Interactive comment on "Evaluation of the CAMS global atmospheric trace gas reanalysis 2003–2016 using aircraft campaign observations" *by* Yuting Wang et al.

Yuting Wang et al.

guy.brasseur@mpimet.mpg.de

We thank the referee#2 for taking the time to read the manuscript and offer helpful comments and suggestions. The referee's comment is repeated with our response in bold. Responses to those comments are listed below: 1. This paper presents an evaluation of several CAMS reanalysis products against a suite of aircraft campaign data for multiple chemical species. Evaluation of reanalyses against independent data is an important activity and the evaluation here is rigorous. While the manuscript does a nice job of evaluating the performance of the CAMS reanalyses, there is only limited explanation or analysis of the causes of mismatches between CAMS and observa-

tions, or the reasons why the newer reanalysis outperforms the earlier reanalyses in some regions. An interesting result of this paper is that for some species and regions of the atmosphere, the reanalysis has only a little improvement, or even weaker performance, than the control simulation. It would be helpful to have more analysis of why this is the case. Overall, the paper would be strengthened by a more detailed exploration of the underlying causes of the biases, as this could provide guidance for future improvements and provide greater scientific insight. Response: We agree with the referee that including further interpretation on the results will give more insight on the model development. We have added more explanations to improve the manuscript as the referee recommended, and in particular we added in the paper a new Section of the concentration ratios of chemically interacting species, so that we can check to what extend the photochemical theory is verified. However, the main point of this paper is to include additional measurements to the routine ones that used by Inness et al. (2019) and Wagner et al. (2019) for the evaluation of the new CAMS reanalysis. The aircraft campaigns provide simultaneous profile measurements of many species such as OH and HO2, which is not within the routine measurements. 2. Line 56: What other species are assimilated? Response: The species that are assimilated are listed in Table 2. 3. Line 57: Which specific satellite observations are assimilated, and how does the assimilation system account for the vertical sensitivities of different satellites? The coarse vertical resolution of satellite data compared to aircraft campaign data is a likely cause of some of the biases against aircraft observations, so this should be discussed in some detail. Response: The satellite used in the assimilations are listed in Table 2. More details for the assimilation have been added in the model description section. 4. Line 88: What does "consistent in time resolution" mean? Response: As shown in Table 1, the time resolution for the analysis fields is six hours for MACCRA and CIRA and three hours for CAMSRA. For the forecast fields, the time resolution is three hours for all the reanalysis versions. So, we used the forecast fields for consistency in the time resolution between the 3 reanalyses. This sentence has been added in the manuscript to make it clear. 5. Line 186: Please define "good". Some rather large biases are

mentioned later in the paragraph. Response: We agree that this expression is not accurate, so we deleted this sentence. 6. Lines 210-212: Is that difference in r2 statistically significant? Also, wouldn't we expect a larger improvement since ozone is being assimilated? Is the limited improvement due to limitations or uncertainties in the observations, or something else? Response: All comparisons show a statistically significant relation between observations and simulations ($p<0.001$). Therefore, although the difference in r2 only in the range of 0.01-0.05, the results are statistically significant. The simulation is improved to a certain degree with assimilation, for example, r2 and slope increase about 0.05 and 0.1 respectively and RMSE decreases above 4 ppb for O3 after assimilation. However, in the case of tropospheric ozone, the influence of the assimilation is weaker than in the case of stratospheric ozone. 7. Line 264: Similar to the comment above, why does the NO2 agreement not improve when NO2 is assimilated? Response: The impact of the assimilation of tropospheric NO2 column retrievals is small because of the short lifetime of NO2. Even though the assimilation lead to large analysis increments this information was not retained by the model, and most of the impact of the data assimilation was lost from one analysis cycle to the next (Inness et al., 2015). We added this explanation to the manuscript. 8. Line 282: Are the differences small because the species are well buffered against changes in O3, NOx, etc., or because the assimilation doesn't change the O3 and NOx concentrations very much? Response: We would think both are the possible reasons. To further analyze the impact of the assimilations of ozone, CO and NO on the other species, more experiment needs to be run, which is out of the scope of this paper. 9. Line 302: The larger biases of CAMSRA with altitude seems like a surprising result since satellite observations of ozone are available in the stratosphere and upper troposphere but not the boundary layer. It would be nice to relate the discussion of the vertical profiles to a description of where observations are available to constrain the reanalysis. Response: This feature is only shown at Hawaii and Mexico, but not the case at Bangor and the Arctic, so we cannot make the universal conclusion that the biases of CAMSRA increase with altitude. The results show that the model's performance varies from region to region.

C3

The information of the assimilated satellite for different layers are listed in Table 2. 10. Line 358: What emission inventory is used? Does it have known biases, or is this a new finding? Response: The emission inventories are listed in Table 1. Their choice has been made by the CAMS project at ECMWF based on what was estimated as the best inventories available at that time. For example, the anthropogenic emissions are based on the MacCity inventory. No systematic bias is known for these inventories, but uncertainties exist. Emission inventories are constantly updated to address uncertainties and to account for changes in emissions. 11. Summary: Can you end with some directions for future improvements and/or a take-home message for the atmospheric chemistry community? Response: We added a sentence in the conclusions calling for an improvement in the adopted surface emissions of hydrocarbons and on the need to assimilate, if possible, some organic species in addition to formaldehyde. 12. Lines 32-33: It seems, then, that the reanalysis covers the period 2003-3018. Response: Yes, the reanalysis now covers 2003-2018, but we only evaluate the first release of the reanalysis, which is 2003-2016. 13. Line 52: misplaced comma Response: Corrected. 14. Lines 308-309: confusing sentence, please reword Response: Corrected. 15. Figs. 9-12: Please use thicker lines so they are easier to see. Response: We have modified the figures as suggested.

C4