Author's comments in reply to the anonymous referee for "An updated tropospheric chemistry reanalysis and emission estimates, TCR-2, for 2005–2018" by K. Miyazaki et al.

We want to thank the referees for the helpful comments. We have revised the manuscript according to the comments, and hope that the revised version is now suitable for publication. Below are the referee comments in italics with our replies in normal font.

Reply to Referee #1

General comments 1. Section 2.2. The covariance matrices of observation and background error are very important for the accuracy of forecasts. Please clarify how the uncertainties are determined in the text.

The following sentences have been added to Section 2.2.
"Then background error covariance is obtained from an ensemble forecast with the updated analysis ensemble, whereas the observation error is obtained from the satellite retrieval uncertainty information (c.f., Section 3.1)."

For the NO2 retrievals, the following information has been added:
"The retrieval uncertainty of individual pixels was calculated based on error propagation in the retrieval."

2. Section 3.1. The super-observation approach is applied to NO2 and CO, but not other species. Is there any specific reason for this?

Super observations were not generated for the TES ozone and MLS ozone and HNO3 retrievals, because of relatively large spatial representativeness of the vertically-integrated information primarily within the free troposphere for the TES measurements and of the upper tropospheric and lower stratospheric concentrations for the MLS measurements. This is clearly explained in the revised manuscript.

The super-observation approach was also applied for the OMI SO_2 measurements, which is mentioned in the revised manuscript.

3. Page 12, line 340, the authors tabulate the bias of modeled NO2 compared to OMI, SCIAMACHY, and GOME-2 observations. I notice that the consistency is significantly better for results relying on single species (e.g., DECSO in Ding et al.) What could be the driver for the bias? Is the uncertainty of the

assimilated system or the conflict caused by satellite observations of different species used for assimilation?

We couldn't find any other results showing smaller biases against assimilated NO2 measurements including the DECSO paper. Nevertheless, there could be a couple of reasons for the remaining biases in general as below.

Firstly, the resolution of our global model is relatively coarse than that of other regional systems including the DECSO. Our recent study (Sekiya et al., in review) demonstrated that increasing model resolution from 1.1 degree to 0.56 degree significantly data assimilation analysis errors of tropospheric NO2 against assimilated measurements by 5–24%.

Secondly, model performance can affect NO2 analysis bias in data assimilation, as demonstrated by Miyazaki et al (2020b) using a multi-model data assimilation framework. The negative NO2 biases in the TCR-1 and TCR-2 reanalyses over the polluted regions could be associated with errors in the model chemical equilibrium states, planetary boundary layer (PBL) mixing, and diurnal variations of chemical processes and emissions. Adjusting additional model parameters such as VOC emissions, deposition, and/or chemical reactions rates by adding observational constraints could help to reduce model errors.

Thirdly, rather than degrading the NO2 analysis, multiple-species measurements provide important information for improving surface NO_X source estimations and improve the chemical consistency including the relation between concentrations and the estimated emissions (Miyazaki and Eskes, 2013; Miyazaki et al., 2017). The improved representation of NO_X emissions was confirmed by the better agreement of simulated ozone concentrations with independent ozonesonde observations using NO_X emissions from multiple-species assimilation than those using NO_X emissions from NO_2-only data assimilation (Miyazaki et al., 2017).

To discuss the possibilities briefly, the following sentences have been added:
"The remaining negative biases could be associated with errors in the model chemical equilibrium states, planetary boundary layer (PBL) mixing, and diurnal variations of chemical processes and emissions. Meanwhile, Sekiya et al. (2020) demonstrated that increasing model resolution from 1.1 degree to 0.56 degree reduced the analysis errors of tropospheric NO2 by 5--24%."

Line 357, Over India, the model negative bias increased with year because of the lack of the emission increases in the a priori emissions. I suggest adding some analysis about the uncertainty associated with the a priori emissions in the text as well.

The following sentences have been added to discuss the negative biases over India:
"The a posteriori NOx emissions in 2018 are up to 90% larger than the a priori emissions over polluted areas at grid scale, whereas the remaining negative NO2 biases suggest that the NOx emission analysis increments are insufficient. We applied a covariance inflation to the emission factors to prevent covariance underestimation caused by the application of a persistent forecast model, by inflating the spread to a minimum predefined value (i.e., 30% of the initial standard deviation (=40%)) at each analysis step. The inflation was essential to maintain emission variability and continue to increase the emissions. The remaining model biases suggest requirements for a stronger covariance inflation, although too large inflation can cause unstable analysis increments."

Section 7.2. As mentioned by the authors VOCs have a pronounced influence on the tropospheric chemistry. Any specific reason for excluding satellite-observed VOC into the assimilated system?

The main reason is the relatively large uncertainty in the current HCHO satellite products for decadal reanalysis (OMI and GOME-2). In most previous studies, these satellite data have been used as monthly averages to provide systematic constraints on emissions with reduced noises. Because of the short-assimilation window, it is not appropriate to use monthly mean data in our assimilation system. With the recent developments in satellite products (e.g., TROPOMI HCHO and CrIS isoprene), we expect that we will be able to update VOC emissions in the TCR-2 framework.

Specific comments: 1. Page 3, line 58, the description of IFS and the following discussion about reanalysis pop up here. What is the relationship between TCR and those models?

These systems are not related to the TCR-2 system, but have been compared with the TCR reanalyses and thus are useful to introduce here. To clarify this, the sentence has been rewritten as:
"Apart from the TCR systems, employing…."

2. Page 21, line 625, the negative increments over India seems to be inconsistent with the recent report. please try to clarify the reason for such inconsistency.

What we show here is data assimilation increment, not trend or temporal changes. Because the increments strongly depend on the a priori emissions, they can differ among data assimilation systems. To clarify this, “The negative increments” have been replaced with “The negative data assimilation increments”.

3. Fig 8 & 9, the legend is too small to see.

Increased.

4. Fig 11, I recommend adding the r and RMSE in the figures.

Added.

5.The authors performed comprehensive evaluations for multiple species using multiple measurements. It is very impressive, but easy for readers to get lost. I recommend a table to summarize the validation results by species and sources of measurements. It would be easier for readers to follow the improvement by listing the global average difference between aprior and post simulations against measurements.

Thank you for the suggestion. Table 10 has been added to summarize the key global statistics of NO2, CO, and ozone evaluation results.