CoCO2

Prototype system for a Copernicus CO_2 service

Ensemble of estimates for assimilation into prototype

Julia Marshall, Antoine Berchet, Nicolas Bousserez, Gregoire Broquet, Dominik Brunner, Hans Chen, Liang Feng, Janne Hakkarainen, Guillaume Monteil, Paul Palmer, and Elise Potier

CoCO2

Prototype system for a Copernicus CO_2 service

D4.7 ENSEMBLE OF ESTIMATES FOR ASSIMILATION INTO PROTOTYPE

Dissemination Level:	Public
Author(s):	Julia Marshall (DLR), Antoine Berchet (CEA), Nicolas Bousserez (ECMWF), Gregoire Broquet (CEA), Dominik Brunner (EMPA), Hans Chen (LUND / now at CHALMERS), Liang Feng (UEDIN), Janne Hakkarainen (FMI), Guillaume Monteil (LUND), Paul Palmer (UEDIN), Elise Potier (CEA / now at LISA)
Date:	24/10/2023
Version:	1.1
Contractual Delivery Date:	30/06/2023
Work Package/ Task:	WP4/ T4.2 and T4.4
Document Owner:	DLR
Contributors:	ECMWF, CEA, EMPA, LUND, FMI
Status:	Final

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 CO2 emissions

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

Published by the CoCO2 Consortium

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

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

Table of Contents

Figures

Tables

1 Executive Summary

This report summarizes the inputs provided from the local and regional inverse modelling systems of WP4 to the global multi-model, multi-scale inversion approach of the prototype system being developed in WP6. The deliverable itself consists of these data sets, this report serves to document these inputs.

2 Introduction

2.1 Background

The goal of the multi-model, multi-scale data assimilation approach that has been developed within the prototype system of the IFS is to integrate the high-resolution transport and inversion information from local and regional inversion systems into the global IFS inversion system. Here a two-way flow of information is foreseen: while the ensemble information from the local and regional systems provides additional information about the transport, inversion and corresponding errors from the higher resolution runs, the prior fields from the global atmospheric inversion system can be used as boundary conditions for the local and regional inversion systems.

From the perspective of the global modelling system, ensemble statistics from the local and regional scale atmospheric inversions are used to extract the information required to assimilate these products as observations in the IFS 4D-Var system. This is summarized in Figure 1.

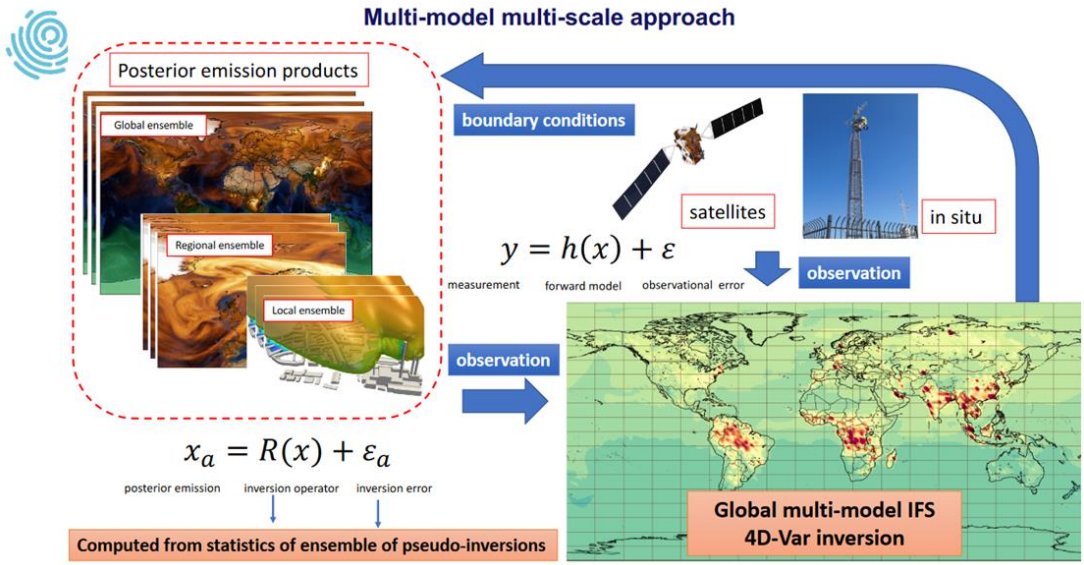

Figure 1: Conceptual diagram describing the multi-model, multi-scale approach.

The system was conceived such that it should be both computationally efficient and non-intrusive from the perspective of the global modelling system. It was foreseen that the required input from the regional modelling systems would be similar for different regional modelling systems, and should be easily generated, which proved not to be the case for some inversion setups.

Input for the global system was considered on two scales, namely point-source emissions from plume-based approaches (such as those developed in Task 4.2 of CoCO2), and regional-

or national-scale inversions (such as those in Task 4.4 of the project). The plume-based approaches are operating at a spatial resolution much higher than can be realized with the global model, and can provide finer scale information about the emission distribution within a global model gridcell. The regional- and national-scale inversions are more comparable in terms of the resultant fluxes, and may be able to better represent different aspects of the uncertainty, such as including country-specific prior fluxes and better representing *in situ* measurements. (Note that the input for the global system in WP6 is the output from the regional- and national-scale modelling systems of WP4).

The required output from models performing point-source emission estimates (as described in Section 3.1) and for spatially-resolved regional fluxes differed slightly, and is summarized as follows:

1) For point-source estimates (plume inversions), which do not rely on Bayesian inversion approaches:
   a) Posterior emission value
   b) Location (longitude, latitude, time) of the emission
   c) Posterior uncertainty (error as a standard deviation)
   d) Location (longitude, latitude, time) of satellite observation used in the inversion (typically the area and neighbourhood of the plume image or transect in the satellite data)

2) For spatially-resolved estimates based on Bayesian inversion approaches:
   a) Gridded prior emissions (x_b)
   b) (Satellite) observation values (vector y) and locations
   c) Ensemble of perturbed gridded prior emissions ($x_{bi}=xb+d_{xi}$, for i=1,...,N)
   d) Corresponding ensemble of perturbed observations ($y_i=y_o+\varepsilon_i$, for i=1,...,N)
   e) Ensemble of gridded posterior fluxes (x_{ai}, for i=1,...,N) based on both perturbed observations and perturbed fluxes.

The desired output proved difficult for some of the models to provide, as most regional ensemble-Kalman-filter-based inversions do not typically perturb the observations, instead using only a perturbation of the prior fluxes. While perturbing the observations is itself a trivial exercise, and for the first optimization step only requires the filtering step to be repeated, the results of this filtering is propagated forward into the next lag step, and influences the creation of the following ensemble. Thus, not only the optimization step would need to be repeated, but also all subsequent transport simulations for each ensemble member. This proved restrictively expensive for modelling groups using e.g.CT-DAS-based inversion systems (such as EMPA's ICON-CT-DAS implementation, or DLR's WRF-CT-DAS) to carry out.

However, while discussing the input with different modelling groups, it became clear that deterministic square-root Kalman ensemble systems that do not perturb the observations could still produce the required error matrix that is needed in order to assimilate posterior fluxes in the IFS system. That is, perturbing the observations may not be strictly necessary in these cases. This approach still needs to be tested, which is not feasible within the time constraints of this deliverable. However, further tests with ensemble-based regional inversions are foreseen in the future.

To ensure the ease of use of this information within the global inversion system, a template input file was provided, such that the output could be interpolated onto a common grid (pixels in the longitude dimension = 1616, pixels in the latitude dimension = 800).

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverable

This deliverable consists primarily of output from various local and regional (national or continental) scale inverse modelling systems from WP4 to test their assimilation into the multi-

model, multi-scale approach being developed in the 4D-Var inversion system of the IFS. This report serves to document these systems and describe the simulations that produced these data sets.

2.2.2 Work performed in this deliverable

Several meetings were held between participants in WP4 and WP6 to clarify the requirements of the global prototype data assimilation system. Additional simulations were carried out by the modelling systems in WP4 who were able to perturb their observations in order to provide the required output. This work was carried out by consortium partner LUND with the LUMIA system, by CEA with the ORCHIDEE-CIF system, and by UEDIN with the GEOS-Chem system. This output was post-processed to put it into a format that could more easily be used by the IFS, and it was provided to WP6. Additionally, plume-based emissions were provided and documented by FMI. The different modelling systems and their output are documented in this deliverable.

2.2.3 Deviations and countermeasures

This deliverable has been delayed. It was originally due in month 18 (June, 2023) and was postponed to month 21 (September, 2023). Part of this delay is related to the development work going on in the contributing modelling systems, not all of which were able to provide output earlier in the project. There were also several iterations with the global modelling team in WP6 to clarify exactly what was needed as input. The specific requirements have evolved in time due to the specific developments of the systems in WP4 and WP6. Few of the regional modelling inversion systems had previously used perturbed observations as part of their inversion framework, and large ensemble experiments are computationally expensive. Providing input for this deliverable required additional development work and simulations from the contributors from WP4. This additional work was done on a best-effort basis by the participating groups in WP4, and provides a first test data set for the multi-model, multi-scale data assimilation approach. Finally, the consortium member who was originally supposed to be responsible for the preparation of this deliverable was unable to do so, and another project partner had to step in to draft the report and coordinate input, which contributed to the delay.

3 Contributions to ensembles of estimates

The sections below briefly describe the modelling systems that have contributed to this deliverable. The datasets themselves are available on the FTP site of the CoCO2 project, and can be accessed as follows:

ftp coco2@ftp.ecmwf.int

password: Can be found on the Confluence site, available upon request to the project coordinator.

cd data-exchange/WP4/WP4_WP6_exchange

3.1 Contribution from a plume-based inversion

3.1.1 Description of modelling system

As a contribution from the FMI plume-based inversion system, we consider two sets of CO_2 emission estimates for the Matimba power station in South Africa based on the results presented by Hakkarainen et al. (2021). The first estimate is derived from NASA's Orbiting Carbon Observatory-2 (OCO-2) CO_2 data (V10r) using the cross-sectional flux (CSF) emission estimation method. The second is derived by multiplying the satellite-derived source-specific NO_x-to-CO_2 emission ratio (calculated from the OCO-2 and TROPOMI data by Hakkarainen et al., 2021) to the monthly NO_x emissions obtained from the Sentinel-5P/TROPOMI NO_2 data

(V1.2) using exponentially modified Gaussian (EMG) emission estimation method. Details of the methodology and the data versions used can be found in Hakkarainen et al. (2021).

We also use emission estimates of six power stations (Kendal, Kriel, Matla, Majuba, Tutuka and Grootvlei) and the largest single emitter of greenhouse gas in the world, the Secunda CTL synthetic fuel plant in the South African Highveld region as derived by Hakkarainen et al. (2023). The emissions are estimated using the CSF method from the NASA's OCO-3 Snapshot Area Map (SAM) CO_2 observations (V10.4r) that cover a target area of 80 km by 80 km in 2 min.

Finally, we consider the emissions from the Bełchatów Power Station in Poland obtained using a Gaussian Plume (GP) model method as described by Nassar et al. (2022). The emission estimates are based on OCO-3 SAM data (V10r).

In all cases, we share estimates of the standard deviation of the uncertainties in the individual emissions estimates to support the protocol of assimilation of these estimates into the global IFS inversion system.

3.1.2 Output from simulation

Table 1 summarizes the point source emission estimate datasets from the plume-based inversion. In all cases, we recommend excluding the respective data (OCO-2, OCO-3, TROPOMI) within a ±2° latitude-longitude box with respect to source location from the global IFS inversions when co-assimilating these point source emission estimates. The point source locations and the 1-sigma uncertainties in the emission estimates were provided in an accompanying Excel spreadsheet. The spreadsheet has been uploaded to the ftp server under **`data-exchange/WP4/WP4_WP6_exchange/plume_inversions.`**

Table 1: An overview of the flux estimation ensembles from the plume-based approach that were supplied to WP6.

Sources	Method	Satellite data	Number of emission estimates	Reference
Matimba	CSF	OCO-2 (V10r), TROPOMI (V1.2)	14	Hakkarainen et al., 2021
Matimba	EMG, NO_x-to-CO_2 ratio	TROPOMI (V1.2), OCO-2 (V10r)	29	Hakkarainen et al., 2021
Kendal, Kriel, Matla, Majuba, Tutuka, Grootvlei and Secuda CTL	CSF	OCO-3 SAM (V10.4r)	11	Hakkarainen et al., 2023
Bełchatów	GP	OCO-3 SAM (V10)	9	Nassar et al., 2022

3.2 Contribution from the Lagrangian-based inversion system LUMIA

3.2.1 Description of modelling system

Monte-Carlo posterior emission estimation were provided from the LUMIA CH_4 inversions that were carried out within WP5.5. The transport was done with the Lagrangian model FLEXPART10.4, which was run on a regional grid (15°W, 33°N to 35°E, 73°N), at 0.25° resolution for the full year of 2018. The FLEXPART footprints were computed using hourly ERA5 meteorological fields.The mixing ratios for the boundary conditions were based on the CAMS v19r1 optimized methane fluxes, which were run forward in the TM5 global tracer transport model with the regional fluxes set to zero to provide a time series of background CH_4 concentrations at the observation sites. The prior fluxes within the regional domain were taken from the CoCO2 WP5.5 CH_4 inversion intercomparison protocol.

The prior uncertainty matrix was constructed by setting the variances proportional to the (absolute) prior emissions with exponentially decaying spatial and temporal correlations of length 50 km and 30 days, respectively. The total uncertainty was set to 5 $TgCH_4$/year. The observations consisted of the 45 sites which are listed in the CH_4 inversion intercomparison protocol of WP5.5. Only afternoon observations were used with the exception of high altitude sites, for which nighttime observations were used. Uncertainties on the measurements were defined as the sum of the measurement uncertainties (when provided) and a constant, site-specific estimation of model uncertainty, based on the quality of the fit of the short-term variability of the model to the short-term variability of the observations.

3.2.2 Output from simulation

An ensemble of extra inversions was performed. For each ensemble member (currently 30), the prior and observation vectors were replaced by a random realization of the prior and observation covariance matrices. The output was stored in five output files:

- **`observations.nc`**, containing:
  - the observation coordinates (**`time`**, **`lat`**, **`lon`**, **`alt`**, **`height`**)
  - the (unperturbed) observations for 45 sites within the domain (**`obs`**)
  - the observation uncertainty (**`err`**)
  - the prescribed boundary condition (**`mix_background`**)
  - the prior and posterior mixing ratios in the main inversion (**`mix_apri`** and **`mix_apos`**)
  - the perturbed observation, prior and posterior mixing ratio in each perturbed experiment (**`obs_pert0`**, **`mix_apri_pert0`**, **`obs_pert1`**, **`mix_apri_pert1`**, etc.)
- **`emissions.nc`**, containing the prior and posterior emissions
- **`perturbed.nc`**, containing the prior and posterior fluxes for the inversions with perturbed priors and observations
- **`emissions.regridded.nc`**, as emissions.nc, but regridded onto the IFS grid
- **`perturbed.regridded.nc`**, as perturbed.nc, but regridded onto the IFS grid

Furthermore, a readme file **`readme.html`** with some sample plots was provided. The output from these simulations have been uploaded to the project ftp server described above into a folder called **`data-exchange/WP4/WP4_WP6_exchange/LUMIA`**. A plot of the perturbed prior and resultant posterior simulations is shown in Figure 2.

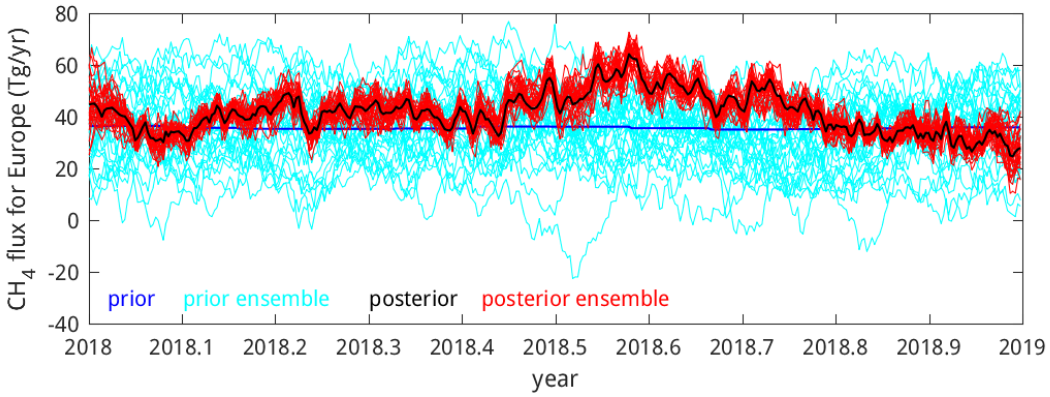

Figure 2: The prior flux is shown in blue for the full domain over 2018, while the 30 ensemble members with perturbed priors are shown in cyan. The posterior fluxes are plotted in black, with the posterior ensemble members shown in red. The fluxes are daily, but shown in TgCH$_4$/yr.

3.3 Contribution from the CIF-CHIMERE inversion system

Contribution from the CIF-CHIMERE inversion system

3.3.1 Description of modelling system

The CIF-CHIMERE inversion configuration for France developed in the framework of task T4.4 (WP4) of the CoCO2 project is used to provide a Monte Carlo national scale estimate of hourly CO_2 flux maps over a full month. 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). The CHIMERE configuration for France covers the domain: 11°W-12°E ; 39,5°N-54,5°N (cf Figure 3). Its zoomed grid has a 10 km horizontal resolution over France (Figure 3). CHIMERE is driven by the ECMWF / IFS operational meteorological forecasts. The CIF-CHIMERE inversions of the CO_2 fluxes in France separately control 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 1-day temporal resolution
- the ocean and terrestrial ecosystem fluxes are controlled at 10-km and 6-hourly resolution

In the frame of T4.4 in WP4, various experiments have been conducted with this CIF-CHIMERE configuration using different products for the prior estimates of the control vector, assimilating different observation datasets, and varying some of the system parameters. A description of the CIF-CHIMERE configuration parameters, of these datasets and of the results will be provided in deliverable 4.6.

Here, a specific set-up of the CIF-CHIMERE configuration is used to generate Monte Carlo ensembles of prior and posterior estimates of the CO_2 fluxes in France in July 2018 with a corresponding ensemble of observation vectors.

This set-up relies on the prior estimate of the anthropogenic emissions and terrestrial ecosystem fluxes from the TNO inventory and Vegetation Photosynthesis Respiration Model (VPRM) simulations delivered in the frame of WP2 (Denier van Der Gon et al., 2022). The prior estimate of the initial and boundary conditions is derived from the CAMS global CO_2 inversions v20r2 (assimilating surface data). The corresponding CIF-CHIMERE inversion

assimilates both 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/), and satellite total column CO_2 (XCO_2) observations from the OCO-2 NASA-JPL mission (the v11 dataset, see Figure 3). For both the in situ and satellite observations the inversion systems 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 observations respectively taken from Broquet et al. (2013) and Potier et al. (2022). The set-up of 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 much higher spatial resolution here) and some other slight differences. The set-up of the part corresponding to the CO_2 anthropogenic emissions in the prior uncertainty covariance matrix assumes a 50% 1-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. This matrix does not include any correlation between the different main components of the control vector (ocean, terrestrial ecosystem and anthropogenic fluxes, and boundary conditions).

A specific driver of the CIF allows for a direct derivation of ensembles reflecting the uncertainties in the prior estimate of the control vector, the model and observation errors and the uncertainties in the posterior estimate of the control vector. They are respectively characterised by an ensemble of perturbed prior estimates of the control vector, an ensemble of perturbed observations and an ensemble of corresponding posterior estimates of the control vector. Therefore, these ensembles fulfil the requirements for the assimilation into the global IFS inversion system.

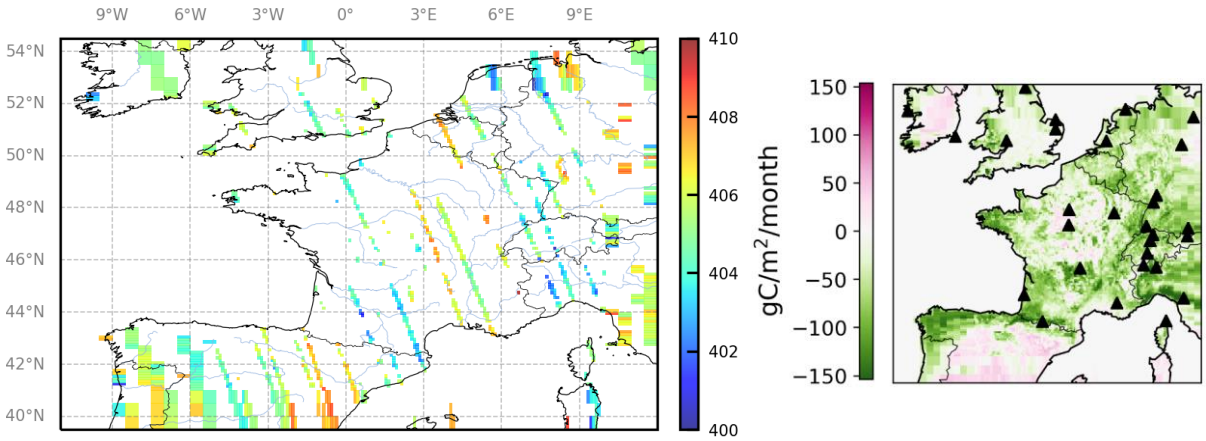

Figure 3: Domain of the CIF-CHIMERE inversions showing for the month of July 2018 (left): binned OCO-2 v11 observations, and (right): prior biogenic CO_2 fluxes from VPRM interpolated onto CHIMERE zoomed grid and ground-based measurement stations used for the inversions.

3.3.2 Output from simulation

An ensemble of 50 perturbed members has been generated at this stage. The flux estimates from these members, as well as the corresponding unperturbed prior and posterior inversion estimates have been reprojected at 1-hour temporal resolution (using the products corresponding to the prior estimates of the fluxes) and regridded on the template of the IFS

grid provided by ECMWF in NetCDF format. They have been uploaded to the project ftp under the folder **data-exchange/WP4/WP4_WP6_exchange/CHIMERE.**

The delivered files from the ensemble of 51 perturbed and unperturbed inversions consist of:

- 51 prior emission files **CHIMERE_prior_<n>_sfc_em.grib**
- 51 posterior emission files **CHIMERE_posterior_<n>_sfc_em.grib**
- 51 directories, each one with two files containing the corresponding CO_2 observations from, respectively, the surface network (**obsvect_<n>/concs/CO2/monitor.nc**) and OCO-2 (**obsvect_<n>/satellites/CO2/monitor.nc**)

where the perturbed inversions are numbered n=1 to 50 (zero-padded to four digits) and where the unperturbed inversion is numbered n=0. The folder **data-exchange/WP4/WP4_WP6_exchange/CHIMERE** also contains three archive (**tar**) files with the prior emission files, the posterior emission files and the observation directories: respectively **prior.tar**, **posterior.tar** and **obsvect.tar**.

3.4 Contribution from UEDIN inversion system

3.4.1 Description of modelling system

For our forward model, 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 to 35 ° E longitude and 34 to 66 ° N latitude) with 47 vertical levels. GEOS-Chem is driven by meteorological reanalysis 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. Sources for CO_2 include 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}). We linearize our CO simulation by using pre-calculated 3-D loss fields of OH generated by a full-chemistry version of the model

For our 2018-2021 *a priori* fluxes, we use a combination of regional and global inventories. 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.

For our nested domain, we use boundary conditions for CO_2 (CO_2^{Trans}) from the CAMS inversion-optimised global greenhouse gas analysis with assimilation of *in situ* observations (version v19r2, 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.
For our inversion we use the Ensemble Kalman Filter (EnKF) approach as discussed in detail by others (e.g., 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 for the mean a posteriori state vector that represents the mean of our 100 ensemble members. 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.

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 covariance 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. 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.

3.4.2 Output from simulation

Output for 2018 has been made available for the joint CO-CO_2 inversion based on surface measurements only, and is available on the project ftp under

`Data-exchange/WP4/WP4_WP6_exchange/…`

`GEOSChem/UoE_Tia_CO2_CO_insitu_inv.nc`.

It contains the prior and posterior ensembles of the CO_2 biogenic and anthropogenic fluxes separately, as well as the CO anthropogenic prior and posterior flux ensembles. The perturbed observation ensembles are included as well.

In addition, the same model output for an inversion using both surface and satellite measurements is available in the same folder, with the file name

`UoE_Tia_CO2_CO_satellite_inv.nc.gz`.

4 Conclusion

This deliverable documents the datasets that were provided from the local- and regional-scale inversions of WP4 as input to the multi-scale, multi-model inversion system foreseen within WP6. The inclusion of not only perturbed prior fluxes but also perturbed observations proved to be computationally prohibitive for some of the Ensemble Kalman-Filter approaches to implement, but the use of output from deterministic square-root ensemble Kalman filter approaches will be tested in the future. Within the scope of this deliverable, example datasets were provided on a variety of temporal and spatial scales, as described in Section 3. This

provides sample datasets in order to test the implementation of the multi-scale, multi-model approach within WP6.

5 References

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.

Bukosa, Beata, Jenny A. Fisher, Nicholas M. Deutscher, and Dylan B. A. Jones. 2023. "A Coupled CH4, CO and CO2 Simulation for Improved Chemical Source Modeling" Atmosphere 14, no. 5: 764. https://doi.org/10.3390/atmos14050764

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.

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

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

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

Feng, L., P. I. Palmer, H. Bösch, and S. Dance (2009), Estimating surface CO2 fluxes from space-borne CO2 dry air mole fraction observations using an ensemble Kalman filter, Atmos. Chem. Phys., 9, 2619–2633, doi:10.5194/acp-9-2619-2009.

Feng, L., Palmer, P. I., Bösch, H., Parker, R. J., Webb, A. J., Correia, C. S. C., Deutscher, N. M., Domingues, L. G., Feist, D. G., Gatti, L. V., Gloor, E., Hase, F., Kivi, R., Liu, Y., Miller, J. B., Morino, I., Sussmann, R., Strong, K., Uchino, O., Wang, J., and Zahn, A.: Consistent regional fluxes of CH4 and CO2 inferred from GOSAT proxy XCH4 : XCO2 retrievals, 2010–2014, Atmos. Chem. Phys., 17, 4781–4797, https://doi.org/10.5194/acp-17-4781-2017, 2017.

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.

Gerbig, C.: Parameters for the Vegetation Photosynthesis and Respiration Model VPRM, https://doi.org/10.18160/R9X0-BW7T, 2021.

Hakkarainen, J., Szeląg, M. E., Ialongo, I., Retscher, C., Oda, T. and Crisp, D.: Analyzing nitrogen oxides to carbon dioxide emission ratios from space: A case study of Matimba Power Station in South Africa, Atmospheric Environment: X, Volume 10, doi:10.1016/j.aeaoa.2021.100110, 2021.

Hakkarainen, J., Ialongo, I., Oda, T., Szeląg, M. E., O'Dell, C. W., Eldering, A., and Crisp, D.: Building a bridge: Characterizing major anthropogenic point sources in the South African Highveld region using OCO-3 carbon dioxide Snapshot Area Maps and Sentinel-5P/TROPOMI nitrogen dioxide columns, Environmental Research Letters, vol. 18, no. 3, doi:10.1088/1748-9326/acb837, 2023.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh (2007), Efficient data assimilation for spatiotemporal chaos: A Local Ensemble Transform Kalman Filter, Phys. D, 230, 112–126.

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

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier van der Gon, H.: CAMS-REG-v4: a state-of-the-art high resolution European emission inventory for air quality modelling, Earth System Science Data, 14, 491–515, https://doi.org/10.5194/essd14-491-2022, 2022.

Lefèvre, N., Tyaquiçã, P., Veleda, D., Perruche, C., & Van Gennip, S. J.: Amazon River propagation evidenced by a CO2 decrease at 8 N, 38 W in September 2013. Journal of Marine Systems, 211, 103419, https://doi.org/10.1016/j.jmarsys.2020.103419, 2020.

Liu, J., Bowman, K. W., and Lee, M. (2016), Comparison between the Local Ensemble Transform Kalman Filter (LETKF) and 4D-Var in atmospheric CO2 flux inversion with the Goddard Earth Observing System-Chem model and the observation impact diagnostics from the LETKF, J. Geophys. Res. Atmos., 121, 13,066–13,087, doi:10.1002/2016JD025100.

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.

Nassar R, Moeini O, Mastrogiacomo J-P, O'Dell CW, Nelson RR, Kiel M, Chatterjee A, Eldering A and Crisp D (2022) Tracking CO_2 emission reductions from space: A case study at Europe's largest fossil fuel power plant. Front. Remote Sens. 3:1028240. doi: 10.3389/frsen.2022.1028240

OCO-2/OCO-3 Science Team, Vivienne Payne, Abhishek Chatterjee (2022), OCO-2 Level 2 geolocated XCO2 retrieval results and algorithm diagnostic information V11, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: September 1, 2023, https://disc.gsfc.nasa.gov/datacollection/OCO2_L2_Diagnostic_11.html

Potier, E., Broquet, G., Wang, Y., Santaren, D., Berchet, A., Pison, I., Marshall, J., Ciais, P., Bréon, F.-M., and Chevallier, F.: Complementing XCO_2 imagery with ground-based CO_2 and $^{14}CO_2$ measurements to monitor CO_2 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.

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

White, E.D., Rigby, M., Lunt, M.F., Smallman, T.L., Comyn-Platt, E., Manning, A.J., Ganesan, A.L., Simon O'Doherty, S., Stavert, A.R., Stanley, K., Williams, M., Levy, P., Ramonet, M., , Forster, G.L., Manning, A.C., and Palmer, P.I.: Quantifying the UK's carbon dioxide flux: an atmospheric inverse modelling approach using a regional measurement network, Atmos. Chem. Phys., 19, 4345-4365, https://doi.org/10.5194/acp-19-4345-2019, 2019.

Document History

Version	Author(s)	Date	Changes
0.1	Julia Marshall (DLR)	14/09/2023	First draft
1.0	Julia Marshall (DLR)	06/10/2023	Updated in response to reviewer comments, edited to include the contribution from UEDIN.
1.1	Julia Marshall (DLR)	24/10/2023	Replotted Figure 2 to provide units, and converted the fluxes into a more commonly-used unit. Details regarding the input from CEA were also included.

Internal Review History

Internal Reviewers	Date	Comments
Wouter Peters (WUR)	28/09/2023	Comments were provided in the document. The data version used for e.g. satellite input and lateral boundary conditions should be clarified in several places. Some text had to be clarified, regarding input/output and the interface between the global and regional modelling systems.
Sander Houweling (VUA)	2/10/2023	A few minor comments were made in the text to clarify some points that were not clear enough for a set of independent eyes.

Estimated Effort Contribution per Partner

Partner	Effort
CEA	1
DLR	1
ECMWF	0.5
FMI	0.5
LUND	1
UEDIN	0.5
Total	**4.5**