

# **Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020**

## **Reply to reviewer 2**

We are very grateful for the insightful comments from the reviewer to improve the documentation of the CAMS greenhouse gas reanalysis for potential users. We have addressed the comments of the reviewer (in bold) to clarify the different aspects of the reanalysis including the selection/quality of the assimilated observations, the different model components and the characteristics of the prescribed surface fluxes.

**Review of "The CAMS greenhouse gas reanalysis from 2003 to 2020" by Agusti-Panareda et al.**

**This article presents a description of a new reanalysis dataset of greenhouse gases (GHGs). This is the first GHG reanalysis dataset ever produced using data assimilation techniques that adjust GHG concentrations rather than GHG fluxes. As such, it is a formidable achievement but at the same time, there is considerable room for improvement (as noted by the authors in sections 4 and 5). Reanalyses are typically improved after feedback from the user community and/or with improved data assimilation techniques or increased numbers of observations. Thus an important aspect of reanalysis production is the dissemination and documentation of the products to encourage and inform the potential user community. That is the main purpose of this article. As such, the comments below are primarily aimed at improving the description of the product to potential users.**

### **Specific comments**

- 1. Section 2.2 and Table 1: It would be useful to know how the satellite instruments were selected. For example, OCO-2 was not used. In section 5, it is mentioned that it will be used in future versions, but it is useful for the user to understand the rationale behind the selection of instruments for this product.**

The main rationale for the selection of the data was the operational availability of NRT data, as the re-analysis is foreseen to catch up and run in near-real time eventually. Currently there is no near-real time product of the OCO-2 data and that is why OCO-2 has not been used in the current CAMS near-real time products (including re-analysis). We are working with LSCE to have a near-real time OCO-2 product based on a neural network retrieval (Breon et al., 2022; <https://amt.copernicus.org/articles/15/5219/2022/amt-15-5219-2022-discussion.html>). This will allow us to use OCO-2 in the next CAMS re-analysis. The rationale for the observation data selection will be clarified in the revised manuscript.

- 2. Line 99 and section 2.3: Figure 2 shows 2-way interaction between the forecast and surface fluxes. Please explain how the forecast influences the surface fluxes in section 2.3 or line 99.**

The two-way arrows indicate the two-way coupling between the atmospheric forecast and the biogenic surface fluxes. The surface fluxes are affected by the forecast of temperature, humidity, radiation and soil moisture and the atmospheric CO<sub>2</sub> in the forecast is affected by the surface biogenic fluxes. This will be clarified in the revised version of the manuscript.

- 3. Section 2.3: There is no mention of Figure 3 in the text. It is useful to understand why this Figure is presented, and the main message behind it.**

Figure 3 shows the seasonal, inter-annual variability and trend of the surface fluxes. This is important because one of the caveats of the re-analysis is that CH<sub>4</sub> emissions are kept fixed from 2010 onwards, while for CO<sub>2</sub> anthropogenic fluxes we apply an extrapolation, and the biogenic fluxes are modelled and therefore have an inter-annual variability. In the revised manuscript, Figure 3 is mentioned and referred to when explaining the differences between different fluxes in terms of inter-annual variability and seasonal cycles.

- 4. Line 184: What is the rationale behind the choice of EDGAR versus higher spatial resolution datasets for anthropogenic emissions?**

EDGAR produces global emissions for both CO<sub>2</sub> and CH<sub>4</sub> at a relatively high resolution of 0.1 degrees (compared to 80km resolution of the CAMS re-analysis). The problem with EDGAR is the extension to NRT and the caveats associated with large inter-annual variability during the COVID period. This will be clarified in the revised manuscript.

- 5. Line 217: Typo: TI255 should be TL255 presumably.**

This will be corrected in the revised manuscript.

- 6. Line 302: I do not see the 20 ppm error in Figure 5. If the vertical scale in Fig. 5c (bottom panel) is linear then the green and red dashed curves seem to overlap after 2019, meaning errors of less than 5 ppm.**

Yes, indeed. There is a typo here, 20ppm should be 2ppm. The sentence “accompanied by a large standard deviation error (~20ppm, cf. dashed lines in same panel and figure)” has been removed as the degradation associated with IASI is mainly detected from the large step change in the averages of the departures.

7. Lines 368-370: Taylor diagrams of Figs. 8b, 11b. What is the normalization used on the radial axes? Presumably it is the observed standard deviation for a given site. Why was this normalization needed? Standard Taylor diagrams do not do a normalization. Presumably it allows for better comparison among sites with very different observed variability.

The reviewer is correct. Assuming we have 20 sites, then there are 20 time series for the model and the observed NDACC ground based data (gb), each with its own variability. In order to plot them together, the standard deviation of the model time series is used to normalize the timeseries (i.e for a site, the time series model or gb is divided with the std of the model time series ): that is why the model time series is always at 1 for each site. The correlations are independent of such a normalization and from the location of a site in the Taylor plot you can deduce if the model has higher variability (the site is plotted with a distance <1 to the origin) or lower (the site is plotted at a distance >1 from the origin) compared to the variability in the gb data. This will be clarified in the revised manuscript.

8. Lines 457-8 and Fig. 13: How was the site for Fig. 13 chosen? Is it a typical example, or a good example? It is nice to see that the overall structure of the profile (esp. for CO<sub>2</sub>) is well captured in boreal summer in France. It would be interesting to see the comparison in the southern hemisphere. Are the general biases of Fig. 12 more evident in New Zealand?

The site of Traînou (France) has been used as a good example of the synoptic variability during the summer, more than the systematic errors. As the reviewer suggested, we have plotted the only profile available from Lauder (New Zealand) (see Figure S1 in supplement) which shows much larger errors in September, particularly near the surface and in the stratosphere. In the revised manuscript we will add an extra figure as a Supplement to illustrate the errors at different sites across different latitude bands.

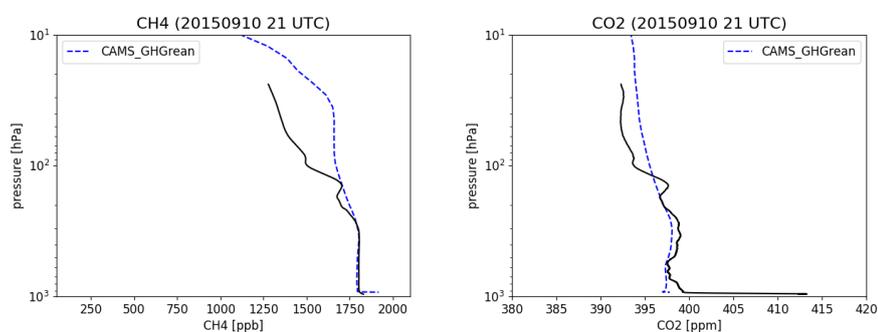


Figure S1. Vertical mole fraction profiles of CO<sub>2</sub> and CH<sub>4</sub> from the CAMS GHG reanalysis (dash line) and AirCore observations (solid line) at Lauder (New Zealand, see Table A3) over the period in June 2019.

**9. Line 496: Since this is a reanalysis, why were the in situ data not assimilated? Would they be enough to better constrain the global growth rate? Will this be done in the future?**

The CAMS reanalysis and generally the reanalysis at ECMWF aim to eventually run in near-real time. Since the availability of quality controlled in situ data is not generally available in near real time and we use a global model with a short window data assimilation (i.e. 12 hours) that adjusts only the GHG concentrations (i.e. not the fluxes), the impact of the sparse in situ data would be small as the increments are very localised in time (i.e at the beginning of the data assimilation window) and space (around the station near the surface). Preliminary tests showed that the increments around the in situ stations are dispersed by the atmospheric transport in less than 12 hours. In the future, we will explore the possibility of using in situ data with the inversion capability currently being developed in the CoCO2 project. The flux adjustment will result in a longer-lived impact of the impact of the in situ observations, as we expect the resulting atmospheric CO<sub>2</sub> and CH<sub>4</sub> corrections will not be as localised in time and space, particularly when using a longer data assimilation window with the inversion capability. All these aspects related to the future potential assimilation of in situ data will be clarified in the revised version of the manuscript.