CO2 Human Emissions

D1.2: Set of global data assimilation input and output products

Liesbeth Florentie

che-project.eu

Co-ordinated by ECMWF

CO2 Human Emissions

D1.2 Set of global data assimilation input and output products

Dissemination Level: Public

Author(s): Liesbeth Florentie, Wouter Peters, Maarten Krol (WUR), Frédéric Chevallier (CEA), Maximilian Reuter, Michael Buchwitz (UB), Alex Webb, Hartmut Boesch (ULEIC)

Date: 20/12/2019
Version: 1.0
Contractual Delivery Date: 31/12/2019
Work Package/ Task: WP1/ T1.2
Document Owner: WUR
Contributors: CEA, UB, ULEIC
Status: Final

CHE: CO2 Human Emissions Project

Coordination and Support Action (CSA)
H2020-EO-3-2017 Preparation for a European capacity to monitor CO2 anthropogenic emissions

Project Coordinator: Dr Gianpaolo Balsamo (ECMWF)
Project Start Date: 01/10/2017
Project Duration: 39 months

Published by the CHE Consortium

Contact:
ECMWF, Shinfield Park, Reading, RG2 9AX,
gianpaolo.balsamo@ecmwf.int

The CHE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776186.

Table of Contents

Figures

Tables

Glossary

ATBD	Algorithm Theoretical Basis Document
CAMS	Copernicus Atmospheric Monitoring Service
COS	Carbonyl Sulfide
CTDAS	CarbonTracker Data Assimilation Shell
CTE	CarbonTracker Europe
CTSF	CarbonTracker Statistical Fit
DA	Data Assimilation
GPP	Gross Primary Production
MVS	Monitoring and Verification System
NEE	Net Ecosystem Exchange
NIR_v	Near Infrared Reflectance of terrestrial vegetation
PSD	Product Specification Document
PUG	Product User Guide
QA/QC	Quality Assurance/Quality Control
TCCON	Total Carbon Column Observing Network

1 Executive Summary

To address challenges towards building an operational Monitoring and Verification System (MVS) for CO_2, we report here on (a) the creation of new sets of input data for data assimilation (DA) systems, and (b) their use in a large number of alternative or new DA configurations. Together, these sets provide new insights on the data processing for XCO_2, their assimilation, and the best method to maximize data usage, minimize biases, separate biogenic and anthropogenic fluxes, while maintaining a fast operating chain. This report documents the technical part of our efforts, and the data availability of a first set of input and output products. More and updated results will be created during Task 1.3 when a benchmarking and QA/QC pipeline will be demonstrated for our DA systems.

We find that a new fast and efficient XCO_2 production algorithm (FOCAL, UB) can create a suitable alternative to full-physics retrievals for OCO-2, and that propagation of error covariances into an XCO_2 retrieval seems feasible (ULEIC & CEA), but that the processing chain is currently not suited for near real-time. Data assimilation of SIF is feasible, but a proper observation operator is missing because of incomplete knowledge of SIF emissions from the canopy, and because the pre-TropOMI SIF products are too coarse to separate ecosystems. COS as a biogenic tracer requires further work on ocean and soil source terms, specifically in the tropical regions. XCO_2-based DA systems can now match in-situ based flux estimates on global to regional scales and suggest a smaller tropical sink and seasonal cycle. They are ready for wider application. Atmospheric transport models need to be upgraded in resolution to match the satellite products, which can be achieved by either (a) moving towards online calculation of meteorology and mass-fluxes, and (b) shortening the DA windows towards the daily-to-weekly scale. These system improvements are scheduled for the second part of task 1.2.

2 Introduction

2.1 Background

Within the CHE project WP1, we aim to test and implement a number of innovations in data assimilation that are needed towards an operational MVS for CO_2. This includes the delivery chain of large amounts of input data, such as from remote sensing instruments and in-situ sensors, of CO_2 and its auxiliary tracers such as COS. These need to feed data assimilation systems equipped to not only handle these data streams, but also to combine them including their biases and formal uncertainties. Finally, an important component is the quality control on the posterior results: how do we quickly assess whether the solution is correct from a statistical, but also from a physical and biogeochemical viewpoint? As this is such an important issue a separate task is devoted to this in WP1. Quality assessment and control of the posterior results will be elaborated upon in detail during the benchmarking phase, in task 1.3.

In task 1.2 we have set up a number of DA configurations of the CAMS (CEA) and CTE (WU) systems, aiming to bring several innovations to the processing chain. To facilitate this new satellite XCO_2 products were created that either (a) follow a very fast processing chain suitable for near real-time delivery (UB), or (b) include auxiliary covariance information to aid their consistent assimilation in CAMS (ULEIC). These XCO_2 products and their availability are briefly documented here. They form the input of the DA chain, along with a number of community-available and standardized inputs that were put in place for a DA protocol that both

CAMS and CTE followed. This protocol is separately documented as Deliverable D1.1 (Sept. 2018).

The DA innovation runs were designed to address four main challenges:

1. Optimally integrate satellite retrieval bias-correction into the data assimilation processing chain,
2. Minimize the impact of transport model errors on surface flux estimates,
3. Improve the model-data fusion techniques for joint natural and anthropogenic CO2 emission estimates,
4. Deploy more advanced surface flux descriptions of natural and anthropogenic CO2 emissions in the data assimilation chain.

The documentation below describes our joint efforts towards these challenges. The report describes the simulations and data availability, but not yet the main outcomes. These are reported elsewhere in literature, but partly also need to be created in the follow-up step in Task 1.3. There, these different DA outcomes will be compared and benchmarked in a newly developed QA/QC chain, designed to uncover the strengths and weaknesses of each approach.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

The objective of this deliverable is to present a set of global satellite-derived XCO2 input products, as well as a first set of carbon flux estimates obtained with our innovative DA systems. These products are prepared according to a community protocol (D1.1) and available as open-access to the scientific community.

Note that these are not yet the final products, as they will be further developed during the project. improvements and updates to these products will be made available during the benchmarking phase in Task 1.3.

2.2.2 Work performed in this deliverable

- A new algorithm for a satellite-based XCO_2 retrieval based on OCO-2 data, FOCAL-OCO-2, has been developed. The resulting set of input data contains observations for the years 2015 to 2018.
- Research has been performed towards the creation of a XCO_2 product with error covariance structures tailored to and consistent with the CAMS atmospheric inversion system.
- The capability to assimilate satellite-based XCO_2 observations has been added to CarbonTracker Europe. Flux inversions have been performed using both the NASA OCO-2 and the FOCAL-OCO-2 retrievals.
- CAMS flux inversions based on the assimilation of NASA OCO-2 XCO_2 observations have been performed.
- Research has been performed towards joint assimilation of surface and space-based observations. A new set-up to perform flux inversions with CAMS using such combined observational constraints has been developed.
- A new long-window inversion system based on the optimization of parameters in a statistical biosphere flux description enhanced with proxy anomalies has been developed. The system has been tested by performing flux inversions using both SIF and NIR_V as proxy data.
- Research has been performed towards the development of a new flux inversion system that incorporates COS data.

2.2.3 Deviations and counter measures

ULEIC

The original plan to improve the quality of satellite XCO_2 retrievals by adopting machine learning methods was abandoned in favour of the development of a modified retrieval that features a tailored covariance structure for consistent assimilation of the XCO_2 data by the CAMS atmospheric inversion system. This is motivated by (i) the need to address inconsistent statistical hypotheses between XCO_2 retrieval and flux inversion schemes, which may be a significant source of error, and (ii) the improvement in quality filter and bias correction present in the NASA OCO-2 XCO_2 retrieval from the v8 release on.

WU

The development of a DA system that features high-resolution online IFS-TM5 transport is ongoing but has not yet reached the point where we can present results. This system will be further developed early 2020.

CEA

The protocol agreed upon in September 2018 (D1.1) has not been closely followed yet for practical reasons linked to the research stage of this work and to the large computational burden required for each test: each inversion needed at least 2-3 weeks of computations on a parallel cluster. Innovations to the CAMS inversion set-up have been tested though and will be repeated according to the protocol in the following months.

Work on a flux inversion system with COS is ongoing but has not yet reached the level of maturity where results can be shared.

3 Description of developed input products

Development of the below described input products has been performed within the framework of the CHE project. Apart from these products, links to external input products for carbon flux inversions have been shared on the data portal on the CHE website (https://www.che-project.eu/data-portal).

3.1 FOCAL-OCO-2 XCO_2

3.1.1 Short description of the FOCAL algorithm

(From the FOCAL website www.iup.uni-bremen.de/~mreuter/focal.php)

The fast atmospheric trace gas retrieval for OCO-2 (FOCAL-OCO-2) has been setup to retrieve XCO2 (the column-average dry-air mole fraction of atmospheric CO2) by analyzing hyper spectral solar backscattered radiance measurements of NASA's OCO2 satellite. FOCAL includes a radiative transfer model which has been developed to approximate light scattering effects by multiple scattering at an optically thin scattering layer. This reduces the computational costs by several orders of magnitude. FOCAL's radiative transfer model is utilized to simulate the radiance in all three OCO-2 spectral bands allowing the simultaneous retrieval of CO_2, H_2O, and solar induced chlorophyll fluorescence. More information can be found in the FOCAL ATBD (Reuter et al., 2019).

3.1.2 Short description of the data product

The FOCAL-OCO-2 XCO_2 data product is a level 2 product and contains detailed information on XCO_2 and its uncertainty, time, latitude, longitude, averaging kernel, a priori profile, etc. for each individual satellite footprint fulfilling the pre- and post-processing quality requirements. Additionally, the data product includes information on XH_2O and SIF, which are by-products of the XCO_2 retrieval. The data format is conform with the satellite-derived Level 2 XCO_2 data products generated in the framework of ESA's Climate Change Initiative (http://www.esa-ghg-cci.org) (e.g., Buchwitz et al., 2017) and, more recently, for the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu) (e.g., Buchwitz et al., 2018). The data format is netCDF-4 (classic) and the product is in-line with CF (Climate and Forecasting) convention 1.6. Each parameter is explained in each file and the product is, therefore, essentially self-explaining. For each day one separate file has been generated. The data format is described in detail by Buchwitz et al. (2014). More information on the data product and its format can be found in the FOCAL PUG (Reuter and Buchwitz, 2019) and PSD (Buchwitz et al., 2014).

3.1.3 Data availability

Four years (2015-2018) of global OCO-2 XCO_2 data have been generated with the FOCAL-OCO-2 XCO_2 level 2 algorithm. The data is made available via the FOCAL website: http://www.iup.uni-bremen.de/~mreuter/focal.php.

3.2 XCO_2 retrieval with covariance structures consistent to the CAMS atmospheric inversion system

3.2.1 Description of the modified XCO_2 product

Current atmospheric inversion schemes assume rather realistic a priori information, while satellite retrieval schemes which try to maximise the measurement contribution in the retrievals give a very weak weight to prior information. Such inconsistent statistical hypotheses between XCO_2 retrieval and atmospheric inversion schemes may be a significant cause of error in atmospheric inversions assimilating satellite data. By using sounding-specific covariances from model covariances representing the prior flux uncertainties of the CAMS model, we can derive satellite CO_2 columns tailored towards and consistent with the specific assumptions of the CAMS atmospheric inversion scheme.

We have applied a framework for swapping the a priori error covariance matrices used in the operational OCO-2 XCO_2 retrieval data with those taken from the CAMS CO_2 inversion model v14.2. This framework has been applied to OCO-2 L2 Diagnostic 8r data for the year 2016, recalculating the CO_2 profiles, averaging kernels, and recalculating the total columns using the new prior information. In addition, a second method has been used where the CAMS a priori covariance matrices have been scaled such that that total column variability agrees with those from operational OCO-2 retrieval. The purpose of this is to remove any offset between the two prior covariances, leaving all remaining differences solely due to the profile structure.

In total, approximately 36 million OCO-2 soundings have been recalculated using both the 'scaled' and 'non-scaled' methods with a small number of soundings failing during the matrix inversion process.

Due to the use of new a priori covariance, the applied OCO-2 bias correction of the operational processing will not be applicable to our new XCO_2 results and a custom bias correction has to be applied following the same method as the JPL OCO-2 team:

- Firstly, we calculate a new footprint bias by finding all scenes where all 8 of the OCO-2 across-track scenes have been successfully retrieved and calculating the bias across these different footprints.
- Secondly, we consider the correlation between variables, which are retrieved simultaneously with XCO_2. We use the same variables as the JPL OCO-2 team and perform a multilinear regression on training subsets of the data which we believe should have a constant XCO_2. This includes OCO-2 data co-located with ground based TCCON measurements, performing a Small Area Analysis of the data, and implementing the Southern Hemisphere Approximation as described in Wunch et al. 2011.
- Thirdly, a global offset between TCCON data and OCO-2 is calculated.

3.2.2 Status

Some differences remain between our new custom bias correction and the bias correction applied by the JPL team and we are in discussions with the JPL team to better understand their method and fine-tune our bias correction.

Our next steps are to refine our bias correction by applying the new filters as advised by JPL, after which we will calculate the bias corrected results for the new datasets. At this stage we will perform a final validation of the datasets.

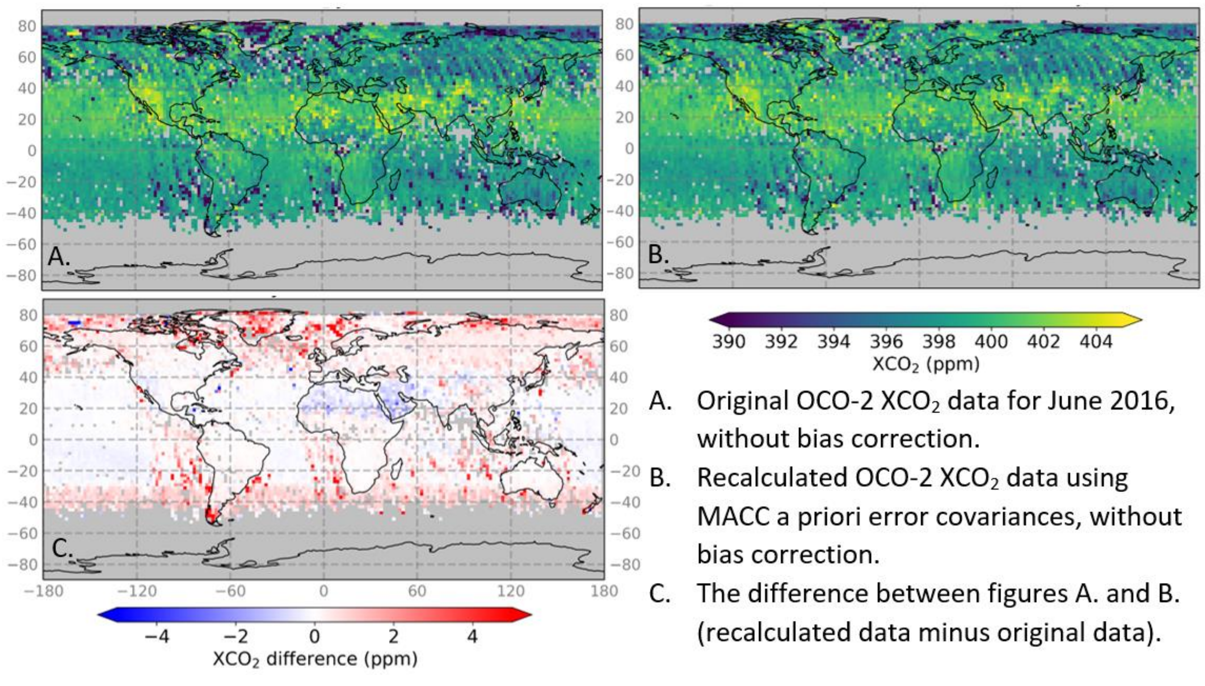

Figure 1: Comparison of the recalculated OCO-2 XCO_2 data to the original retrieval product by NASA for one month of observations (June 2016), both without bias correction.

4 Innovative data assimilation systems for global carbon flux estimates

The work performed as part of T1.2 consists of innovations to existing flux inversion frameworks, or the development of completely new set-ups. A common factor in the development work is the use of space-based observations, mainly XCO_2. When assimilating XCO_2 observations, the information content of the retrieved parameters is considered by applying the retrievals averaging kernel to the simulated XCO_2 samples. This way (i) the partitioning of information originating from observed radiances and prior information included in the retrieval algorithm, and (2) the error covariance and vertical resolution of the retrieval, can be properly considered during assimilation.

Table 1: Overview and status of the scheduled flux inversions. Type of the output product is baseline (B), innovation run (I) or new inversion set-up (N).

Data provider	Type	Short description	Status
WU	B	CTE default set-up (surface CO_2 only)	Completed
WU	I	CTE inversion using satellite-based XCO_2 observations (FOCAL OCO-2)	Completed
WU	I	CTE inversion using satellite-based XCO_2 observations (NASA OCO-2)	Completed
WU	I	CTE inversion with short window and satellite-based XCO_2 observations (NASA OCO-2)	Completed
WU	I	CTE inversion with combination of surface CO_2 and satellite-based XCO_2 observations	Ongoing
WU	N	CTE with coupled IFS-TM5 transport	Planned
WU	I	CTE with short window and coupled IFS-TM5 transport	Planned
WU	N	CTSF inversion using NIRv	Ongoing
WU	I	CTSF inversion using SIF anomalies and related shorter window length	Ongoing
CEA	B	CAMS default set-up (surface CO_2 only)	Completed
CEA	I	CAMS inversion using satellite-based XCO_2 observations (FOCAL OCO-2)	Planned
CEA	I	CAMS inversion using combination of surface CO2 and FOCAL OCO-2 XCO_2 observations	Planned
CEA	I	CAMS inversion using satellite-based XCO_2 observations (OCO-2 with tailored error structures)	Planned
CEA	I	CAMS inversion using combination of surface CO_2 and NASA OCO-2 XCO_2 observations	Completed
CEA	N	COS + surface CO_2 inversion	Ongoing
CEA	I	COS + surface CO_2 + FOCAL OCO-2 XCO_2 observations	Planned

Apart from XO2, Satellite-based observations of SIF and COS are/will be included in the developed innovative inversion systems. Note that satellite retrievals typically contain other by-products (e.g. surface pressure, albedo, aerosols) that can contain important information regarding the spatial covariance of errors. These are not considered in the current effort, but inclusion of these products in inversion systems is investigated in e.g. the VERIFY (Verifying Greenhouse Gas Emissions) project.

An overview of the planned flux inversions to be performed with these new or modified systems is contained in Table 1. As specified in the output protocol D1.1 we have aimed, where possible, to introduce changes to systems gradually such as to facilitate analysis of the impact of particular system changes. Each flux inversion therefore has a type indicator that designates whether it is a default set-up (baseline, B), a completely new set-up (new, N), or a run containing an innovative feature compared to either the baseline or new set-up (innovation run, I).

4.1 Satellite-based CO_2 flux estimates with CarbonTracker Europe (CTE)

4.1.1 Short description of the CTE data assimilation system for CO_2

The CarbonTracker data assimilation shell (CTDAS) (Van der Laan-Luijkx et al., 2017) is a modular data assimilation framework written in Python. In its default CarbonTracker Europe (CTE) setup this system uses observations of atmospheric CO_2 mole fractions to estimate the global carbon fluxes between the atmosphere, land biosphere and oceans. An ensemble Kalman smoother with a 5-week fixed-lag assimilation window (Peters et al., 2005) is used to optimize a state vector that contains weekly scaling factors for the biosphere and ocean flux prior estimates, which consist of pre-calculated space-time patterns. The optimized scaling factors vary per ecoregion, except for the Northern hemisphere land regions where they vary per 1°x1° grid box with exponentially decaying correlation between grid boxes belonging to the same ecoregion. Atmospheric transport is calculated with the two-way nested TM5 transport model (Krol et al., 2005, Huijnen et al., 2010), using a resolution of 3°x2° globally, with 1°x1° zoom regions over North America and Europe.

4.1.2 Description of system innovations

New modules have been implemented in CTDAS and TM5 to allow for the assimilation of satellite-based XCO_2 observations. The CTDAS module reads the XCO_2 observations from a netCDF file, handles the I/O communication with the atmospheric transport model, and afterwards reads the simulated samples. This new module is kept as general as possible such as to be useable for general total-column observations and both level- and layer-based data products. The new TM5 sampling module allows for the calculation of simulated XCO_2 samples, thereby directly applying the desired (level- or layer-based) retrieval's averaging kernel.

As satellite-based XCO_2 observations typically have a footprint much smaller than the scale (in both space and time) resolvable by the atmospheric transport model, they are pre-processed into 10-second bins prior to assimilation. These so-called super-observations are created in a 2-step inverse-variance-weighted averaging procedure, according to the method described in Crowell et al., 2019. The smoothed super-observations created this way are representative on a spatial scale similar to a typical transport model resolution.

In addition to the uncertainty provided with the retrieval product, a rough estimate for the model error is added to each observation i equal to $\epsilon_i = 0.8 \exp\left(\frac{90+lat_i}{300}\right)$, based on a parametrization by Frédéric Chevallier (personal communication, July 3, 2019).

4.1.3 Overview of resulting output products

Baseline

Our baseline carbon flux inversion consists of the default CTE setup with the following specifications:

- Biosphere prior flux: SiBCASA
- Ocean prior flux: ocean inversion by Jacobsen et al., 2007
- Fire imposed flux: GFED4s
- Fossil fuel imposed flux: EDGAR + IER, scaled to CDIAC
- Atmospheric observations: Hourly resolution (well-mixed conditions) ObsPack GLOBALVIEWplus v4.2 & NRTv4.4
- Period: from January 1, 2000 to January 31, 2019

This is the same flux inversion as submitted for the Global Carbon Budget 2019 (Le Quéré et al., 2019)

Innovation run 1

This flux inversion follows the same setup as the baseline run, but using a different type of observations, namely satellite-based XCO_2 from the FOCAL OCO-2 v08 retrieval. Due to availability of the observational data this inversion started at January 1, 2015. The global CO_2 concentration field required to initialize the atmospheric transport model is extracted from the simulated field obtained with the baseline run in order to minimize the required spin-up time.

Innovation run 2

Similar to innovation run 1, but using the NASA OCO-2 L2 lite FP v9r XCO_2 retrieval. Note that only nadir and glint observations over land are assimilated due to the possibility of a remaining bias over oceans. Due to availability of the observational data this inversion started at October 1, 2014.

Innovation run 3

Similar to innovation run 2, but using a shorter window length of 3 instead of 5 weeks (by reducing the lag from 5 to 3 weeks).

4.1.4 Global flux results obtained with the different CTE set-ups

Below a first comparison of the results obtained with the different CTE flux inversions is presented. Results with the label SURF indicate flux results obtained with our baseline inversion set-up. Figure 2 shows the distribution of ocean and biogenic fluxes over the northern, tropical and southern regions, whereas in Figure 3 a validation of the obtained atmospheric CO_2 concentrations is presented based on independent observation locations. Overall, we observe that our XCO_2-based DA system can now match our in-situ based flux estimates on global to regional scales, and suggest a smaller tropical sink and seasonal cycle.

Note the deviating result for the southern region obtained with the FOCAL retrieval for the year 2015. This is a spin-up effect caused by the fact that XCO_2 observations for FOCAL are only available from January 2015 on. However, we also note deviations in simulated mole fractions at stations that are sensitive to calculated ocean fluxes. In particular the seasonal bias at the South Pole Observatory station (SPO) is notable, suggesting a remaining bias over (southern) ocean areas in the FOCAL retrieval.

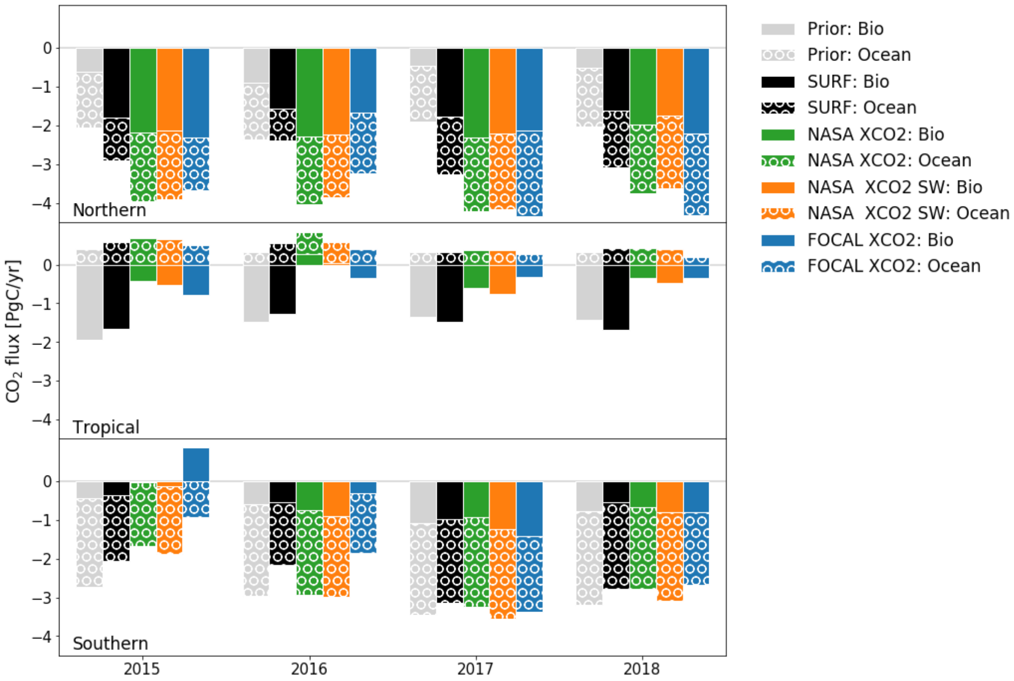

Figure 2: Comparison of yearly total fluxes as obtained with the baseline and innovation CTE flux inversion set-ups. The regions are aggregated based on TRANSCOM land and ocean regions.

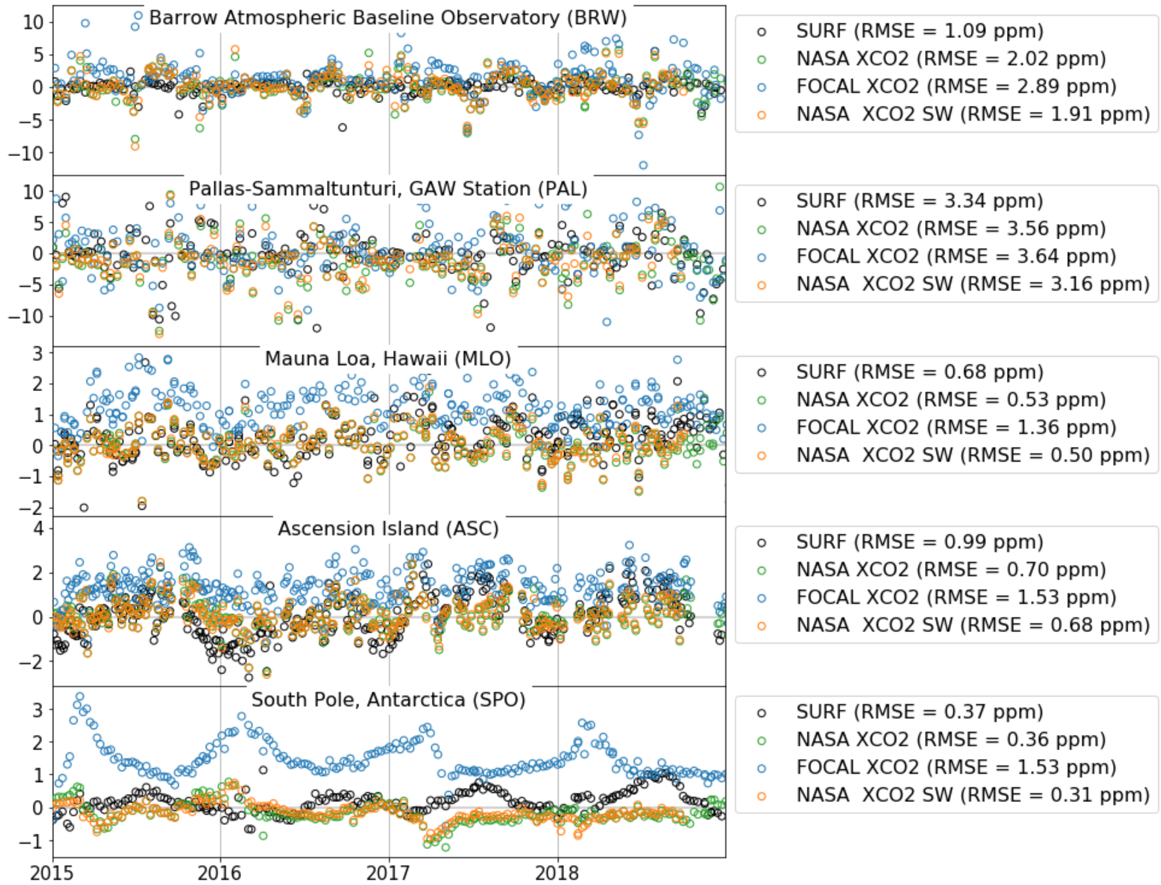

Figure 3: Residuals in atmospheric CO_2 mole fraction concentrations at several background in-situ observation locations. Measurements from these in-situ observation sites are not assimilated in the XCO_2-based inversions.

4.1.5 Data availability

Global flux inversion results can be downloaded from ftp://ftp.wur.nl/carbontracker/data/SpecialProjects/CHE/.

4.2 Optimization of surface-flux parameters in a long-window inversion

4.2.1 Description of the CarbonTracker Statistical Fit (CTSF) flux-inversion system

A new, long-window, inversion set-up for optimizing biosphere fluxes is implemented in CTDAS. This new system is designed to study long-term trends and interannual variations in NEE, and to exploit the spatio-temporal variations present in photosynthesis proxies, like sun-induced fluorescence (SIF). It is based on earlier work by Christian Rödenbeck, who made use of temperature anomalies to study the long-term and interannual variability of NEE (Rödenbeck et al., 2018).

It is assumed that the long-term mean and temporal trends of NEE can be described by a statistical function consisting of a second-order in time polynomial enhanced with *N* yearly harmonics whose amplitude can vary linearly in time. Interannual and spatial variability in NEE is introduced though a linear regression of proxy anomalies onto NEE anomalies. The statistical description for NEE thus has the form

$$NEE(x,y,t) = p_0 + p_1 t + p_2 t^2 + \sum_{i=1}^{N}\left[(a_i + b_i t)\cdot \sin\left(\frac{2\pi}{T} it\right) + (c_i + d_i t)\cdot \cos\left(\frac{2\pi}{T} it\right)\right] + \gamma^P \Delta P(x,y,t)$$

where T represents the yearly period of the harmonics in days, ΔP is the anomaly field of the considered photosynthesis proxy, and the model parameters in blue ($p_0, p_1, p_2, a_i, b_i, c_i, d_i$ and γ^P) are optimized during the inversion. One set of parameters per ecoregion is optimized for the full window of the inversion, spanning several years. The regression coefficients γ^P additionally vary per calendar month. For $N = 4$ the state vector thus contains a total of 4154 elements, which are assumed to be uncorrelated.

The optimization of the state vector is done with the CTDAS Ensemble Kalman smoother approach, but this time using a single window. The cycle length thus equals the full timespan of interest and there is no lag. Compared to our default CTE setup a reduced set of observations is assimilated. The included observations contain flask measurements from marine boundary layer sites only, which capture variations in the background atmospheric CO_2 concentration. Local spatio-temporal flux variations are based on observational data of the chosen photosynthesis proxy. This choice for the assimilated observations implies that we can suffice with evaluating the transport model at a rather coarse resolution of 6°x4° globally. Tests with a higher resolution have been performed, but indeed the effect on the obtained posterior fluxes was very limited.

4.2.2 Overview of the resulting output products

New inversion set-up

A CTSF inversion for the years 2000-2018 has been performed using NIR_V as proxy (Badgley et al., 2017), with the following specifications:

- Atmospheric transport: TM5 on 6°x4° resolution globally
- Ocean imposed flux: optimized ocean flux resulting from the CTE baseline inversion
- Fire imposed flux: GFED4s
- Fossil fuel imposed flux: EDGAR + IER, scaled to CDIAC

- Atmospheric observations: Reduced set of hourly resolution (well-mixed conditions) observations from ObsPack GLOBALVIEWplus v4.2 & NRTv4.4, containing flask measurements only
- Period: from January 1, 2000 to December 31, 2018

Innovation run 1

This biosphere flux inversion follows the same setup as the above described CTSF run, but using satellite-based SIF instead of NIR_V anomalies as proxy. The SIF observations are from the SIFTER v2 product (Sanders et al., 2016 and Van Schaik 2016), which is retrieved from the GOME-2 instrument aboard the MetOp satellite. Due to the availability of the SIFTER product, this inversion is performed for the period from January 1, 2007 to December 31, 2016 only.

4.2.3 Status / Data availability

The system is up and running, but further validation of the results has to be performed before the datasets can be made available. We expect that they will be available for download from ftp://ftp.wur.nl/carbontracker/data/SpecialProjects/CHE/ early 2020.

4.3 Satellite-based CO_2 flux estimates with CAMS

4.3.1 Description of the flux inversions

As a baseline, we use inversion CAMS v18r3 that was provided by CAMS and is being post-processed for the needs of CHE. CAMS v18r3 covers the period 1979-2018 and assimilated surface data from mainly GLOBALVIEWplus v5.0, rather than from GLOBALVIEWplus v4.0 exclusively.

As an incremental version, we assimilated the same surface measurements together with NASA's OCO-2 retrievals v9 in the configuration chosen for the OCO-2 Model Intercomparison Project 2 (i.e. prescribed prior fossil fuel emissions and binned observations within 10-s windows). In order to allow some flexibility for the inversion system to compensate for possible inconsistencies between the two observation types as seen by the transport model, we have added a monthly global offset to the retrievals in the control vector. The adjoint of the observation operator has been consistently adapted. The period covered is September 2014-May 2019. It has not been possible yet to run this incremental version with the same prior fossil fuel emissions.

The OCO-2 retrievals prepared by the University of Bremen have been delivered recently and are being processed. A first inversion where they have been assimilated alone (i.e. without surface measurements) should be ready by the end of the year. The period covered is January 2015-December 2018. An incremental run should follow in the beginning of the year.

The OCO-2 retrievals prepared by the University of Leicester will also be assimilated alone (i.e. without surface measurements) when they are ready.

4.3.2 Data availability

Monthly means and 3-hourly flux fields as resulting from the NASA OCO-2 XCO_2 + SURF inversion can be downloaded from http://dods.lsce.ipsl.fr/invsat/CHE/WP1/.

Results from the baseline run will soon be made available on the CAMS server (https://apps.ecmwf.int/datasets/data/cams-ghg-inversions/).

4.4 Flux inversion with assimilation of carbonyl sulfide (COS) observations

We aim to assimilate measurements of Carbonyl sulfide (COS) that is related to the Gross Primary Production (GPP) of terrestrial vegetation through the leaf-scale relative uptake ratio of COS to CO_2 fluxes. Other sources or sinks of COS are from the ocean, the wetlands, the soils, some industrial processes, fire and oxidation in the atmosphere. Each term is uncertain and current best estimates of some of them are still likely biased. In order to investigate how to exploit the GPP information of COS within this complex framework, we have designed an analytical system that allows us to make many tests rapidly once the matrix of atmospheric transport is computed. We assimilate surface measurements of COS for the period 2009-2015 to optimize monthly fluxes of GPP, respiration and the other sources of COS in large regions of the globe. We currently try to assimilate them together with surface CO_2 measurements, before trying to assimilate them together with CO_2 satellite retrievals at a later stage.

5 Conclusion

The creation of a new set of input data for data assimilation systems is presented, consisting of a new fast and efficient XCO_2 production algorithm, and an XCO_2 retrieval with tailored error covariance structures for use in a flux inversion system. Together, these sets provide new insights on the data processing for XCO_2, their assimilation, and the best method to maximize data usage, minimize biases, separate biogenic and anthropogenic fluxes, while maintaining a fast operating chain.

In addition, the development of innovative flux inversion output data products is presented. These consist of both innovations applied to existing DA set-ups, like adapting them for use with the new XCO_2 input data, as well as the development of new flux inversion systems, like the optimization of surface-flux parameters in a long-window inversion setup or the construction of a COS inversion system. The first global flux results obtained with these systems are made available for benchmarking.

The results of task 1.2 will be the input for a thorough benchmarking and quality assessment in task 1.3. Details about the chosen strategy of quality control, as well as the assessment of the innovative DA systems presented in the current document, will be presented in deliverable D1.3 in month 39. Development work to the input and output data products of task 1.2 will continue alongside the aforementioned benchmarking process of the already available products. Additional and updated products will be made available in the course of Task 1.3.

Apart from WP1, the current work is also connected to WP5: the lessons learned from the current task, as well as the progressive insights we expect from task 1.3, form essential input for respectively the progress and final reports on service elements for CO_2 earth observation integration, CO_2 emission and transport model integration, and requirements for data assimilation methodology.

6 References

Buchwitz, M., R. Detmers, M. Reuter, H. Boesch, F. Chevallier, B. Dils, C. Frankenberg, O. Hasekamp, S. Houweling, G. Lichtenberg, J. Marshall, R. Parker, O. Schneising, C. O'Dell, ESA Climate Change Initiative (CCI) Product Specification Document (PSD) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), 6-June-2014, Version 3, 2014. Link: http://www.esa-ghg-cci.org/index.php?q=webfm_send/160

Buchwitz, M., Reuter, M., Schneising, O., Hewson, W., Detmers, R. G., Boesch, H., Hasekamp, O. P., Aben, I., Bovensmann, H., Burrows, J. P., Butz, A., Chevallier, F., Dils, B., Frankenberg, C., Heymann, J., Lichtenberg, G., De Mazière, M., Notholt, J., Parker, R., Warneke, T., Zehner, C., Griffith, D. W. T., Deutscher, N. M., Kuze, A., Suto, H., and Wunch, D.: Global satellite observations of column-averaged carbon dioxide and methane: The GHG-CCI XCO2 and XCH4 CRDP3 data set, Remote Sensing of Environment 203, 276–295, http://dx.doi.org/10.1016/j.rse.2016.12.027, 2017.

Buchwitz, M., Reuter, M., Schneising, O., Bovensmann, H., Burrows, J. P., Boesch, H., Anand, J., Parker, R., Detmers, R. G., Aben, I., Hasekamp, O. P., Crevoisier, C., Armante, R., Zehner, C., Schepers, D., Copernicus Climate Change Service (C3S) Global Satellite Observations of Atmospheric Carbon Dioxide and Methane, Adv. Astronaut. Sci. Technol., https://doi.org/10.1007/s42423-018-0004-6, 2018.

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., Liu, J., Deng, F., Feng, L., McKain, K., Chatterjee, A., Miller, J. B., Stephens, B. B., Eldering, A., Crisp, D., Schimel, D., Nassar, R., O'Dell, C. W., Oda, T., Sweeney, C., Palmer, P. I., and Jones, D. B. A.: The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network, *Atmos. Chem. Phys*., 19, 9797–9831, https://doi.org/10.5194/acp-19-9797-2019, 2019

Huijnen, V., Williams, J., van Weele, M., van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., de Laat, J., Boersma, F., Bergamaschi, P., van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W., The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445–473, https://doi.org/10.5194/gmd-3- 445-2010, 2010

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, *Atmos. Chem. Phys*., 5, 417– 432, https://doi.org/10.5194/acp-5-417-2005, 2005

Peters, W., J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski, L. Bruhwiler, and P. P. Tans, An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, *J. Geophys. Res*., 110, D24304, https://doi.org/10.1029/2005JD006157, 2005

Reuter et al., 2019: M.Reuter and IUP's GHG CCI team: Algorithm Theoretical Basis Document Version 1 (ATBDv1) - Retrieval of XCO2 from the OCO-2 satellite using the Fast Atmospheric Trace Gas Retrieval (FOCAL), ESA GHG-CCI+ project, 2019

Reuter and Buchwitz, 2019: M. Reuter and M. Buchwitz: ESA Climate Change Initiative "Plus" (CCI+) Product User Guide (PUG) for the FOCAL XCO2 OCO-2 Data Product CO2_OC2_FOCA (v08) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG), ESA GHG CCI+, 2019

Rödenbeck, C., Zaehle, S., Keeling, R., and Heimann, M.: How does the terrestrial carbon exchange respond to inter-annual climatic variations? A quantification based on atmospheric CO2 data, *Biogeosciences*, 15, 2481–2498, https://doi.org/10.5194/bg-15-2481-2018, 2018.

Sanders A., Verstraeten W., Kooreman M., van Leth T., Beringer J., Joiner J., Spaceborne sun-induced vegetation fluorescence time series from 2007 to 2015 evaluated with Australian flux tower measurements, *Remote. Sens.* 8, 895, doi:10.3390/ rs8110895, 2016

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

van Schaik E., Retrieving sun-induced fluorescence from the Global Ozone Monitoring Experiment 2. See https://goo.gl/fgAMG7. Wageningen University MSc Internship report for Royal Netherlands Meteorological Institute (KNMI), 2016

Wunch, D., et al, A method for evaluating bias in global measurements of CO2 total columns from space, *Atmos. Chem. Phys*., 11, 12317-12337, doi:10.5194/acp-11-12317-2011, 2011.

Document History

Version	Author(s)	Date	Changes
0.1	Liesbeth Florentie (WU)	03/12/2019	
0.2	Liesbeth Florentie (WU)	10/12/2019	Updated overview CEA inversions in table 1 + added description of COS inversion
1.0	Liesbeth Florentie (WU)	18/12/2019	- Addition of a glossary on p.6 - Section 1: added a line about planned work on upgrading modelled transport in second part of task 1.2 - Section 2.1: added a line about the QA/QC planned for task 1.3 - Throughout: Replaced MACC by CAMS - Section 2.2.3: updated the text about contribution of ULEIC - Extended the introduction of section 4 by including more information about the application of the retrievals averaging kernel to deal with the inconsistency in retrieved and simulated XCO_2, and about considered retrieval by-products. - Section 4.2.1: added extra info about chosen transport model resolution & assimilated observations - Added a line to section 4.3.1 about CEA inversion with tailored error structures - Extended the conclusion: description of connection current deliverable to other tasks and WPs.

Internal Review History

Internal Reviewers	Date	Comments
Denis Simeoni and Sandrine Mathieu (Thales Alenia Space)	13/12/2019	Approved with comments
Laure Brooker (Airbus SAS)	17/12/19	Approved with comments

Estimated Effort Contribution per Partner

Partner	Effort
WU	12
CEA	4.2
UB	4.03
ULEIC	3.41
Total	**23.64**

This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.