



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

## **Assimilation of atmospheric CO<sub>2</sub> observations from space can support national CO<sub>2</sub> emission inventories**

Downloaded from: <https://research.chalmers.se>, 2026-04-06 02:14 UTC

Citation for the original published paper (version of record):

Kaminski, T., Scholze, M., Rayner, P. et al (2022). Assimilation of atmospheric CO<sub>2</sub> observations from space can support national CO<sub>2</sub> emission inventories. *Environmental Research Letters*, 17(1). <http://dx.doi.org/10.1088/1748-9326/ac3cea>

N.B. When citing this work, cite the original published paper.

ENVIRONMENTAL RESEARCH  
LETTERS

## LETTER

Assimilation of atmospheric CO<sub>2</sub> observations from space can support national CO<sub>2</sub> emission inventories

## OPEN ACCESS

## RECEIVED

3 September 2021

## REVISED

1 November 2021

## ACCEPTED FOR PUBLICATION

24 November 2021

## PUBLISHED

4 January 2022

Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Thomas Kaminski<sup>1,\*</sup>, Marko Scholze<sup>2</sup>, Peter Rayner<sup>3</sup>, Michael Voßbeck<sup>1</sup>, Michael Buchwitz<sup>4</sup>, Maximilian Reuter<sup>4</sup>, Wolfgang Knorr<sup>1,2</sup>, Hans Chen<sup>2</sup> , Anna Agustí-Panareda<sup>5</sup>, Armin Löscher<sup>6</sup> and Yasjka Meijer<sup>6</sup>

<sup>1</sup> The Inversion Lab, Hamburg, Germany

<sup>2</sup> Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

<sup>3</sup> University of Melbourne, Melbourne, Australia

<sup>4</sup> University of Bremen, Institute of Environmental Physics (IUP), Bremen, Germany

<sup>5</sup> European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

<sup>6</sup> ESA, Noordwijk, The Netherlands

\* Author to whom any correspondence should be addressed.

E-mail: [Kaminski@Inversion-Lab.com](mailto:Kaminski@Inversion-Lab.com)

**Keywords:** CO<sub>2</sub> Monitoring Mission, fossil fuel emissions, Monitoring and Verification System, Paris Agreement, inverse modelling

Supplementary material for this article is available [online](#)

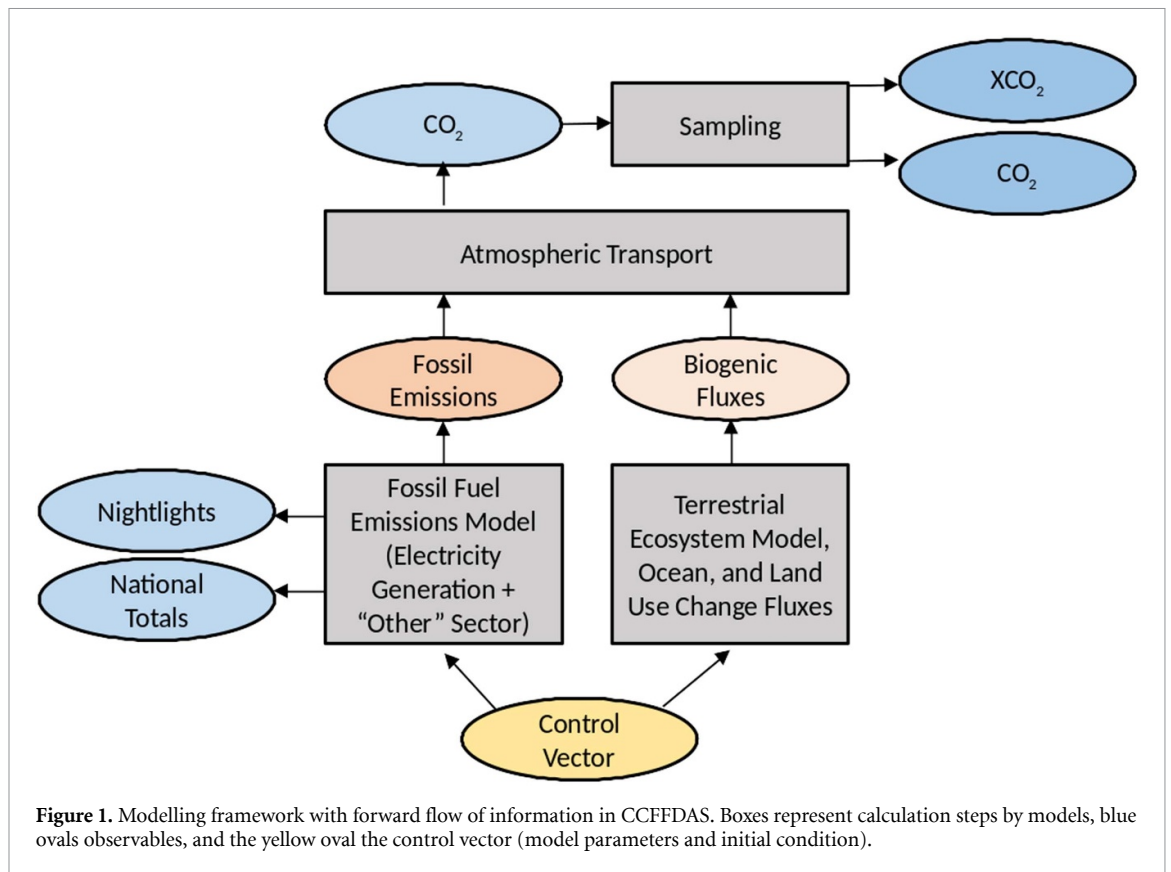
**Abstract**

The Paris Agreement establishes a transparency framework for anthropogenic carbon dioxide (CO<sub>2</sub>) emissions. Its core component are inventory-based national greenhouse gas emission reports, which are complemented by independent estimates derived from atmospheric CO<sub>2</sub> measurements combined with inverse modelling. It is, however, not known whether such a Monitoring and Verification Support (MVS) capacity is capable of constraining estimates of fossil-fuel emissions to an extent that is sufficient to provide valuable additional information. The CO<sub>2</sub> Monitoring Mission (CO2M), planned as a constellation of satellites measuring column-integrated atmospheric CO<sub>2</sub> concentration (XCO<sub>2</sub>), is expected to become a key component of such an MVS capacity. Here we provide a novel assessment of the potential of a comprehensive data assimilation system using simulated XCO<sub>2</sub> and other observations to constrain fossil fuel CO<sub>2</sub> emission estimates for an exemplary 1-week period in 2008. We find that CO2M enables useful weekly estimates of country-scale fossil fuel emissions independent of national inventories. When extrapolated from the weekly to the annual scale, uncertainties in emissions are comparable to uncertainties in inventories, so that estimates from inventories and from the MVS capacity can be used for mutual verification. We further demonstrate an alternative, synergistic mode of operation, with the purpose of delivering a best fossil fuel emission estimate. In this mode, the assimilation system uses not only XCO<sub>2</sub> and the other data streams of the previous (verification) mode, but also the inventory information. Finally, we identify further steps towards an operational MVS capacity.

**1. Introduction**

The Parties to the 2015 Paris Agreement (UNFCCC 2015) have agreed to reduce global emissions of greenhouse gases and implement a transparency framework building on regular national greenhouse gas inventories. As part of the 2019 refinement (Calvo Buendia *et al* 2019) of their guidelines for national greenhouse gas inventories (IPCC 2006), the Intergovernmental Panel on Climate Change recommends

the implementation of systems for additional scientific validation that support emission inventories by an independent approach based on atmospheric observations and inverse modelling. A first attempt to derive CO<sub>2</sub> fluxes for the period 1990–2018 from both anthropogenic and the natural sources and sinks has recently been performed for the European Union and UK by Petrescu *et al* (2021). The study uses a combination of state-of-the-art process-based models and inventories (so-called bottom-up estimates)



and inverse modelling approaches (also called top-down methods). They concluded that independent approaches based on currently available data are consistent with the inventory data, however, their uncertainty is too large to allow a verification.

The European Commission's Copernicus programme is preparing an operational Monitoring and Verification Support (MVS) with CO2M as an essential component (Janssens-Maenhout *et al* 2020). CO2M is planned as a constellation of multiple polar orbiting satellites with imaging capability sampling XCO<sub>2</sub> at a resolution of 4 km<sup>2</sup> (ESA 2018). The MVS capacity's goal is to derive fluxes both at the national scale as well as the scale of megacities and certain emission hot spots (Pinty *et al* 2017). The attribution of such atmospheric concentration data to specific emission sources requires comprehensive inverse modelling systems, which ideally include modules that accurately represent the processes behind fossil and biogenic fluxes (Ciais *et al* 2015).

The potential performance of CO2M to quantify CO<sub>2</sub> emissions from cities and power plants globally lumped into so-called emission clumps has been assessed by Wang *et al* (2020). Using a plume monitoring inversion framework based on a Gaussian plume model they estimated that future CO2M observations can constrain, ten times within the 1 year study period, large clumps with annual emissions larger than 2 MtC yr<sup>-1</sup> with a relative posterior uncertainty of less than 20%. In another study focussing on large point sources Kuhlmann *et al*

(2021) used a plume detection algorithm combined with a mass-balance approach to assess the potential performance of CO2M to quantify CO<sub>2</sub> emissions of large point sources, for which they used 15 large power plants in the border region of the Czech Republic, Germany and Poland. For constraining the shape of the CO<sub>2</sub> emission plume, CO2M's capability of also observing NO<sub>2</sub> was essential and emissions could be estimated with 39%–150% uncertainty from a single CO2M overpass. However, they concluded that the development of advanced inversion systems is required to further reduce the uncertainties.

Here, we have implemented a comprehensive prototype of a global Carbon Cycle Fossil Fuel Data Assimilation System (CCFDAS) from existing components (see figure 1) for assessing the added value of CO2M observations in constraining national scale emissions. These components are the Fossil Fuel Data Assimilation System (FFDAS) from Asefi-Najafabady *et al* (2014), which includes two fossil fuel emission models, one for the electricity generation sector (at point scale of the plants) and one for all other fossil fuel emission sectors combined (termed 'other sector'), and the Carbon Cycle Data Assimilation System from Kaminski *et al* (2017), which includes a model of the terrestrial carbon cycle, land use change fluxes, and exchange fluxes with the ocean. To link simulated surface fluxes to atmospheric CO<sub>2</sub> concentrations we use the atmospheric transport model TM3 (Heimann and Körner 2003). The components of the modelling chain are described in more detail in section 2.

Our aim is to provide, for the first time, an estimate of the expected uncertainty range of fossil fuel emissions estimates of a future fully operational MVS capacity targeting the national scale emissions. For this purpose, we employ the quantitative network design technique (QND) (Kaminski and Rayner 2017), which assesses the benefit of a set of (real or hypothetical) observations by propagating uncertainties through a modelling chain (see section 1 of the supplementary text (available online at [stacks.iop.org/ERL/17/014015/mmedia](https://stacks.iop.org/ERL/17/014015/mmedia))). The results of the assessment are posterior uncertainties on selected target quantities, in our case national sectorial fossil fuel emissions for five countries (see table 1). These countries were selected to cover both hemispheres and the corresponding seasons, varying levels of industrialisation and of uncertainties in inventories (see section 2 of the supplementary text). They also needed to be large enough to encompass several TM3 grid cells. We first target the weekly scale, at which typically no information from inventories is available. More specifically, we use estimates of the uncertainty of atmospheric CO<sub>2</sub> concentration data over the first week of June 2008 to constrain emissions over that week. This falls during the northern growing season when biogenic fluxes impact strongly atmospheric concentrations. In addition, we compare the impact of XCO<sub>2</sub> data as provided by CO2M to the impact of *in-situ* CO<sub>2</sub> data as provided by the surface network.

## 2. Methodology

Our CCFDAS is based on a modelling chain (figure 1) that includes models of fossil fuel emissions and terrestrial natural fluxes and dedicated observation operators (Kaminski and Mathieu 2017) linking the state of the model to each of the data streams it can assimilate. The fossil fuel CO<sub>2</sub> emissions model follows Asefi-Najafabady *et al* (2014) and distinguishes fossil fuel emissions from two sectors: an electricity generation sector and a sector for all other fossil fuel emissions denoted as the ‘other’ sector. Details on the fossil fuel CO<sub>2</sub> emissions model, the underlying data sets and the corresponding observation operators are provided in section 2 of the supplementary text. The terrestrial biosphere model is conceptually similar to the one used by Kaminski *et al* (2002) and based on Knorr and Heimann (1995). It was previously employed for assimilation of XCO<sub>2</sub> (Kaminski *et al* 2017) and is described in section 3 of the supplementary text. The atmospheric component of our modelling chain is the global atmospheric transport model TM3 (Heimann and Körner 2003). Its setup is presented in section 4 of the supplementary text, together with the estimates of the uncertainty in the atmospheric CO<sub>2</sub> measurements following Buchwitz *et al* (2013).

The assessments of the observational scenarios within the prototype assimilation system apply the QND technique. QND performs a formal uncertainty propagation through the modelling chain and relies on sensitivity information of our modelling chain, which is efficiently provided by the automatic differentiation tool TAPENADE (Hascoët and Pascual 2013). For details we refer to section 1 of the supplementary text.

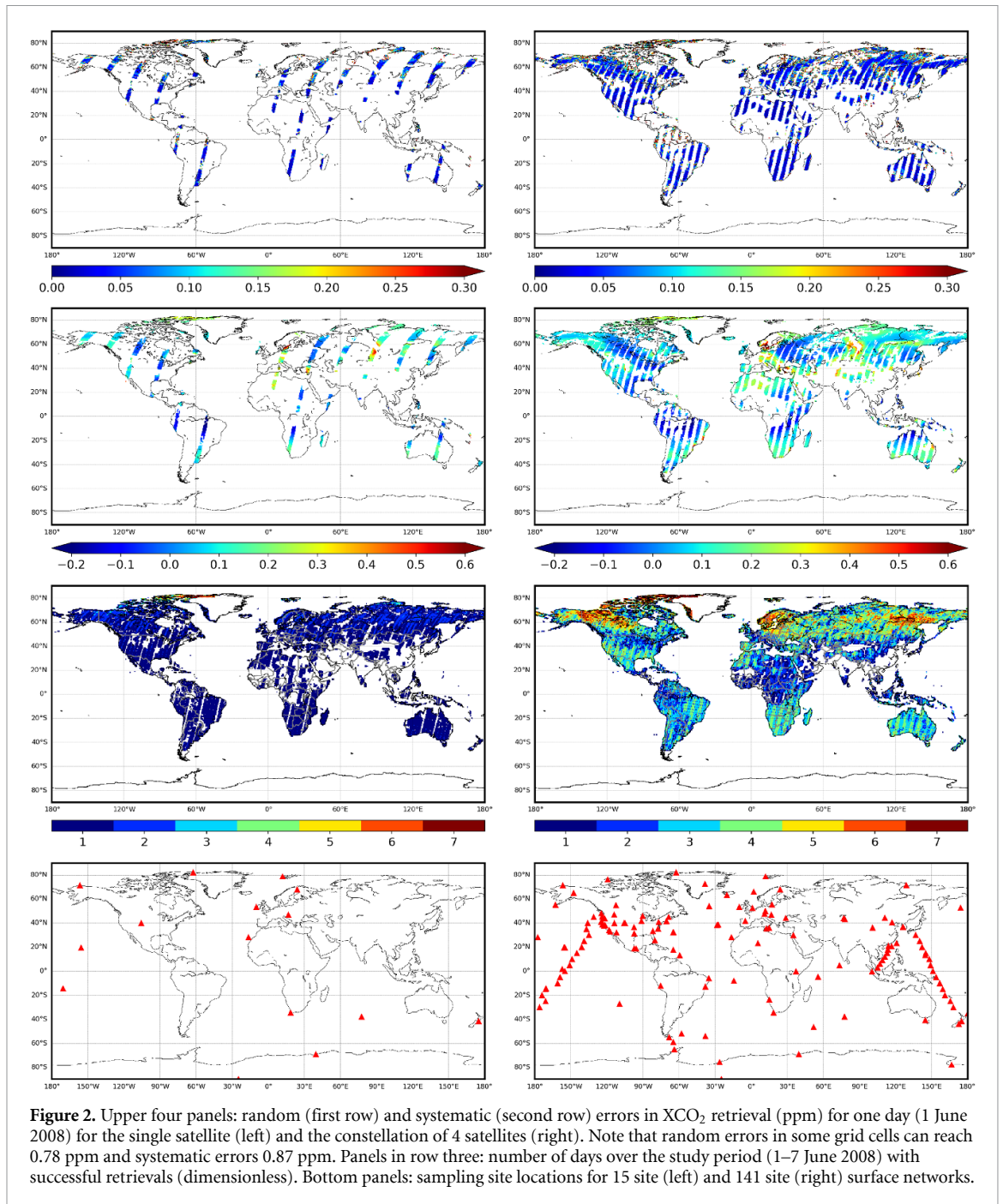
Ideally the assimilation system would be set up to assimilate XCO<sub>2</sub> over a longer period (at least a year) to infer fossil fuel emissions in weekly (or shorter) temporal resolution. In such an ideal setup, atmospheric observations made in one week could help to constrain (potentially remote) emissions from the preceding weeks, and after a sufficiently long spinup the spatial structures in the initial atmospheric CO<sub>2</sub> concentration would be negligible. For this initial analysis, however, we use a computationally pragmatic setup which restricts the analysis period to a single week, for which we assume constant emissions. As in the above ideal setup we neglect the sensitivity of the XCO<sub>2</sub> observations to the initial concentration structure. While this ignores an important uncertainty for our experimental setup, this uncertainty diminishes rapidly for longer assimilation periods (Peylin *et al* 2005). This is compensated by the fact that in the present setup, constraints from observations that would arise from subsequent weeks are being ignored. This second effect is probably stronger meaning our simplified setup would tend to underestimate the constraint available in a real case.

## 3. Results and discussions

The prototype can simultaneously assimilate the following data streams (see figure 1): XCO<sub>2</sub>, *in-situ* atmospheric CO<sub>2</sub>, nightlight intensity, and sectorial national emission totals (typically derived by inventories). There are two principal modes of operating the system. The first is the verification mode, in which the sectorial national emission totals are excluded from the list of data streams to be assimilated. In the verification mode the system delivers an estimate of the sectorial national emissions that is independent of the inventory-derived national emission totals. This independence would be limited, if the system’s sectoral emission models relied on information that was also used in the compilation of the inventories. While our model of the other sector is clearly independent, it is more difficult to rule out potential overlaps between our data base for the power plant emissions and the data entering the inventories. The second mode of operating the prototype is the synergistic mode, in which the inventory-derived sectorial national emission totals are assimilated into the system. In the synergistic mode the system delivers a best estimate

**Table 1.** Posterior uncertainty (including atmospheric and other observational constraints) of fossil fuel emission rates (1 sigma) during first week of June 2008 for ‘other sector’ and electricity generation sector for five countries shown by their International Organization for Standardization (ISO) country codes, and annual average weekly emission rate (not uncertainty) for 2008 from national inventories for comparison (bottom row).

Scenario	Description	Other sector					Electricity generation sector					
		AUS	BRA	CHN	DEU	POL	AUS	BRA	CHN	DEU	POL	
1	Surface 15 sites	9.03	16.70	177.31	12.18	4.70	0.28	0.17	2.36	0.43	0.23	
2	Surface 141 sites	4.57	8.21	8.29	2.60	2.10	0.28	0.17	2.36	0.43	0.23	
3	1 satellite (default)	0.30	0.42	3.43	0.97	0.38	0.27	0.17	2.21	0.43	0.23	
4	4 satellites	0.25	0.29	2.38	0.79	0.33	0.26	0.17	2.07	0.43	0.23	
5	Default with ocean	0.29	0.41	2.93	0.94	0.37	0.27	0.17	2.20	0.43	0.23	
6	Default with repr. error	0.35	0.68	4.68	1.36	0.62	0.28	0.17	2.28	0.43	0.23	
7	Default for emission change	0.25	0.43	2.88	0.75	0.40	0.27	0.17	2.14	0.43	0.23	
8	Default with national total	0.03	0.16	1.84	0.08	0.05	0.04	0.06	1.43	0.07	0.05	
9	141 sites with national total	0.03	0.18	2.30	0.08	0.05	0.04	0.06	1.58	0.07	0.05	
10	No atmospheric observations	9.03	16.70	177.33	24.32	7.29	0.28	0.17	2.36	0.43	0.23	
—	National inventory	0.90	1.67	17.73	2.43	0.73	1.15	0.22	16.36	1.76	0.83	
					Annual average weekly emission rate (MtC week <sup>-1</sup> )							

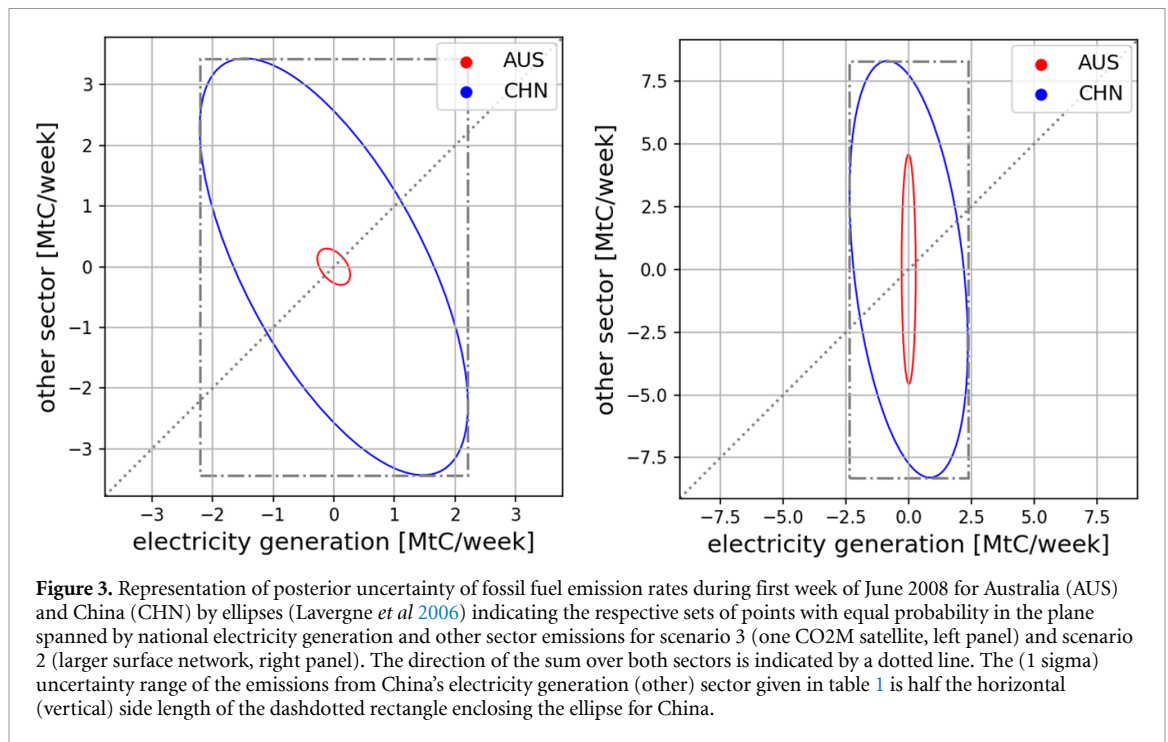


of the sectorial national emissions that combines the information in the inventory-derived national emission totals with the information provided by the atmospheric and other observations and by the model.

For our analysis we define ten observational scenarios (table 1), each of which uses subsets of these observations to produce posterior (after data assimilation) uncertainties for country-scale sectorial fossil fuel emission rates during the first week of June 2008. For scenarios 1–7 we operate the system in verification mode and for scenarios 8 and 9 we operate it in synergistic mode. The 2008 annual average weekly fossil fuel emission rate based on inventories

(International Energy Agency 2011) is provided for comparison (see section 2 of the supplementary text).

The first two scenarios assume the assimilation of continuous observations of atmospheric CO<sub>2</sub> from global surface networks. Scenario 1 uses a network of 15 *in-situ* measurement sites (figure 2, bottom row, left hand panel) and leads to posterior uncertainties of the other sector emissions that amount to between five (Germany, Poland) and ten times the annual average weekly emission rate. We explain the somewhat better performance for the two countries in Europe by strategically well positioned sites in and around that continent.



Indeed, for a reference scenario without any atmospheric observations (scenario 10), the constraint on fossil fuel emissions from the two European countries is at the level of the other three countries. Scenario 2 is based on an existing large network (141 sites, [www.esrl.noaa.gov/gmd/ccgg/flask.php](http://www.esrl.noaa.gov/gmd/ccgg/flask.php)), but with the additional assumption that all sites provide continuous observations of atmospheric CO<sub>2</sub>. In reality, the vast majority of these sites (many of them points on ship tracks) currently provide only approximately weekly flask samples. But even for this much larger network, posterior uncertainties of fossil fuel emission rates are above the inventory-derived annual average weekly emission rate (other sector), with the exception of China. But even the posterior uncertainty of China ( $\approx 50\%$  of inventory-based emissions) is still too large to be relevant for verification purposes.

The results are markedly different for scenario 3 (default satellite scenario) using a single CO<sub>2</sub>M satellite (figure 2, second and third rows, left hand panels) providing observations only over land (nadir viewing mode), and using our standard CO<sub>2</sub>M data uncertainty (see section 4 of the supplementary text). The posterior uncertainties for the other sector are all well below the inventory based average rate. This gives a first indication that CCFFDAS in combination with XCO<sub>2</sub> data has the potential to provide useful information for the purpose of cross validation with national inventories, as opposed to a system using surface stations only. Note, however, that the situation is different for the electricity generation sector, for which more prior information is available to constrain emission estimates, even though our prior uncertainty on power plant emissions is not

particularly tight (see section 2 of the supplementary text). Hence atmospheric observations have a larger impact on the other sector. Our finding regarding the electricity generation sector illustrates an important aspect of how CCFFDAS gains information about separate sectors. Let us assume that we identify a sub-sector (with a suitable sectorial emission model) of what is currently called the 'other sector' that can be constrained by additional independent observations, while we retain our current emission model and observations for the remainder of the 'other sector'. Because satellite observations capture the sum over all sectors, the additional observations will result in reduced uncertainties not only for the new subsector, but also for the remainder of the 'other sector'. Furthermore, after splitting off the new subsector, both the emissions of the remainder, i.e. the new and relatively uncertain 'other sector', and their posterior uncertainty will be reduced via recalibration of the sectorial model within the CCFFDAS.

A single CO<sub>2</sub>M satellite achieves a strong uncertainty reduction for the other sector of all countries compared to the surface network, and a slight uncertainty reduction for the electricity generation sectors of the larger countries China and Australia. We also note a strong negative uncertainty correlation between the two sectors of a given country (e.g.  $-0.66$  for China or  $-0.44$  for Australia). As illustrated by the rotation of the respective ellipses in the left hand panel of figure 3, such negative uncertainty correlation means that the sum of the sectors (direction indicated by dotted line) is better constrained than the difference of the sectors, a result of atmospheric mixing of the sectorial emissions. In other words: if we underestimate one target quantity we are likely to

overestimate the other. This aspect of the problem is useful for the MVS capacity, because it offers the opportunity to reduce uncertainty in one sector by constraining the remaining sector(s) through additional observations. The impact of the atmospheric constraint on the negative correlations between the uncertainties of the emission rates from the two sectors is illustrated by the right hand panel of figure 3. The weaker atmospheric constraint afforded by the surface network (scenario 2) results in much smaller uncertainty correlations (e.g.  $-0.36$  for China and  $-0.03$  for Australia).

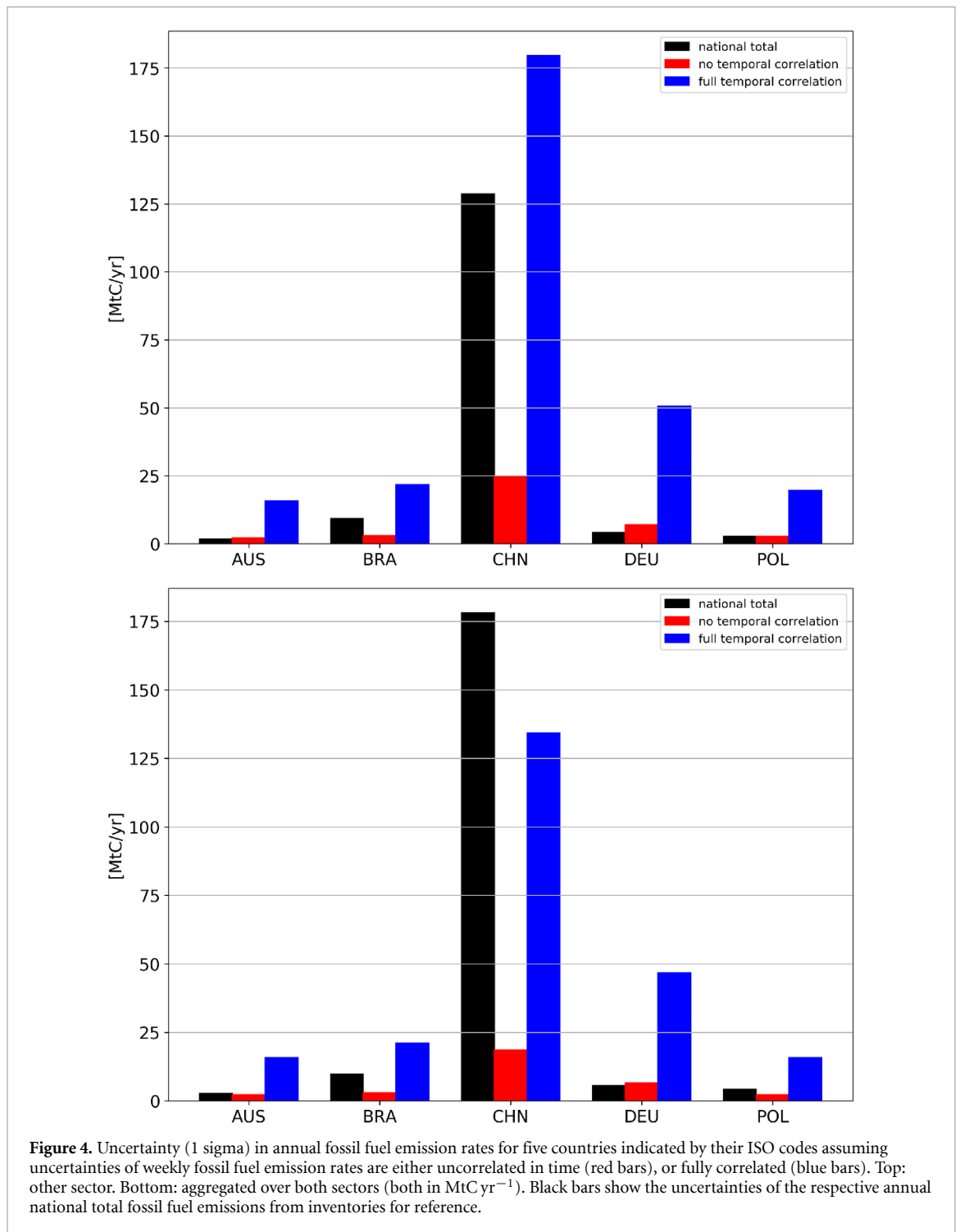
Increasing the number of satellites in the constellation will yield a significant increase in spatiotemporal coverage (third row of figure 2) which in turn will lead to an improved MVS capacity. This added value of the extra satellites is more pronounced for the other sector, where the uncertainty over Brazil and China is reduced by about 30% when going from a single satellite to the four-satellite constellation (scenarios 3 and 4). We note that the exact performance gain for each additional satellite depends on the individual orbits in the week we analyse. This is because atmospheric transport and aspects that influence the retrieval (including the quality filtering), such as cloud cover or aerosol load vary in time (see section 4 of the supplementary text for construction of XCO<sub>2</sub> data sets). The glint mode provides XCO<sub>2</sub> over the oceans and coastal land regions in addition to land areas, constituting an important area for a MVS capacity, because many power plants and cities are located along the coast. Adding those to the observations over land (scenario 5) results in some gain in performance, especially for the other sector of China, which is only partially covered by observations over land (figure 2, third row).

Our standard assumptions for XCO<sub>2</sub> data uncertainty (see section 4 of the supplementary text) have deliberately ignored the effect of model error and in particular of the error we make when comparing the average of potentially inhomogeneously distributed small-scale observations with the XCO<sub>2</sub> in our grid cell (representation error (Heimann and Kaminski 1999, Pillai *et al* 2010)). Scenario 6 explores the effect of including an estimate of representation error (based on a high-resolution transport simulation (Agustí-Panareda *et al* 2019), see section 4 of the supplementary text) in the uncertainty specification of our XCO<sub>2</sub> product. As the representation error will decrease when resolution of the transport model is increased, this assessment cannot be generalised to a CCFFDAS that uses a transport model with finer resolution (resolving in the ideal case the scales of the XCO<sub>2</sub> observations, i.e. 4 km<sup>2</sup>). The effect of representation error varies between countries and depends on aspects that vary in time such as atmospheric transport, small-scale variability in the high-resolution XCO<sub>2</sub> simulation, and factors that

determine the systematic retrieval errors and quality filters, such as clouds or aerosol load. We therefore find that representation error will be important in a CCFFDAS, but also acknowledge the potential to decrease it using higher resolution transport models and better characterisation of observation conditions.

Our next scenario (# 7) addresses the quantification of changes in the long-term emission trends based on the default scenario, i.e. with one CO<sub>2</sub>M satellite. For this purpose, we define a second analysis period in the following year, i.e. the first seven days of June 2009. Between this and the current analysis period, we would expect no or little changes in several factors that determine systematic sampling errors in CO<sub>2</sub> observations (see section 4 of the supplementary text), and therefore systematic errors should be very similar between the two periods. As for our default scenario (# 3), we assume no model error. For our emission change scenario we also neglect the systematic component of the retrieval error and only use the random component of the retrieval error to compute the random error in the difference of the CO<sub>2</sub> observations between the two periods (see section 4 of the supplementary text). This is certainly an over-optimistic choice. For the other sectors of Australia, China, or Germany the system yields posterior uncertainties in emission changes that are between 15% and 25% lower than the posterior uncertainties in absolute emissions (scenario 3). For the other sectors of Brazil and Poland the posterior uncertainties in emission changes are slightly larger (between 2% and 6%) than the posterior uncertainties in absolute emissions.

Our next scenario (scenario 8) is intended to demonstrate and quantify the potential for simultaneous use of inventories and CO<sub>2</sub>M in a CCFFDAS. As the prototype is not yet set up for the annual scale (see section 2), where inventory information is routinely available, we construct a hypothetical scenario on the weekly time scale, in which we assume availability of weekly inventories with the same relative uncertainty than on the annual scale (see section 2 of the supplementary text). We note, however, ongoing work on sub-annual inventories (Andres *et al* 2011, 2016, Oda *et al* 2018). The addition of this inventory information as a further observational data stream to the default scenario would result in a strong reduction in posterior uncertainty (scenario 8) for both sectors, which indicates a strong potential for synergy or complementarity between atmospheric (and other) observations and the inventories in a CCFFDAS. Combining the hypothetical inventory information with the larger of the two surface networks (scenario 9), instead of a single CO<sub>2</sub>M satellite (scenario 8), reduces the performance for the other sectors of Brazil ( $\approx 10\%$ ) and China ( $\approx 20\%$ ) and the electricity generation sector of China ( $\approx 10\%$ ). The specification of larger uncertainties on the hypothetical



weekly inventories would have resulted in larger posterior uncertainty. It would have also resulted in larger differences in posterior uncertainty between the two scenarios that use the inventories (scenarios 8 and 9). The larger the assumed uncertainty of the hypothetical weekly inventories the more we approach the respective posterior uncertainties of the scenarios that do not make use of the sectoral national totals derived by inventories (scenarios 2 and 3).

An operational CCFDAS would obtain fossil fuel emission estimates for all 52 weeks of a given year,

allowing the generation of annual fossil fuel emissions data. We estimate the uncertainty of annual fossil fuel emissions assuming all weekly emissions uncertainties to be equal and uncorrelated. Results for the default scenario are shown in figure 4, and compared to the uncertainties of the corresponding inventories (see section 2 of the supplementary text). For the other sector, the resulting posterior uncertainty is similar to the uncertainty in the national total for Australia and Poland, somewhat higher for Germany, and considerably lower for Brazil and China.

The assumption of uncorrelated weekly uncertainties is certainly optimistic, as there may be systematic errors in the prior or the retrievals that persist for longer than a week. To explore the effect of temporal uncertainty correlation we repeated the calculation for a case where the uncertainty of the weekly estimates has a temporal correlation coefficient of 1. We stress that this is an extreme and unrealistic case, as it assumes all errors are systematic and that repeated XCO<sub>2</sub> measurements throughout the year will provide no additional information. The results should thus be regarded as an upper bound. Overall the CCFDAS can provide an estimate of the national totals that is independent from the inventories and exhibits uncertainties of the same order of magnitude. We also computed the uncertainty in annual national emission rates aggregated over both sectors. As expected from the existence of a negative cross-sector uncertainty correlation in particular for Australia and China, the CCFDAS generally performs better for aggregated totals.

#### 4. Conclusions

We have analysed two modes of operation of a CCFDAS using observations from the CO2M mission. The first mode of operation is independent of national inventories (verification mode). Our analysis shows that, at country and annual scales, this mode can provide uncertainty ranges in sectorial fossil fuel emission rates of the same order of magnitude as the uncertainties in the national inventories. Consequently, comparison of inventories with CCFDAS estimates has the potential to detect and resolve inconsistencies between the two data sources.

The second mode of operation is the direct inclusion of sectoral national totals derived by inventories as a further ‘observational’ data stream to be assimilated into the CCFDAS. Compared to the verification mode, this synergistic mode results in a considerable reduction of uncertainty ranges in estimated emission rates through the inventory information. Or, taking the inventory perspective, the inventory-derived national emission totals are improved through assimilation of atmospheric (and other) observations and their uncertainty reduced. Due to limitations imposed on the current prototype by the resolution of its atmospheric transport component, our analysis focused on large countries. This approach is, however, equally applicable to smaller countries and at the sub-country level when using a high-resolution transport model. In the ideal case, the transport model would have the same resolution as the sensor, i.e. 4 km<sup>2</sup>, and the system could take full advantage of the CO2M XCO<sub>2</sub> imaging capabilities. This will allow us to obtain information of localised emission sources such as power plants via observed emission plumes. Kuhlmann *et al*

(2021) have demonstrated the potential of using the CO2M observations at sensor resolution in a simplified mass-balance approach. Furthermore, our analysis indicates scope for improving the performance of an MVS capacity, by increasing the number of satellites in the CO2M constellation and the use of XCO<sub>2</sub> observations over the ocean. Each of these additional observational data streams contributes to a further reduction of the posterior uncertainty ranges in sectorial fossil fuel emission rates. This holds for the synergistic mode of operation as well as for the validation mode. CO2M is just one of several constellations of satellites either flying or planned. Combinations of CO2M with other CO<sub>2</sub> missions Crisp *et al* (2018), virtual constellation, see, e.g. could, provided good calibration was maintained, further reduce this uncertainty. We expect the effect of increased spatial coverage to be more pronounced in a system with a high-resolution transport component.

On the basis of our analysis, we expect that the MVS will quantify emission changes more accurately than absolute emissions, owing to recurrence of several sources of systematic errors. Planned aerosol and cloud measurements from the same platform will further reduce the magnitude of such systematic errors (ESA 2018). We have deliberately ignored the effect of potential errors in our modelling chain, to allow generalisation of our results beyond our prototype CCFDAS. As such model errors will degrade the performance of an MVS capacity, the development and validation of high-quality process representations of biogenic and fossil fuel emissions and atmospheric transport is of paramount importance for an operational MVS capacity. Residual model errors can be systematically assessed in comparisons of component models or comparisons with alternative inversion approaches. Equally important are additional observational data streams (both *in-situ* and remotely sensed), which can be assimilated into the MVS component models as further constraints. Any observation that can further constrain a fossil emission sector or a natural flux component will, through the atmospheric constraint on the integrated signal, also reduce uncertainties on the other components. Potential for improvement of the fossil fuel emission models lies in further sectorial differentiation, e.g. residential (McNorton *et al* 2021) or transport, where sector specific information (e.g. local climate or national and local scale socioeconomic data) and atmospheric measurements of NO<sub>2</sub> (Reuter *et al* 2014, 2019, Konovalov *et al* 2016), CO (Rayner *et al* 2014, Konovalov *et al* 2016) or radiocarbon (Levin *et al* 2003, Rayner *et al* 2010) can be utilised as additional constraints. Examples of data streams suitable for further constraining biogenic fluxes in a CCFDAS (and thus indirectly the fossil fuel emissions) are additional remotely sensed variables (Scholze *et al* 2017) such as Sun-induced fluorescence (Norton *et al*

2018) or biomass (Thum *et al* 2017, Quegan *et al* 2019), or *in-situ* observations such as eddy covariance measurements (Williams *et al* 2005).

Here we have investigated the constraint provided by XCO<sub>2</sub> observations over a week on emissions over that week. Provided that the required input data streams are available, (preliminary) emission estimates could in principle be provided shortly after the end of each analysis period, e.g. to monitor progress towards local reduction targets. We expect, however, a performance gain from an extended data assimilation setup that uses observations for a considerably longer assimilation window to infer the temporal variations in the fossil fuel emissions within this assimilation window. Such a setup would be capable of exploiting the sensitivity of atmospheric observations made in one week to emissions in the preceding weeks to better constrain such emissions by also including longer-range constraints.

### Data availability statement

The L3e files (random and systematic XCO<sub>2</sub> errors) are available on request from Michael Buchwitz (<https://Michael.Buchwitz@iup.physik.uni-bremen.de>). The CHE project tier 1 nature run data are available on request from Anna Agustí-Panareda (<https://Anna.Agusti-Panareda@ecmwf.int>). All other data supporting the findings of this study are available within the paper.

The data that support the findings of this study are available upon reasonable request from the authors.

### Acknowledgment

This study was funded by the European Space Agency under Contract No. 4000122014 and through the European Union's Horizon 2020 research and innovation programme under Grant Agreements No. 776186 (the CHE project) and No. 958927 (the CoCO<sub>2</sub> project). We thank Christiane Textor for valuable comments on the manuscript and Nadine Gobron and Monica Robustelli for providing full resolution JRC-TIP products. We thank Bernard Pinty for his enthusiastic support of this study which reflected his strong commitment to promote spaceborne CO<sub>2</sub> measurements. We thank Gregoire Broquet and Philippe Ciais for helpful discussions on the study. M S and W K acknowledge support from the three Swedish strategic research areas Modelling the Regional and Global Earth system (MERGE), the e-science collaboration (eSENCE), and Biodiversity and Ecosystems in a Changing Climate (BECC).

### Author contributions

T K, M S, P J R conceived the study. M B and M R prepared XCO<sub>2</sub> error data set. A A-P provided simulated high-resolution XCO<sub>2</sub> data set. M V prepared further

input data sets. T K conducted experiments and prepared first draft of manuscript. All authors contributed to discussion and presentation of results.

### Code availability

The core CCFDAS code is not publicly available. Python scripts for the processing of the L3e files (figure 2) and of the posterior uncertainties (figures 3 and 4) are available from the corresponding author upon reasonable request. The FFDAS code is available from Peter Rayner (<https://prayner@unimelb.edu.au>) upon reasonable request. The TM3 model code is available from Martin Heimann (<https://Martin.Heimann@bgc-jena.mpg.de>) or Christian Rödenbeck (<https://croeden@bgc-jena.mpg.de>) upon reasonable request.

### ORCID iD

Hans Chen  <https://orcid.org/0000-0002-8601-6024>

### References

- Agustí-Panareda A *et al* 2019 Modelling CO<sub>2</sub> weather—why horizontal resolution matters *Atmos. Chem. Phys.* **19** 7347–76
- Andres R J, Boden T A and Higdon D M 2016 Gridded uncertainty in fossil fuel carbon dioxide emission maps, a CDIAC example *Atmos. Chem. Phys.* **16** 14979–95
- Andres R J, Gregg J S, Losey L, Marland G and Boden T A 2011 Monthly, global emissions of carbon dioxide from fossil fuel consumption *Tellus B* **63** 309–27
- Asefi-Najafabady S, Rayner P J, Gurney K R, McRobert A, Song Y, Coltin K, Huang J, Elvidge C and Baugh K 2014 A multiyear, global gridded fossil fuel CO<sub>2</sub> emission data product: evaluation and analysis of results *J. Geophys. Res. Atmos.* **119** 10213–31
- Buchwitz M *et al* 2013 Carbon Monitoring Satellite (CarbonSat): assessment of atmospheric CO<sub>2</sub> and CH<sub>4</sub> retrieval errors by error parameterization *Atmos. Meas. Tech.* **6** 3477–500
- Calvo Buendia E *et al* 2019 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (available at: [www.ipcc-nggip.iges.or.jp/public/2019rfl/index.html](http://www.ipcc-nggip.iges.or.jp/public/2019rfl/index.html))
- Ciais P, Crisp D, van der Gon H, Engelen R, Heimann M, Janssens-Maenhout G, Rayner P and Scholze M 2015 Towards a European operational observing system to monitor fossil CO<sub>2</sub> emissions—final report from the expert group (European Commission, Copernicus Climate Change Service) (available at: [www.copernicus.eu/sites/default/files/library/CO2\\_Report237\\_22Oct2015.pdf](http://www.copernicus.eu/sites/default/files/library/CO2_Report237_22Oct2015.pdf))
- Crisp D *et al* 2018 A constellation architecture for monitoring carbon dioxide and methane from space *CEOS Atmospheric Constellation Greenhouse Gas Team, Version 1.0*
- ESA 2018 *Copernicus CO<sub>2</sub> Monitoring Mission Requirements Document* (European Space Agency)
- Hascoët L and Pascual V 2013 The Tapenade automatic differentiation tool: principles, model and specification *ACM Trans. Math. Softw.* **39** 1–43
- Heimann M and Kaminski T 1999 Inverse modeling approaches to infer surface trace gas fluxes from observed atmospheric mixing ratios *Approaches to Scaling of Trace Gas Fluxes in Ecosystems* ed A F Bouwman (Amsterdam: Elsevier) ch 14, pp 275–95
- Heimann M and Körner S 2003 The global atmospheric tracer model TM3 *Technical Report 5* (Jena, Germany: Max-Planck-Institut für Biogeochemie)

- International Energy Agency 2011 CO<sub>2</sub> emissions from fuel combustion *OECD/IEA, Paris, 1997* (available at: [www.iea.org/](http://www.iea.org/))
- IPCC 2006 2006 *IPCC Guidelines for National Greenhouse Gas Inventories* (Intergovernmental Panel on Climate Change) (available at: [www.ipcc-nggip.iges.or.jp/public/2006gl/index.html](http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.html))
- Janssens-Maenhout G *et al* 2020 Towards an operational anthropogenic CO<sub>2</sub> emissions monitoring and verification support capacity *Bull. Am. Meteorol. Soc.* **101** E1439–51
- Kaminski T, Knorr W, Rayner P and Heimann M 2002 Assimilating atmospheric data into a terrestrial biosphere model: a case study of the seasonal cycle *Glob. Biogeochem. Cycles* **16** 14–1–16
- Kaminski T and Mathieu P-P 2017 Reviews and syntheses: flying the satellite into your model: on the role of observation operators in constraining models of the Earth system and the carbon cycle *Biogeosciences* **14** 2343–57
- Kaminski T and Rayner P J 2017 Reviews and syntheses: guiding the evolution of the observing system for the carbon cycle through quantitative network design *Biogeosciences* **14** 4755–66
- Kaminski T, Scholze M, Voßbeck M, Knorr W, Buchwitz M and Reuter M 2017 Constraining a terrestrial biosphere model with remotely sensed atmospheric carbon dioxide *Remote Sens. Environ.* **203** 109–24
- Knorr W and Heimann M 1995 Impact of drought stress and other factors on seasonal land biosphere CO<sub>2</sub> exchange studied through an atmospheric tracer transport model *Tellus B* **47** 471–89
- Konovalov I B, Berezin E V, Ciais P, Broquet G, Zhuravlev R V and Janssens-Maenhout G 2016 Estimation of fossil-fuel CO<sub>2</sub> emissions using satellite measurements of “proxy” species *Atmos. Chem. Phys.* **16** 13509–40
- Kuhlmann G, Henne S, Meijer Y and Brunner D 2021 Quantifying CO<sub>2</sub> emissions of power plants with CO<sub>2</sub> and NO<sub>2</sub> imaging satellites *Front. Remote Sens.* **2** 689838
- Lavergne T, Voßbeck M, Pinty B, Kaminski T and Giering R 2006 Evaluation of the two-stream inversion package, EUR 22467 EN (European Commission—DG Joint Research Centre, Institute for Environment and Sustainability)
- Levin I, Kromer B, Schmidt M and Sartorius H 2003 A novel approach for independent budgeting of fossil fuel CO<sub>2</sub> over Europe by <sup>14</sup>C observations *Geophys. Res. Lett.* **30** 2194
- McNorton J R, Arduini G, Bousseret N, Agustí-Panareda A, Balsamo G, Boussetta S, Choulga M, Hadade I and Hogan R J 2021 An urban scheme for the ECMWF integrated forecasting system: single-column and global offline application *J. Adv. Model. Earth Syst.* **13** e2020MS002375
- Norton A J, Rayner P J, Koffi E N and Scholze M 2018 Assimilating solar-induced chlorophyll fluorescence into the terrestrial biosphere model BETHY-SCOPE v1.0: model description and information content *Geosci. Model Dev.* **11** 1517–36
- Oda T, Maksyutov S and Andres R J 2018 The open-source data inventory for anthropogenic CO<sub>2</sub>, version 2016 (ODIAC2016): a global monthly fossil fuel CO<sub>2</sub> gridded emissions data product for tracer transport simulations and surface flux inversions *Earth Syst. Sci. Data* **10** 87–107
- Petrescu A M R *et al* 2021 The consolidated European synthesis of CO<sub>2</sub> emissions and removals for the European Union and United Kingdom: 1990–2018 *Earth Syst. Sci. Data* **13** 2363–406
- Peylin P, Rayner P J, Bousquet P, Carouge C, Hourdin F, Heinrich P and Ciais P (AEROCARB Contributors) 2005 Daily CO<sub>2</sub> flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology *Atmos. Chem. Phys.* **5** 3173–86
- Pillai D, Gerbig C, Marshall J, Ahmadov R, Kretschmer R, Koch T and Karstens U 2010 High resolution modeling of CO<sub>2</sub> over Europe: implications for representation errors of satellite retrievals *Atmos. Chem. Phys.* **10** 83–94
- Pinty B *et al* 2017 An operational anthropogenic CO<sub>2</sub> emissions monitoring & verification support capacity—baseline requirements, model components and functional architecture, EUR 28736 EN (European Commission Joint Research Centre) (<https://doi.org/10.2760/08644>)
- Quegan S *et al* 2019 The European space agency BIOMASS mission: measuring forest above-ground biomass from space *Remote Sens. Environ.* **227** 44–60
- Rayner P J, Raupach M R, Paget M, Peylin P and Koffi E 2010 A new global gridded data set of CO<sub>2</sub> emissions from fossil fuel combustion: methodology and evaluation *J. Geophys. Res. Atmos.* **115** D19306
- Rayner P J, Utembe S R and Crowell S 2014 Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study *Atmos. Meas. Tech.* **7** 3285–93
- Reuter M, Buchwitz M, Hilboll A, Richter A, Schneising O, Hilker M, Heymann J, Bovensmann H and Burrows J P 2014 Decreasing emissions of NO<sub>x</sub> relative to CO<sub>2</sub> in East Asia inferred from satellite observations *Nat. Geosci.* **7** 792–5
- Reuter M, Buchwitz M, Schneising O, Krautwurst S, O’Dell C W, Richter A, Bovensmann H and Burrows J P 2019 Towards monitoring localized CO<sub>2</sub> emissions from space: co-located regional CO<sub>2</sub> and NO<sub>2</sub> enhancements observed by the OCO-2 and S5P satellites *Atmos. Chem. Phys.* **19** 9371–83
- Scholze M, Buchwitz M, Dorigo W, Guanter L and Quegan S 2017 Reviews and syntheses: systematic earth observations for use in terrestrial carbon cycle data assimilation systems *Biogeosciences* **14** 3401–29
- Thum T, MacBean N, Peylin P, Bacour C, Santaren D, Longdoz B, Loustau D and Ciais P 2017 The potential benefit of using forest biomass data in addition to carbon and water flux measurements to constrain ecosystem model parameters: case studies at two temperate forest sites *Agric. For. Meteorol.* **234–5** 48–65
- UNFCCC 2015 *Conference of the Parties (COP), Adoption of the Paris Agreement. Proposal by the President* (available at: <http://unfccc.int/resource/docs/2015/cop21/eng/l09.pdf>)
- Wang Y *et al* 2020 PMIF v1.0: assessing the potential of satellite observations to constrain CO<sub>2</sub> emissions from large cities and point sources over the globe using synthetic data *Geosci. Model Dev.* **13** 5813–31
- Williams M, Schwarz P A, Law B E, Irvine J and Kurpius M R 2005 An improved analysis of forest carbon dynamics using data assimilation *Glob. Change Biol.* **11** 89–105