

Response to Review #1

We thank reviewer #1 for the review, the insightful comments to the paper and for his/her endurance to read the long paper.

We will respond to the review point by point. The reviewer's comments are included in bold italics.

I think this is a standard paper describing a reanalysis product. It is likely to be useful to the scientific community; however, it is a bit difficult to read, partly owing to its length.

We appreciate the reviewer's concern that the paper is long. However, we prefer to present all considered species in one paper because they were produced by the CAMS system in one combined assimilation experiment.

The paper should be accepted for publication once the authors address a series of points, detailed in the specific comments and technical comments. They largely concern quantification and/or clarification of statements made.

Specific comments:

L. 32: Indicate why ozone at the surface cannot be improved by the assimilation.

The surface values could potentially be changed by the following processes : (i) direct addition of observation increments close to the surface, (ii) the impact of non-surface observations or the observation of other species by means of the model backward error co-variances and (iii) the downward transport of ozone from levels where the assimilation changed the ozone fields.

We think that the impact of any of these three factors was eliminated at the lowest model levels by the strength of the ozone dry deposition and titration with NO near the surface.

CAMSiRA did not assimilate any surface observations nor satellite retrievals for the lower troposphere. Only total columns and stratospheric profile data were assimilated. The background error co-variances calculated with the NMC method did not provide enough impact for strong non-local vertical influence, which would have led to an alteration at surface. Also, species-to-species back-ground error correlation were not implemented in the applied 4D-VAR method.

We added in the abstract:

“ ... because of the strong impact of surface processes such as dry deposition and titration with nitrogen monoxide (NO), which were both not changed by the assimilation. “

We add in the conclusion section when we discussed the reasons for the small influence (L1147)

“ ... nor that the vertical correlations in the model background errors were strong enough to cause a correction of the surface levels based on the levels above. “

L. 56: Quantify the “sufficient accuracy”.

We can not quantify this rather general statement but we changed it as follows:

“ .. with an accuracy sufficient to have an impact during the assimilation. “

L. 58: Provide details of the surface properties.

We clarified the statement by replacing “surface properties” with “surface albedo”.

L. 109: List the key species.

We changed the text as follows:

“... the aerosol variables and key chemical species such as ozone, HNO₃, N₂O₅, NO, NO₂, PAN and SO₂ only”.

L. 208: Why did you use scenario 8.5 instead of another one?

The scenario was chosen by the producers of the MACCITY data set (Granier et al. ,2011)

We added “... obtained in the MACCcity emissions ...”

L. 286-287: Did you use the averaging kernels for data other than MOPITT? Explain your choices.

We add

“For the ozone retrieval averaging kernels were not used because they were not provided or did not improve the analysis. For example, the high vertical resolution of the MLS ozone retrievals in the stratosphere made the use of AK not necessary. “

The L. 291: The data used are flagged “good” or not flagged “bad”?

Yes, this is the case. The retrieval data include a quality flag given by the providers

L. 375: Does the decrease in the burden indicate a positive result from the assimilation?

Yes, because there is a better agreement with the MOPITT retrievals.

L. 397: Explain in the text why you do not assimilate MOPITT observations over the Arctic.

Larger biases and errors in the retrievals occur at high latitudes because of low thermal contrast. It is a recommendation by the data providers not to assimilate data in high latitudes.

We added: “... because of the higher biases of the MOPITT data in this region.”

L. 451: Why is there only a little effect on the surface? Why are there no changes between CR and CAMSiRa from the assimilation?

See our response above

L. 471: Is it reasonable to calculate a linear trend? What assumptions do you make?

It is a valid to comment to question the underlying assumption (i.e. a linear trend) for any type of trend analysis. A detailed trend analysis is beyond the scope of the paper. However, the linearity of the trend seems not an unreasonable choice when looking at the graphs. Our focus is the comparison of trends of different data sets using a unified but simple approach.

We will add:

“... and, for reasons of simplicity, only the linear ...”

We also point out that linear trends are often expressed in units of %/yr in the paper. We concede that this unit is technically not consistent with a linear trend. We obtain the linear trend as

percentage by normalising the linear trend (e.g. Tg/year) with the average of the quantity over the whole period, i.e. all years.

We add at line 361

“The linear trend is as expressed as percentage with respect to the mean of the burden over the whole period.”

L. 516: Provide references for this statement.

We added the following reference:

(Eskes et al. , 2015).

Eskes, H., Huijnen, V., Arola, A., Benedictow, A., Blechschmidt, A.-M., Botek, E., Boucher, O., Bouarar, I., Chabrillat, S., Cuevas, E., Engelen, R., Flentje, H., Gaudel, A., Griesfeller, J., Jones, L., Kapsomenakis, J., Katragkou, E., Kinne, S., Langerock, B., Razinger, M., Richter, A., Schultz, M., Schulz, M., Sudarchikova, N., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Validation of reactive gases and aerosols in the MACC global analysis and forecast system, *Geosci. Model Dev.*, 8, 3523-3543, doi:10.5194/gmd-8-3523-2015, 2015.

L. 523: Is the comparison with MACCRA and CAMSiRA within the errors of these datasets?

Unfortunately, we cannot provide error estimates of the global burden of the two analysis sets.

L. 535: Why is there an exaggeration of the sea salt emission?

As pointed out in the supplement, sea salt emissions were close to the median of the AeroCom models. They were only at the high end of the values given in Boucher et al. (2015)

We amend the text as follows:

“The simulated sea salt emissions of C-IFS were within the reported range in the literature (see supplement). This suggests that the loss processes of sea salt were underestimated in C-IFS in comparison to other models.”

L. 594: Discuss why this seems unrealistic.

We add the following

“ ..., given that the global SO₂ emission are only less than 2% of the total aerosol emissions (see supplement)”

L. 662: Quantify the trends. Explain (or remind the reader) how you test for significance. Same for L. 751 and L. 757.

The significance of the linear trends was estimated at the 95 confidence interval. We now repeat this information in each section.

L. 674: Provide further details of the artificial accumulation of sulphate by the assimilation.

We added:

“The increase in sulphate was probably caused by underestimated loss processes for sulphate and SO₂ in the free and upper troposphere away from the emissions sources. The relative increase in

sulphate with respect to the other aerosol species could not be corrected by the assimilation of AOD.”

L. 767: Why is this remarkable? Because unexpected? Please avoid subjective comments.

We replace the statement with:

“Despite its simplicity, the Cariolle scheme in CR reproduced the...”

L. 852-865: What is the fidelity of the GOZCARDS dataset?

The standard error of the GOZCARDS data set is given as part of the data set. The values of the error are in the range of 10-20 ppb on the considered region, which is about 1%. However this error does not reflect biases. As we already mention in the text, the inter-comparison of different satellite retrievals by Tegtmeier et al. (2013) shows that MLS is biased low above 5 hPa (5-10%) and ACE-FTS is biased high above 10 hPa and biased low below 10 hPa with respect to the multi-instrument-mean. Since ACE-FTS contributes more to the GOZCARDS product in this region, we assume that the GOZCAD biases are controlled by the ACE-FTS biases.

We will quantify the biases in the text:

“ACE-FTS is biased high (5-10%) above 10 hPa and biased low (5-10%) below 10 hPa against the median of various retrievals. “

We corrected all technical comments: