Author's comments in reply to the anonymous referee for "Evaluation of a multi-model, multi-constituent assimilation framework for tropospheric chemical reanalysis" 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

This study presents an unprecedented work consisting in applying the same data assimilation framework to different Chemical Transport Models. This allows to treat the different flavor of model error terms such as chemistry, transport and deposition schemes in a consistent manner. The study is well written and presents a comprehensive work. The evaluation against independent measurements is greatly appreciated. The discussion on the model error is quite interesting, this could be further compared to the observations error, which has been effectively used in the analysis procedure. It is stated that the analysis error can be overestimated, which is not common in EnKFs where the spread in the ensemble is usually too small. The ensemble spread and the analysis error of the model mean against TES can be as low as 1 ppb (Fig. 6). One can expect that the use of a multi-model estimate (inter-model differences in the spread) of the model error fits nicely the observation error. In general, throughout the paper, error bars can be added to the observations.

Thank you for the helpful comments. We have added the following comments about the relationship between the observation error and the obtained results. Since the detailed observation error information are not provided for most of the independent observations, we have not added error bars in the figures.

In Section 2.4.1
"The observation error is 5-10 % between the surface and 30 km (Smit et al., 2007)."
In Section 4.1.2
"The obtained mean errors and multi-model spreads of the data assimilation analysis are comparable to or smaller than the ozonesonde observation errors (5-10 %). "
In Section 4.1.1
"The mean retrieval uncertainty of the TES measurements is typically between 5-10 ppb in the SH and 10-15 ppb in the NH, which are larger than the multi-model spread and the mean model errors after data assimilation."

In Section 4.2
"The OMI NO_2 super-observation error is typically about 20-50 % of the tropospheric NO_2 columns over polluted areas, which are comparable to or larger than the analysis error."

P9L25: "The state vector includes the chemical concentrations of various species as well as the surface sources of NOx and CO and LNOx sources". And P10L6: "We also applied covariance localization for different variables in the state vector (Kang et al., 2011), by setting the covariance among non- or weakly related variables to be zero." Could you be more specific and describe which observations are used to estimate which state variable? In particular in the case of the Ozone responses to NOx perturbation, it would be interesting to know which state variables are optimized while assimilating ozone.

The following sentences have been added:
"Concentrations of NOy species and ozone were optimized from TES ozone, OMI and SCIAMACHY NO_2, and MLS ozone and HNO_3 observations."

Minor comments: For future work, you could look at the impact of assimilation on the O3-CO correlations and dO3/dCO enhancement ratios (Zhang et al. 2006).

Thanks for the suggestions. I agree that ozone-CO correlations would be another interesting model response that could be measured using the MOMO-Chem framework. In the current data assimilation setting, we didn't optimize CO concentrations using ozone observations to avoid sampling errors (c.f., Section 2.6). To note the future possibility, the following sentences have been added:
"In addition, tropospheric ozone shows strong correlations with other species such as CO (Zhang et al., 2006) over regions such as continental outflow regions. The relationship can be included in the state vector to improve the tropospheric ozone analysis. The uncertainty information in the CO response to ozone obtained from the MOMO-Chem can be expected to provide useful information on model diagnostics and future predictions."

P5L4: the family name of the author is "Olivier" not "Oliver" according https://themasites.pbl.nl/tridion/en/themasites/edgar/publications/index-2.html

Corrected.

P9L24: The data assimilation settings were almost same among the systems as follows. Should it be a "the" before same? this sentence can be improved.

Rewritten as
"Almost the same data assimilation settings were used for the four systems as follows."

P12L26 "lowover", is there a space missing ?

Added.

P15L12: "the large multi-model spreads (25–55%) suggest that individual models have large uncertainty in representing strong ozone productions, for instance, associated with VOCs emissions and chemistry." Is it really an uncertainty or is it just because the chemical regime is not in favor of a clear and linear relationship between Ozone and NOx emissions.

It is possible that the chemical regime differs among the models because of, for instance, different level of VOCs. Also, chemical and transport mechanisms affecting ozone formation can differ among the models. To clarify it the sentence has been rewritten as follow:
"the large multi-model spreads (25–55%) suggest that individual models have large uncertainty in representing strong ozone productions, for instance, associated with VOCs emissions and chemistry that could results in different chemical regimes."

P15L16: "the mean of the individual model estimates (solid while lines)". I guess you mean 'white line' instead of 'while line'.

Corrected

P21 L34: "Fig.13 compares the global distributions of annual and tropospheric mean OH concentrations." How do you define tropospheric? – also in legend of Fig. 13.

The tropospheric mean OH concentrations were calculating by averaging OH from the surface to 300 hPa, which is explained in the revised manuscript. Note that the averages between the surface to 300 hPa and between the surface to the tropopause made only small differences.

Figure 3: You should define what the dashed lines are in legend.

Defined.

Figure 5: you could replace 'model' by 'control'

In all the figures, we labeled "Model" rather than "Control". Although either is fine, we would like to keep the consistent labeling (Model).

Figure 6: "(a) Standard Deveation among the models"

Collected.