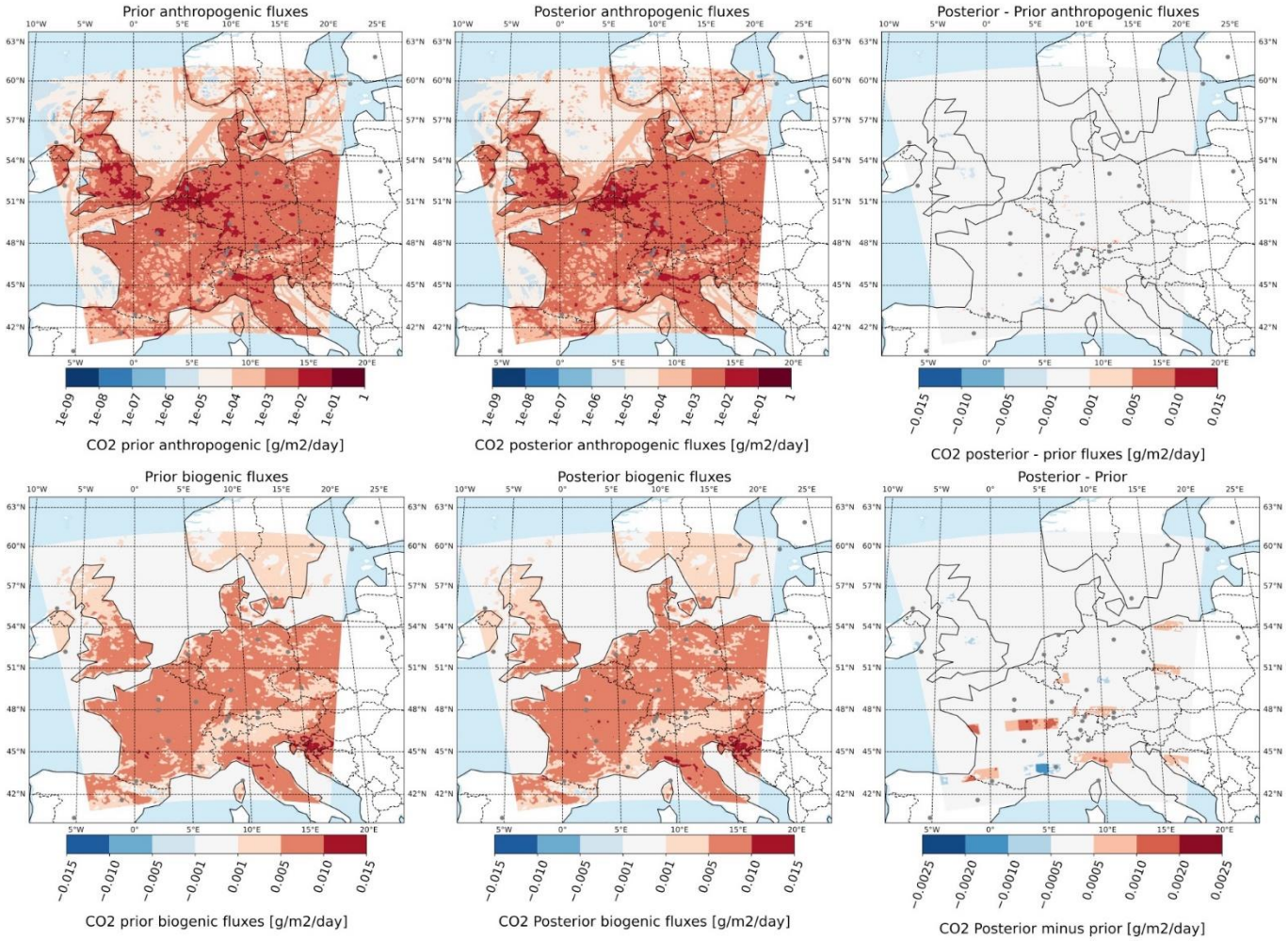


Figure 74: Upper panel: Average CO_2 anthropogenic fluxes. Lower panel: Average CO_2 biogenic fluxes, over western Europe (d02) and for February 2018. From left to right: Prior, posterior fluxes and the difference between posterior and prior. The grey dots indicate the in-situ observations used for the inversions.

Figure 75 shows country-scaled posterior anthropogenic and biogenic CO_2 fluxes for all the countries included in the domain.

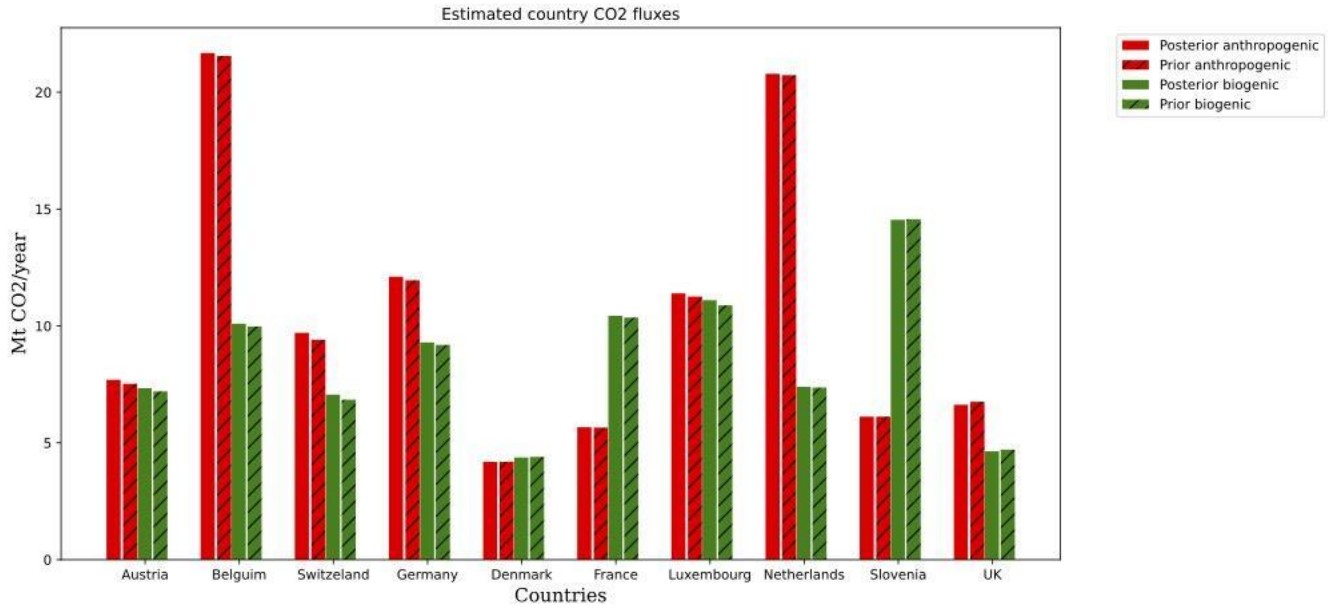


Figure 75: Averaged prior and posterior anthropogenic (in red) and biogenic fluxes (in green) for February 2018 and for the different countries included in the simulation domain (d02).

4 Synthesis and conclusions

Task 4.4 and this deliverable represent a major inverse modelling activity with the development of national-scale inverse modelling systems and their application over periods ranging from one month to several years. This offers a strong basis for the assessment of national-scale inversions. In order to tackle the national-scale targets of the task, the participants had to develop new capabilities or expand existing systems to enable:

- the separate control of the natural and anthropogenic CO_2 fluxes, which is innovative, since previous CO_2 regional inversion configurations generally fixed the anthropogenic emissions and focused on controlling the biogenic CO_2 fluxes
- the assimilation of satellite data (in addition to *in situ* data) at regional scale. While such data have been used in global and local scale inversions for some years, their use for regional-scale inversions has emerged more recently
- the (co-)assimilation of *in situ* and satellite observation of species co-emitted with CO_2 during fossil fuel combustion (CO and NOx)
- the modelling of the transport and the control of the fluxes at horizontal resolutions down to 10 km; continental to national-scale inverse modelling configurations were often based on coarser resolutions, notably 0.5° resolution for European-scale inversions

The different systems use different data assimilation techniques (specifically implementations of variational, Ensemble Kalman Filter and analytical approaches), strategies for the definition of the control vector (with varying temporal, spatial and/or sectoral resolutions), statistics of the uncertainties in the prior estimate of the fluxes, statistics of the observation and model errors, types of transport models, data selection, and components of the prior and fixed flux datasets (in practice, even though the main flux components were taken from the common database established with the T4.4 protocol). Some of them cover all of Europe (or the USA in the case of one system), others have domains focused on one or a few countries, with spatial resolutions ranging from 0.5° to 5 km for the transport, and from 0.5° to 10 km for the inversion.

Therefore, this deliverable provides good insights on the variety of the current options and practices for regional and national-scale flux estimation. The heterogeneity of approaches makes it difficult to attribute differences amongst the results from the different systems to specific characteristics of the inverse modelling setup. However, despite these differences, many of the systems provide consistent insights regarding the key aims of this deliverable, as detailed below. An in-depth quantitative inter-comparison of such a large number of inversions would have been too ambitious in the frame of a single project task, especially since the time required for the system development delayed the delivery of many of the inversion results to the last months of the project. The results demonstrate the need for a systematic analysis of the impact of the different inverse modelling components and design choices as initiated in Tasks 5.3 and 5.6, requiring extensive ensemble analysis with modular inverse modelling platforms such as the Community Inversion Framework (CIF; see D5.3). The results from the different inversion systems are now available for further analysis beyond Task 4.4 of CoCO2.

The following discusses the major insights from the set of inversions in D4.6.

The lack of control of the CO2 anthropogenic emissions at the annual to monthly / national scale

This main highlight from D4.6 is supported by the results from five (CIF-CHIMERE, ICON-ART CTDAS, WRF-analytical, WRF-CTDAS and GEOS-CHEM EnKF) out of six of the CO_2 inversion systems using real data. (The seventh CO_2 inversion system, i.e., TRACE EnKF, is only used for OSSEs without a separate control of the anthropogenic and natural CO_2 fluxes). This major conclusion applies in inversions assimilating the existing surface data, satellite data or both surface and satellite data. Only the biofos inversions with the LUMIA system apply major corrections to the CO_2 anthropogenic emissions, but these corrections appear to be unrealistic, and are attributed to the lack of robust statistical characterization of the respective uncertainties in the anthropogenic emissions and natural fluxes (see Section 3.6).

The current capabilities of the use of the co-emitted species do not enhance significantly the potential to monitor the CO_2 anthropogenic emissions at large spatial and temporal scales despite promising initial steps towards their co-assimilation. Actually, the use of the co-emitted species, and their co-assimilation bring new dimensions of uncertainties associated with the chemistry modelling to the already large uncertainties in the current satellite CO and NO_2 observations and the emission ratios across species (see Sections 3.2 and 3.9).

One might assume that the lack of correction of the CO_2 anthropogenic emissions by the inversions is due to the high accuracy of the inventories used as prior estimates of these emissions at annual and national scales in Europe. However, the lack of correction also applies at the monthly scale and at sub-national scales over most of the territories of the European countries, despite the significant uncertainties in the spatial and temporal disaggregation of the emissions at such scales.

However, systems such as CIF-CHIMERE, ICON-ART-CTDAS and GEOS-CHEM show some local control of the CO_2 anthropogenic emissions (see Sections 3.2, 3.3 and 3.9), and some sensitivity of the results for the NEE to the local anthropogenic emissions (see Section 3.2.2), especially in the areas with the largest local sources, in northwest Germany and Benelux, and especially when the transport model resolution is fine (at 5 to 10 km resolution, see Section 3.5.2). This implies that the capability to monitor national-scale CO_2 anthropogenic relies on exhaustively covering the local CO_2 anthropogenic sources. The current inverse modelling approach lacks the capability to exploit large scale variations in the observations (at least when using the existing observation systems) for the derivation of anthropogenic emissions estimates since the signal from the natural fluxes dominates over such scales. This still applies when analyzing the observations at 10 km resolution, which motivates the use of transport models at finer resolution, able to simulate the plumes from individual sources or clusters of sources, and for inverting the emissions at finer resolution, such as when tackling local scale

inversions. It also motivates the use of observations strongly connected to specific local sources in the national-scale inversion systems, such as those from urban/peri-urban stations dedicated to a given urban area. The ICOS stations that have been used throughout the simulations reported here were not planned with a focus on anthropogenic emission monitoring, which may contribute to the difficulty of controlling these fluxes using these data. Further observational constraint on local sources should come through the exploitation of observed plumes in spaceborne data, such as CO2M.

Ultimately, this requires a strong coupling between national and local inversions, the former relying much on the latter for the quantification of the CO_2 anthropogenic emissions. In parallel we need to progress on the inversions of co-emitted species and on their co-assimilation.

Analysis at high resolution over national domains is challenging: the accurate description of the statistics of uncertainties in the prior inventories of the emissions becomes more challenging when the spatial resolution increases, due to the heterogeneity of the emissions. This adds to the need for improved characterization of the correlations across the prior uncertainties in the emissions of the different co-emitted species when tackling the co-assimilation of co-emitted species.

The situation is different for the CH_4 emissions since the anthropogenic emissions are not known a priori as well as in case of CO_2, and since the anthropogenic and natural CH_4 emissions have more balanced and distinct spatial distributions and atmospheric signature. Therefore, the CH_4 inversions demonstrate some capabilities to infer these two types of emissions separately even at annual and national scale.

The current uncertainties the inversion of the CO_2 terrestrial ecosystem fluxes and CH_4 emissions

The second major highlight from the set of inversions conducted in T4.4 is the large spread of the CO_2 NEE and CH_4 emission estimates across the inversions.

Despite using the same or similar prior estimates of the NEE, the ICON-ART CTDAS CO_2 inversions, the CIF-CHIMERE CO_2 reference inversions, the LUMIA “bio” inversion and the GEOS-CHEM EnKF CO_2 inversions provide annual NEE budgets for France in 2018 ranging from -537MtCO_2/yr to -135MtCO_2/yr (the LUMIA biofos inversion yielding a net source of 35 MtCO_2/yr). Most of these inversions tend to reduce the likely too large sinks modelled by VPRM. However, the results vary not only across the inversion systems, but also within a given system when using the surface data, the satellite data or both. The sole convergence of the results is shown with the tests of sensitivity to the prior estimates of the fluxes with the CIF-CHIMERE system: with a given observation dataset and several prior estimates of NEE, the spread of the inversion results tends to be smaller than that of the prior estimates (see Section 3.2.2.2). As also highlighted by the results in Task 5.6, the situation is similar for the CH_4 inversions but here, none of the CH_4 inversion assimilates satellite data and only two of them provide annual emissions.

This reveals that despite the use of finer resolution systems here than in previous inter-comparisons (such as in EUROCOM, Monteil et al., 2020 and Thompson et al., 2020 or VERIFY, McGrath et al., 2023), and despite the regular progress in the inversion codes and configurations, the uncertainties in the regional-scale inversion of the CO_2 natural fluxes or of the CH_4 emissions remains high. The inter-comparison of the national scale inversions here continues the efforts conducted in these past projects and in WP5 (Tasks 5.3 and 5.6).

The relative weight of the satellite OCO-2 data and surface data for the CO_2 inversions highly depends on the inversion systems and configurations. While, overall, the weight of the surface data appears to be larger than that of the satellite data in the CIF-CHIMERE reference inversions, and while the joint assimilation of the two types of datasets does not alter the fit to each dataset in these inversions, it is the opposite for the ICON-ART CTDAS CO_2 inversions. Some patterns of the ICON-ART CTDAS inversions are very different when using the satellite data or the surface data, which reveals some biases between the information from the

assimilation of these two datasets (see Section 3.3.2). The regional-scale assimilation of the satellite data being more recent than that of the surface data, the configurations of the inversions to tackle these data tend to be less robust than those when assimilating surface data, which benefited from much experience from past studies. There is probably a need to progress on the assimilation of satellite XCO_2 data in regional systems before more consistent and complementary results from the two datasets and improved co-assimilation of the satellite and surface data can be realised.

The insights brought by this deliverable for the development of the CO2MVS multi-scale inversion configuration, and how they can serve as a basis for the development of national operational systems is discussed further in Deliverable 4.8.

5 References

Ahmadov, R., C. Gerbig, R. Kretschmer, S. Koerner, B. Neininger, A. J. Dolman, and C. Sarrat (2007), Mesoscale covariance of transport and CO2 fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model, J. Geophys. Res., 112, D22107, doi:10.1029/2007JD008552.

Beck, V., T. Koch, R. Kretschmer, J. Marshall, R. Ahmadov, C. Gerbig, D. Pillai, and M. Heimann, (2011): The WRF Greenhouse Gas Model (WRF-GHG). Technical Report No. 25, Max Planck Institute for Biogeochemistry, Jena, Germany. https://www.bgc-jena.mpg.de/bgc-systems/pmwiki2/uploads/Download/Wrf-ghg/WRF-GHG_Techn_Report.pdf

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, https://doi.org/10.5194/gmd-14-5331-2021, 2021.

Bergamaschi, P., Segers, A., Brunner, D., Haussaire, J.-M., Henne, S., Ramonet, M., Arnold, T., Biermann, T., Chen, H., Conil, S., Delmotte, M., Forster, G., Frumau, A., Kubistin, D., Lan, X., Leuenberger, M., Lindauer, M., Lopez, M., Manca, G., Müller-Williams, J., O'Doherty, S., Scheeren, B., Steinbacher, M., Trisolino, P., Vítková, G., and Yver Kwok, C.: High-resolution inverse modelling of European CH_4 emissions using the novel FLEXPART-COSMO TM5 4DVAR inverse modelling system, Atmos. Chem. Phys., 22, 13243–13268, https://doi.org/10.5194/acp-22-13243-2022, 2022.

Broquet, G., F. Chevallier, P. Rayner, C. Aulagnier, I. Pison, M. Ramonet, M. Schmidt, A. T. Vermeulen, and P. Ciais(2011), A European summertime CO2biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements,J. Geophys. Res.,116, D23303, doi:10.1029/2011JD016202.

Broquet, G., Chevallier, F., Bréon, F.-M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO2 ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates, Atmos. Chem. Phys., 13, 9039–9056, https://doi.org/10.5194/acp-13-9039-2013, 2013.

Chen, H. W., F. Zhang, T. Lauvaux, K. J. Davis, S. Feng, M. P. Butler, and R. B. Alley (2019): Characterization of regional-scale CO2 transport uncertainties in an ensemble with flow-dependent transport errors. Geophysical Research Letters, 46, 4049–4058, doi:10.1029/2018GL081341.

Chen, H. W., F. Zhang, T. Lauvaux, M. Scholze, K. J. Davis, and R. B. Alley (2023): Regional CO2 inversion through ensemble-based simultaneous state and parameter estimation: TRACE framework and controlled experiments. Journal of Advances in Modeling Earth Systems, 15, e2022MS003208, doi:10.1029/2022MS003208.

Chevallier, F., Broquet, G., Zheng, B., Ciais, P., & Eldering, A. (2022). Large CO2 emitters as seen from satellite: Comparison to a gridded global emission inventory. Geophysical Research Letters, 49, e2021GL097540. https://doi.org/10.1029/2021GL097540

Denier van der Gon et al., CoCO2 Deliverable 2.1, https://coco2-project.eu/sites/default/files/2022-03/CoCO2-D2.1-V1-0.pdf

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, https://doi.org/10.5194/gmd-14-2939-2021, 2021.

Fortems-Cheiney, A., Broquet, G., Potier, E., Plauchu, R., Berchet, A., Pison, I., Denier van der Gon, H. A. C., and Dellaert, S. N. C.: The CO anthropogenic emissions in Europe from 2011 to 2021: insights from the MOPITT satellite data, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-1981, 2023.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled “online” chemistry within the wrf model. Atmospheric Environment, 39 (37), 6957–6975

Hersbach, H. et al. (2017) Complete ERA5 from 1940: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service (C3S) Data Store (CDS), DOI:10.24381/cds.143582cf

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023a): ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.adbb2d47 (Accessed on 15-04-2023)

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023b): ERA5 hourly data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.bd0915c6 (Accessed on 15-04-2023)

ICOS RI, Bergamaschi, P., Colomb, A., De Mazière, M., Emmenegger, L., Kubistin, D., Lehner, I., Lehtinen, K., Lund Myhre, C., Marek, M., Platt, S.M., Plaß-Dülmer, C., Schmidt, M., Apadula, F., Arnold, S., Blanc, P.-E., Brunner, D., Chen, H., Chmura, L., Conil, S., Couret, C., Cristofanelli, P., Delmotte, M., Forster, G., Frumau, A., Gheusi, F., Hammer, S., Haszpra, L., Heliasz, M., Henne, S., Hoheisel, A., Kneuer, T., Laurila, T., Leskinen, A., Leuenberger, M., Levin, I., Lindauer, M., Lopez, M., Lunder, C., Mammarella, I., Manca, G., Manning, A., Marklund, P., Martin, D., Meinhardt, F., Müller-Williams, J., Necki, J., O'Doherty, S., Ottosson-Löfvenius, M., Philippon, C., Piacentino, S., Pitt, J., Ramonet, M., Rivas-Soriano, P., Scheeren, B., Schumacher, M., Sha, M.K., Spain, G., Steinbacher, M., Sørensen, L.L., Vermeulen, A., Vítková, G., Xueref-Remy, I., di Sarra, A., Conen, F., Kazan, V., Roulet, Y.-A., Biermann, T., Heltai, D., Hensen, A., Hermansen, O., Komínková, K., Laurent, O., Levula, J., Pichon, J.-M., Smith, P., Stanley, K., Trisolino, P., ICOS Carbon Portal, ICOS Atmosphere Thematic Centre, ICOS Flask And Calibration Laboratory, ICOS Central Radiocarbon Laboratory, 2023. European Obspack compilation of atmospheric carbon dioxide data from ICOS and non-ICOS European stations for the period 1972-2023; obspack_co2_466_GLOBALVIEWplus_v8.0_2023-04-26. https://doi.org/10.18160/CEC4-CAGK

Manders, A. M. M., Builtjes, P. J. H., Curier, L., Denier van der Gon, H. A. C., Hendriks, C., Jonkers, S., Kranenburg, R., Kuenen, J. J. P., Segers, A. J., Timmermans, R. M. A., Visschedijk, A. J. H., Wichink Kruit, R. J., van Pul, W. A. J., Sauter, F. J., van der Swaluw, E., Swart, D. P. J., Douros, J., Eskes, H., van Meijgaard, E., van Ulft, B., van Velthoven, P., Banzhaf, S., Mues, A. C., Stern, R., Fu, G., Lu, S., Heemink, A., van Velzen, N., and Schaap, M.: Curriculum vitae of the LOTOS–EUROS (v2.0) chemistry transport model, Geosci. Model Dev., 10, 4145–4173, https://doi.org/10.5194/gmd-10-4145-2017, 2017.

McGrath, M. J., Petrescu, A. M. R., Peylin, P., Andrew, R. M., Matthews, B., Dentener, F., Balkovič, J., Bastrikov, V., Becker, M., Broquet, G., Ciais, P., Fortems-Cheiney, A.,

Ganzenmüller, R., Grassi, G., Harris, I., Jones, M., Knauer, J., Kuhnert, M., Monteil, G., Munassar, S., Palmer, P. I., Peters, G. P., Qiu, C., Schelhaas, M.-J., Tarasova, O., Vizzarri, M., Winkler, K., Balsamo, G., Berchet, A., Briggs, P., Brockmann, P., Chevallier, F., Conchedda, G., Crippa, M., Dellaert, S. N. C., Denier van der Gon, H. A. C., Filipek, S., Friedlingstein, P., Fuchs, R., Gauss, M., Gerbig, C., Guizzardi, D., Günther, D., Houghton, R. A., Janssens-Maenhout, G., Lauerwald, R., Lerink, B., Luijkx, I. T., Moulas, G., Muntean, M., Nabuurs, G.-J., Paquirissamy, A., Perugini, L., Peters, W., Pilli, R., Pongratz, J., Regnier, P., Scholze, M., Serengil, Y., Smith, P., Solazzo, E., Thompson, R. L., Tubiello, F. N., Vesala, T., and Walther, S.: The consolidated European synthesis of CO2 emissions and removals for the European Union and United Kingdom: 1990–2020, Earth Syst. Sci. Data, 15, 4295–4370, https://doi.org/10.5194/essd-15-4295-2023, 2023.

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, https://doi.org/10.5194/gmd-6-981-2013, 2013.

Monteil, G., Broquet, G., Scholze, M., Lang, M., Karstens, U., Gerbig, C., Koch, F.-T., Smith, N. E., Thompson, R. L., Luijkx, I. T., White, E., Meesters, A., Ciais, P., Ganesan, A. L., Manning, A., Mischurow, M., Peters, W., Peylin, P., Tarniewicz, J., Rigby, M., Rödenbeck, C., Vermeulen, A., and Walton, E. M.: The regional European atmospheric transport inversion comparison, EUROCOM: first results on European-wide terrestrial carbon fluxes for the period 2006–2015, Atmos. Chem. Phys., 20, 12063–12091, https://doi.org/10.5194/acp-20-12063-2020, 2020.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., ... & Tans, P. P. (2005). An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations. Journal of Geophysical Research: Atmospheres, 110(D24).

Petrescu, A. M. R., Qiu, C., McGrath, M. J., Peylin, P., Peters, G. P., Ciais, P., ... & Zaehle, S. (2022). The consolidated European synthesis of CH4 and N2O emissions for EU27 and UK: 1990–2020. Earth System Science Data Discussions, 2022, 1-97, https://doi.org/10.5194/essd-15-1197-2023

Pisso, I., et al. The Lagrangian particle dispersion model FLEXPART version 10.4, Geosci. Model Dev., 12, 4955–4997, https://doi.org/10.5194/gmd-12-4955-2019, 2019.

Potier, E., Broquet, G., Wang, Y., Santaren, D., Berchet, A., Pison, I., Marshall, J., Ciais, P., Bréon, F.-M., and Chevallier, F.: Complementing XCO2 imagery with ground-based CO2 and 14CO2 measurements to monitor CO2 emissions from fossil fuels on a regional to local scale, Atmos. Meas. Tech., 15, 5261–5288, https://doi.org/10.5194/amt-15-5261-2022, 2022.

Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmos. Chem. Phys., 9, 5331–5342, https://doi.org/10.5194/acp-9-5331-2009, 2009.

Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., Reum, F., Keeling, R. F., and Heimann, M.: Interannual sea–air CO2 flux variability from an observation-driven ocean mixed-layer scheme, Biogeosciences, 11, 4599– 4613, https://doi.org/10.5194/bg-11-4599-2014, 2014.

Rodgers, C. D., Inverse Methods for Atmospheric Sounding: Theory and Practice, 238 pp., World Sci., River Edge, N. J., 2000.

Scholze, M., Kaminski, T., Knorr, W., Voßbeck, M., Wu, M., Ferrazzoli, P., et al. (2019). Mean European carbon sink over 2010–2015 estimated by simultaneous assimilation of atmospheric CO_2, soil moisture, and vegetation optical depth. *Geophysical Research Letters*, 46, 13796–13803. https://doi.org/10.1029/2019GL085725

Scholze, M., Houweling, S., Karstens, U., Meesters, A.G.C.A., Monteil, G. Quantification of uncertainty ranges from European multi-model inversions and ways to benchmark inversion systems. CoCO2 Deliverable 5.6, 2023.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, Z. Liu, J. Berner, W. Wang, J. G. Powers, M. G. Duda, D. M. Barker, and X.-Y. Huang, 2019: A Description of the Advanced Research WRF Version 4. NCAR Tech. Note NCAR/TN-556+STR, 145 pp. doi:10.5065/1dfh-6p97

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C., and Mukai, H.: The CO2 release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, Atmos. Chem. Phys., 11, 6855–6870, https://doi.org/10.5194/acp-11-6855-2011, 2011.

Steiner, M., Peters, W., Luijkx, I., Henne, S., Chen, H., Hammer, S., & Brunner, D. (2023). European CH 4 inversions with ICON-ART coupled to CarbonTracker Data Assimilation Shell. EGUsphere, 2023, 1-40.

Stohl, A., Hittenberger, M., and Wotawa, G.: Validation of the Lagrangian particle dispersion model FLEXPART against largescale tracer experiment data, Atmos. Environ., 32, 4245–4264, 1998, https://doi.org/10.1016/S1352-2310(98)00184-8

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 CO2 and CO to support inverse modelling and network design, Atmos. Chem. Phys., 20, 1795–1816, https://doi.org/10.5194/acp-20-1795-2020, 2020.

Super, I. , Scarpelli. T., Droste, A., and Palmer, P. I.: Improved definition of prior uncertainties in CO2 and CO fossil fuel fluxes and the impact on a multi-species inversion with GEOS-Chem (v12.5), under review at GMD

Thompson, R. L., Broquet, G., Gerbig, C., Koch, T., Lang, M., Monteil, G., Munassar, S., Nickless, A., Scholze, M., Ramonet, M., Karstens, U., Schaik, E. V., Wu, Z., & Rödenbeck, C. (2020). Changes in net ecosystem exchange over Europe during the 2018 drought based on atmospheric observations. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2019.0512.

van Der Laan-Luijkx, I. T., Van Der Velde, I. R., Van Der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., Fang Zhang, H., Liu, Y., He, W., Chen, H., Masarie, K. A., Krol, M. C. and Peters, W.: The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: Implementation and global carbon balance 2001-2015, Geosci. Model Dev., 10(7), 2785–2800, doi:10.5194/gmd-10-2785- 2017, 2017.

Document History

Version	Author(s)	Date	Changes
	Name (Organisation)	dd/mm/yyyy	
1	Julia Marshall and Friedemann Reum (DLR), Gregoire Broquet, Elise Potier, Antoine Berchet and Audrey Fortems-Cheiney (CEA), Erik Koene and Dominik Brunner (EMPA), Sander Houweling and Eleftherios Ioannidis (VUA), Arjo Segers (TNO), Andrea Kaiser-Weiss, Anusha Sunkisala, Valentin Bruch and Thomas Rösch (DWD), Michał Gałkowski and Piotr Sekuła (AGH), Guillaume Monteil and Hans Chen (LUND), Paul Palmer and Tia Scarpelli (UEDIN), Anteneh Getachew Mengistu and Tuula Aalto (FMI)	15/12/2023	Compilation and harmonization of reports from all participating groups.

Internal Review History

Internal Reviewers	Date	Comments
Name (Organisation)	dd/mm/yyyy	

Estimated Effort Contribution per Partner

Partner	Effort
Organisation	effort in person month
Total	**0**