Reply to Comments by Owen R. Cooper (TOAR Scientific Coordinator of the Community Special Issue) on:

Assessing the relative impacts of satellite ozone and its precursor observations to improve global tropospheric ozone analysis using multiple chemical reanalysis systems

Sekiya, T., Emili, E., Miyazaki, K., Inness, A., Qu, Z., Pierce, R. B., Jones, D., Worden, H., Cheng, W. Y. Y., Huijnen, V., and Koren, G.

EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-2426
Discussion started: 20 August 2024; Discussion closes 01 Oct., 2024

We would like to thank Dr. Owen R. Cooper for his comments on our manuscript, which greatly improved the quality of the manuscript. We have revised the manuscript accordingly, and the main changes are as follows:

1) An analysis was added to highlight the seasonal cycles in the improvements in ozone analysis by data assimilation (Figure 5 in the revised manuscript).
2) A more in-depth discussion was added on the impacts of discontinuity in the current satellite observing system on tropospheric ozone analysis, as well as potential strategies to address this issue after the termination.
3) An analysis was added to the Supplementary material to explain the reasons behind the large differences between TCR-2 and GEOS-Chem adjoint based on only the OMI NO_2 assimilation.

Individual comments (in black) and point-to-point replies to them (in blue) are listed below. Revised text (*italicized font*) from the updated manuscript is in quotes below.

Note that we have added line types in Figure 4 in the revised manuscript to assist readers with color vision deficiencies according to the comments from Editorial Support.

This review is by Owen Cooper, TOAR Scientific Coordinator of the TOAR-II Community Special Issue. I, or a member of the TOAR-II Steering Committee, will post comments on all papers submitted to the TOAR-II Community Special Issue, which is an inter-journal special issue accommodating submissions to six Copernicus journals: ACP (lead journal), AMT, GMD, ESSD, ASCMO and BG. The primary purpose of these reviews is to identify any discrepancies across the TOAR-II submissions, and to allow the author teams time to address the discrepancies. Additional comments may be included with the reviews. While O. Cooper and members of the TOAR-II Steering Committee may post open comments on papers submitted to the TOAR-II Community Special Issue, they are not involved with the decision to accept or reject a paper for publication, which is entirely handled by the journal's editorial team.

General Comments:

TOAR-II has produced two guidance documents to help authors develop their manuscripts so that results can be consistently compared across the wide range of studies that will be written for the TOAR-II Community Special Issue. Both guidance documents can be found on the TOAR-II webpage:

https://igacproject.org/activities/TOAR/TOAR-II

The TOAR-II Community Special Issue Guidelines: In the spirit of collaboration and to allow TOAR-II findings to be directly comparable across publications, the TOAR-II Steering Committee has issued this set of guidelines regarding style, units, plotting scales, regional and tropospheric column comparisons, tropopause definitions and best statistical practices.

Guidance note on best statistical practices for TOAR analyses: The aim of this guidance note is to provide recommendations on best statistical practices and to ensure consistent communication of statistical analysis and associated uncertainty across TOAR publications. The scope includes approaches for reporting trends, a discussion of strengths and weaknesses of commonly used techniques, and calibrated language for the communication of uncertainty. Table 3 of the TOAR-II statistical guidelines provides calibrated language for describing trends and uncertainty, similar to the approach of IPCC, which allows trends to be discussed without having to use the problematic expression, "statistically significant"

We sincerely appreciate your comments.

The color scheme specified by the guidelines is applied to the spatial maps in the manuscript. For the ozone mixing ratio unit, "nmol mol^{-1}" is used according to the guidelines. This study focuses on the year 2010 and does not analyze trends. Moreover, the expression "statistically significant" is not used.

Specific Comments:

Line 220

It's not clear which TOAR surface ozone product was used. Are you using the pre-complied 2x2 degree gridded product, available from the PANGAEA repository? If so the PANGAEA citation needs to be provided, as follows:
Schultz, Martin G; et al. (2017): Tropospheric Ozone Assessment Report, links to Global surface ozone datasets [dataset publication series]. PANGAEA, https://doi.org/10.1594/PANGAEA.876108, Supplement to: Schultz, MG et al. (2017): Tropospheric Ozone Assessment Report: Database and Metrics Data of Global Surface Ozone Observations. Elementa - Science of the Anthropocene, 5:58, 26 pp, https://doi.org/10.1525/elementa.244

Yes, we used the precompiled gridded data from the PANGEA repository and have included the citation for it in the revised manuscript.

The model/reanalysis evaluation is conducted across the full year of 2010. As reported by Logan (1999), ozonesondes show strong seasonal cycles of ozone in all latitude bands, with mid-tropospheric ozone at northern mid-latitudes increasing by more than 50% from winter to summer. Was there any seasonal dependence regarding the improvement achieved by assimilating satellite data?

Yes, we found a distinct seasonal dependency in the improvements by data assimilation as well as in the reanalysis performance, which varied among the reanalysis data sets and locations. To clearly illustrate the seasonality of the reanalysis bias and the impact of data assimilation, we have added monthly timeseries of biases in the reanalysis and control runs to Figure 5 and the corresponding description in section 3.2.2 as follows:

(p. 10, line 305–p. 11, line 308)
"*As shown in Figures 5 and S1, over the northern mid and high latitudes, RAQMS showed larger seasonal amplitudes in model bias, with a maximum in boreal spring, compared with CAMS, TCR-2, and IASI-r. Over the southern mid and high latitudes, CAMS and RAQMS showed larger negative biases in austral summer and fall compared to other reanalyses, while TCR-2 and IASI-r exhibited maximum positive biases in austral spring.*"

(p. 11, lines 312–314)
"*The seasonal variations of model biases were approximately 20% larger in IASI-r compared to the other models. IASI-r showed maximum negative biases over the northern*

midlatitudes and maximum positive biases over the southern high latitudes during summer."

(p. 11, lines 318–320)
"*IASI-r exhibited larger biases compared to other models over the mid and high latitudes in summer in both hemispheres, whereas RAQMS and TCR-2 showed larger negative biases over the mid and high latitudes in winter in both hemispheres.*"

(p. 11, line 325)
"*Seasonally dependent biases were also improved by data assimilation over all the latitude bands.*"

(p. 11, lines 331–333)
"*The seasonal variations in model biases were also reduced. However, multi-system mean biases in winter remained over the high latitudes in both hemispheres, likely due to the limited number of assimilated observations in these regions during winter.*"

(p. 11, line 339)
"*In these regions, the seasonal dependency of model biases was not improved by data assimilation.*"

The model output and the reanalysis were evaluated against ozonesondes using all observations collected in broad latitude bands, rather than comparing models and observations at individual monitoring sites. Can you comment on why this approach was chosen? I assume it's because you need a large sample size, due to the fact that most individual monitoring sites only sample the atmosphere once per week, which fails to provide accurate monthly means, as reported by Logan (1999). A new paper published in the TOAR-II Community Special Issue discusses the challenges of detecting long-term ozone trends based on once-per-week sampling and their Figure 1 shows the errors associated with trying to characterize monthly mean ozone with just 4 profiles per month (Chang et al., 2024).

Our focus was on validating ozone profiles on broader spatial scales to provide regional-scale reanalysis performance that reflects regional characteristics. This approach was chosen instead of evaluating reanalysis performance at individual observation sites, which can be affected by sparse temporal sampling, limited spatial coverages, and the influence of local-scale processes. To achieve this, validation was performed using ozonesonde observations aggregated across broad latitude bands. This methodology is supported by Tilmes et al. (2012), who emphasized that regional aggregates of individual sites provide a more representative characteristics of larger regions. Nevertheless, we acknowledge that the number of observations within each latitudinal band may not always be sufficient to fully capture regionally representative model performance, as discussed in Miyazaki and Bowman (2017). This limitation is more carefully discussed in the revised manuscript as follows. Note that temporal trends, including challenges related to limited sampling frequency, will be analyzed in companion papers from the TOAR-II Chemical Reanalysis Working Group (Jones et al., submitted).

The following sentence has been added in Section 2.3.2.
(p. 8, lines 237–244)
"*The validation was conducted against ozonesonde observations collected for five latitude bands to evaluate the global reanalysis performance in a manner that reflects regional characteristics. This approach was chosen instead of evaluating reanalysis performance at individual observation sites, which can be influenced by sparse temporal sampling, limited spatial coverage, and the influence of local processes. Aggregating individual ozonesonde sites with similar characteristics provide a more representative view of larger regions, as demonstrated by Tilmes et al. (2012). However, we acknowledge that the number of observations within each latitudinal band may not always be sufficient to fully capture regionally representative model performance (Miyazaki and Bowman, 2017) or to accurately evaluate long-term trends (Chang et al., 2024).*"

Minor Comments:
Figures 1, 2 and 3 are very hard to read because the panels are so small. The panels could be made larger if they were arranged vertically, in 3 columns; also, the color bars can only be shown once, instead of repeating for each panel.

According to your suggestion, Figs. 1, 2, and 3 have been modified as follows:
(1) Panels were enlarged and arranged vertically (3 columns x 7~6 rows).
(2) Color bars in each panel were removed and shown once below the panels.
(3) X-axis and y-axis labels were removed and shown once for each column and row

(only Fig. 1).
(4) Titles on top of each panel were removed and shown once for each column and row.

Similar changes were made for other figures according to the comments from anonymous reviewer #2.

Similarly, Figure 5 is hard to read because the panels are too small. The lat/lon numbers can be removed and the color bars can only be shown once.

We have enlarged individual panels, removed lon/lat labels and color bars, and showed the color bar once below the panels.

line 25
aircrafts should be aircraft, which can be either plural or singular

This phrasing has been corrected.

line58
please explain the perfect model assumption

We have added an explanation of this assumption as follows:
(p. 3, lines 60–62)
"*the perfect model assumption, which assumes that the forecast step within the data assimilation does not add systematic errors through model processes (Lahoz et al., 2010). Hence reanalyses inherit underlying model biases to an extent that depends on the frequency and sparseness of observations.*"

Line 217
(30°N–30°N) should be (30°S–30°N)

This error has been corrected

References:

Miyazaki, K. and Bowman, K.: Evaluation of ACCMIP ozone simulations and ozonesonde sampling biases using a satellite-based multi-constituent chemical reanalysis, Atmos. Chem. Phys., 17, 8285–8312, https://doi.org/10.5194/acp-17-8285-2017, 2017.

Tilmes, S., Lamarque, J.-F., Emmons, L. K., Conley, A., Schultz, M. G., Saunois, M., Thouret, V., Thompson, A. M., Oltmans, S. J., Johnson, B., and Tarasick, D.: Technical Note: Ozonesonde climatology between 1995 and 2011: description, evaluation and applications, Atmos. Chem. Phys., 12, 7475–7497, https://doi.org/10.5194/acp-12-7475-2012, 2012.