Response to Reviewers of Manuscript

The authors have considerably improved their manuscript and many of my mentioned points have been addressed. However, some issues are still open from my point of view. Line and figure numbers refer to the tracked changes manuscript.

The authors are grateful to the reviewer for the time and thought clearly put into the review and comments regarding our paper. We have addressed all their comments below and including changes accordingly into our revised manuscript, which has led to substantial improvement in clarity and completeness. The original comments from the referees are in black and our responses in blue. Please see below our point-by-point replies to the comments of the reviewers.

Remaining major points:

1) I think it is crucial to connect the TCR-2 model results with measurements, as intended by the manuscript. However, this connection should be done in a way that can be easily understood by the reader. In Fig. 5 (and for O3 also 3,6,7), this is not done in a good way: IASI+GOME2 ozone from surface to 3 km is presented next to TCR-2 ozone at 850 hPa, and TCR-2 CO at 850 hPa is presented next to IASI total column CO.

Reply 1: Agreed and corrected. For presenting TCR-2 model simulations in consistency with measurements, we have modified figures 3, 5, 6 and 7 so as to always present both ozone and CO maps: (i) with satellite data and model simulations next to each other, respectively on the left- and right-hand side panels, (ii) with the same color scale ranges and units and also (iii) for the same atmospheric columns as derived from satellite (i.e. surface to 3 km of altitude (a.s.l.) partial columns in ppb for O_3 and total columns in molecules/cm^2 for CO).

2) Figures 3,5,6,7: Apples could be compared to apples, but it is not done here: From a model run, all necessary data to construct a quantity comparable to the measured quantity should be available. In particular, it is an easy task to calculate (partial) columns or average VMRs from multiple model levels. In the current state, the reader sees major differences between measured and simulated O3 and CO, but it is not clear if this is due to the different kind of quantity, or if the model simulates a different world than the world the measurements were taken in. If the latter was true, all following conclusions drawn on TCR-2 results would be highly questionable. So it is very important to show that the model agrees to the measurements first.

Reply 2: Agreed and done. As described in reply 1, we have followed the reviewer's suggestion. The revised manuscript presents satellite and TCR-2 simulation maps of ozone and carbon monoxide for the same atmospheric columns, units and scales.

3) Color bars of TCR-2 CO and IASI CO have different maximum VMRs, which is very misleading, once the quantities are in a comparable shape.

Reply 3: Agreed and corrected. The revised manuscript uses the same color bars for IASI and TCR-2 CO (fig. 5).

4) Usage of TCR-2 in this manuscript: I do not agree with the answer of the authors that the usage of TCR-2 is only "complementary" and does not need to match the IASI+GOME2 observations. I think that it is a strength of this study that observations and simulations are analyzed together. Pure model studies have the weakness that they potentially could describe a fictional world, so it is much better to also include measurements, as it is done here. But once

measurements are used, it needs to be assured that measurements and models are telling the same story, or explain the differences.

Reply 4: Agreed and done as requested. As recommended, we have calculated from TCR-2 the lowermost tropospheric ozone (surface–3 km) partial columns and total columns of CO to clearly show the differences with IASI+GOME2 ozone and IASI CO, in adjacent panels with respect to satellite data. In some cases, TCR-2 simulates similar features than the observations, in orders some differences are seen mostly in absolute terms. These last ones are mostly linked to uncertainties in the model and differences in the sensitivity of the satellite retrieval. Please find below explanations of these differences provided in the revised manuscript.

> Lines 431–437: "*However, those ozone plumes transported to the Iberian Peninsula are only partially depicted by IASI+GOME2 observations. These differences could be associated with misrepresentation of these ozone plumes in the simulations or lack of sensitivity (or spatial coverage) in the satellite data (particularly over the ocean). This last aspect is related to a reduction of the sensitivity of IASI+GOME2 over the ocean, due to smaller thermal contrasts than over land (Cuesta et al., 2013). The 7-day back-trajectories confirm that the air masses come from the west over the North Atlantic and offshore transport from Western Africa in the middle troposphere (Fig. 3f). It suggests that the link between those ozone plumes and emissions in North America is unclear.*"
>
> Lines 509–517: "*The tropospheric reanalysis also depicts a denser ozone plume over Northern Spain, Western France and the nearby Atlantic. This last one is not depicted by IASI+GOME2 at the LMT nor by in situ measurements at the surface. Ozone concentrations simulated at surface level (not shown) are rather moderate at this location, but higher over Eastern France, thus in better agreement with satellite data. At the upper troposphere (between 6 and 12 km of altitude, not shown), IASI+GOME2 does depict an ozone plume over Western France and located slightly north of the one simulated by TCR-2. This suggests uncertainties in the vertical and horizontal location of this lofted ozone plume in the model. A limitation of TCR-2 may come from the fact that the satellite ozone observations assimilated in TCR-2 are derived from TES measurements with a coarse horizontal resolution (only nadir pointing) and most sensitivity to ozone at the free troposphere at lowest.*",
>
> Lines 555–557: "*In this location, TCR-2 reanalyses depict rather weak increases in background amounts of CO between the east coast of England and Eastern Europe (Fig. 5d). This overestimation can be attributed to errors in the surface emission, chemical productions and losses, long-range transport from North America. On the other hand, the amount of CO over the Black Sea and the Eastern Mediterranean is comparable.*"
>
> Lines 617–625: "*On the other hand, TCR-2 simulates a temporal decrease of ozone concentrations for this plume. The differences between the model and the satellite data are probably associated with the representation of biomass burning emissions and its impact on the simulation of ozone production. Simulations of these processes and their quantification are currently challenging, as shown by the large discrepancies between different state-of-the-art global tropospheric ozone reanalysis products (e.g., Huijnen et al., 2020). Near-surface ozone concentrations are weakly and mainly indirectly constrained by the satellite observations used for assimilation in reanalysis such as TCR-2. In addition, only very few ground-based stations monitor air pollutants over Eastern Europe, as their geographical coverage is far from homogenous across the globe.*

Particularly in those locations, satellite observations of lowermost tropospheric ozone as those derived by IASI+GOME2 are particularly appreciated for filling the gap of ground-based observations."

Remaining minor points:
1) Figure 5: The sorting of the panels is inconsistent, since model data is on the right side in the first row, and on the left side in the second row.

Reply 5: Agreed and done. Sorting of the panels of Fig. 5 has been modified as requested (fig. 5c–d).

2) IASI CO total column is presented in "ppb". Does that mean that the total column is again divided by the total air column to get an average VMR?

Reply 6: Clarified and modified. For showing the original units of the satellite retrieval and compare it with model data, we have changed the unit of total columns of CO to "molec. cm^{-2}" (fig. 5c).

3) Regarding "Reply 11": I still do not understand, which data point of the gridded satellite data set is used for comparison with the in situ data. The nearest grid point? An interpolation of the four nearest data points?

Reply 7: Clarified. The statistics considers the nearest point of IASI+GOME2 1x1° grided data.

This is clarified in the RM as (line 135–136) "*Figure 1 shows the Pearson correlation coefficients (R) between daily in situ surface observations at each individual station (squares or circles) and the nearest grid point of IASI+GOME2 retrievals*."

4) Regarding "Reply 12": It should be mentioned, which kind of significance test was used, if such quantities as P-values are discussed.

Reply 8: Clarified.

This is clarified in the RM as (line 138) "*(with P-values of the Student's t test < 0.05)"*

5) Regarding "Reply 17": See major points 2 & 4

Reply 9: Please see reply 2 and 4.

6) Regarding "Reply 23": In my opinion, language is important here: If the authors want to say that the VMR is higher than the background, I would suggest to use formulations like "relatively high concentration", or "higher than background concentration". Otherwise: If there was a certain threshold used to identify the air mass, the threshold should be mentioned in the manuscript.

Reply 10: Agreed and clarified. We did not set a specific threshold to define ozone plumes in this study, therefore we have used the word "relatively high".

7) Regarding "Reply 27": It is nice that the authors now marked the altitude range, which was shown in panel a. However, I rather meant to mark the horizontal overlap of panels b and c, instead. Still, this is not very important in my view and was only a suggestion to guide the reader's eyes.

Reply 11: Done. We have marked the positions of intersection in the Fig. 9b-c.