Manuscript # acp-2022-678

Responses to Referee #1
 # Summary

The authors have thoughtfully responded to my requests. The model performance results really highlights the currently poor performance. Although the performance does not inspire confidence, the authors have clearly documented the performance and future readers may make up their own minds. I do have a few minor recommendations that do not require subsequent review.

The model performance section should be improved a bit before publication. As written, the performance figure is in the supplement, but it should be moved to the main body. The authors statement that it "can capture the O3 seasonality" is not particularly compelling. The R values from 0.21 to 0.45 are also not particularly good. That being said, the authors are currently comparing sites to a coarse grid. Obviously, any sub grid scale variability will not be captured. The authors could consider some sort of monitor averaging to account for this before calculating R.

Response:

Thank you for the suggestions. We have moved the performance figure to the main body. Both the model and observations show high values in summer and low values in winter. This pattern can be captured by the model. We have revised the description as “It can capture the seasonal pattern of O_3 that high mixing ratios in summer and low mixing ratios in winter.” We agree with the review that monitor averaging method could be better, but is unlikely to largely improve the statistics.

The authors also seem to have moved their primary result of trends to a supplemental table (Table S1). I hope that was a mistake as the values should be presented in the main body.

Response:

We have moved it to the main body.

With those very minor updates, I support publication of these results. I'll note, however, that these results largely call for higher resolution simulations or a focus on metrics (e.g., MDA8 or MDA1) that the model simulates better.

Response:

Thank you. Since the output of the current simulation results is monthly average data, we are unable to calculate MDA8 or MDA1. We will take your suggestion into consideration in our future studies with higher model spatial and temporal resolutions.

Line-by-Line

171: We [do] not

Response:

We have revised it.

Observation methods:
The new figure S4 shows the observations sites and highlights that the observation method section needs a bit of work. Right now, you say that each site must have

Response:

We guess the reviewer mean the sentence for the site selection. We have modified the description as “Seasonal mean for any site that has less than 50% data availability in any month of a season is discarded following Lin et al. (2017). O_3 trends is calculated only when the seasonal data availability is greater than 85% during the analyzed period (more than 22 years).”

Manuscript # acp-2022-678

Responses to Referee #2

The authors did good job with their detailed repose to my comments/questions. The manuscript, however, needs (at least) minor revisions. Most of the additional analyses and information were just put to the conclusion & discussion section or the Supplement. In addition, the changes in the manuscript need a more detailed proof-reading, because it is very hard to follow some of the newly added paragraphs. Finally, at least to my opinion, the revised manuscript is partly very confusing because it jumps between (sub)-figures.

We thank the reviewer for all the insightful comments. We have now made detailed proof-reading throughout the text and reordered and referenced all figures in the manuscript. Please see our point-by-point response below.

Please find my detailed points below:

1) The order of figures is not the same order as they are referenced in the manuscript; this does also apply for the Supplement. Please also check if all supplementary data is referenced in the manuscript. For example, a reference to Table 2 seems to be missing.

Response:

Thank you for the suggestion. We have now made corresponding adjustments to all figures and tables in the main text and the supplement.

For example:

"The time series of the source contributions from NO_x and reactive carbon emissions are shown in Fig. 5 and the O_3 trends in the U.S. attributed to different emission source sectors are shown in Fig. 6."

"Time series of the source contributions are shown in Fig. 7 and the O_3 trends in the U.S. attributed to different emission source regions are presented in Fig. 8."

Table 1 (the original Table 2) has now been mentioned in the *Emissions and Observations* section.

2) To my opinion Table S1/Fig. S4 are an essential part of the model evaluation and should be part of the manuscript. Do you agree?

Response:

Agree. They have been moved into the manuscript.

3) The discussion about the trends of the emissions is placed at the end of the

manuscript. To my opinion this discussion would fit much better to the parts where the emissions/Fig 3 are discussed. In addition, I am missing a discussion about how differences in the emissions might affect the results of the study. Are your estimates of the trends from domestic/Asian emissions are likely at the upper or lower end (given differences in the emissions).

Response:

Thank you for the suggestion. We have modified the emission description and moved it to the *Emissions and Observation* at the *Methods* section, as following:

"Many studies have reported that the previous CEDS version 20160726 (hereafter $CEDS_{2016}$) has large biases in the regional emission estimates (e.g., Cheng et al., 2021; Fan et al., 2018). In this study, the CEDS version 20210205 is used (hereafter $CEDS_{2021}$), which builds on the extension of the CEDS system described in McDuffie et al. (2020) and extends the anthropogenic emissions to year 2019. It updates country-level emission inventories for North America, Europe and China and has considered the significant emission reductions in China since the clean air actions in recent years. The global total NO_x emission from $CEDS_{2021}$ is lower than that of $CEDS_{2016}$ after 2006 and it shows a fast decline since then. In 2014, the global total anthropogenic emission of NO_x in $CEDS_{2021}$ is about 10% lower than the $CEDS_{2016}$ estimate. This difference is mainly reflected in the NO_x emissions in China and India. $CEDS_{2021}$ has a lower estimate of the global NMVOCs emission than $CEDS_{2016}$ by more than 10% during the recent decades, attributed to lower emissions from Africa, Central and South America, the Middle East and India. The using of the $CEDS_{2021}$ emission inventory in this study could reduce the contributions of NO_x emissions from East Asia and South Asia to the U.S. O_3 mixing ratios and trends, as compared to $CEDS_{2016}$. However, recent study reported a difference in aviation emission distribution of NO_x between CMIP5 and CMIP6 related to an error in data pre-processing in CEDS, leading to a northward shift of O_3 burden in CMIP6 (Thor et al., 2023). Therefore, the contribution of the aircraft emissions of NO_x to the O_3 mixing ratios could be overestimated at high latitude regions."

4) Further, I am missing a discussion on how the model biases might affect the derived trends. In l229ff the authors write "..which will be discussed in the following section" but I am missing this discussion.

Response:

Here, "The model can produce the sign of the changes, but has large biases in magnitudes, which will be discussed in the following section." Then we describe the biases in magnitudes between observation and simulation based on Table 1. The discussion of model biases is mainly in the conclusions and discussions as the following:

“Compared to observations, the decreasing trend of O_3 mixing ratios over WUS in summer and increasing trend over EUS in winter are overestimated in the CAM4-chem model. Because most O_3 monitors are located in urban areas and these areas generate strong O_3 during the day and have strong oxidation titration at night, the daily and grid averaged O_3 mixing ratios output by the model could be inconsistent with the urban observations. The overestimate of O_3 trend in the EUS might be related to a potential biased model representation of vertical mixing in winter. Large uncertainties existing in the emissions also result in the biases in the O_3 simulation. Lin et al. (2017) found that the contribution from increasing Asian emissions offset that from the U.S. emission reductions, resulting in a weak O_3 trend in WUS. In this study, the Asian NO_x emissions only contribute to 0.6 ppb/decade of the total positive trend in WUS in summer, much lower than the 3.7 ppb/decade decrease attributable to the domestic emission reductions, suggesting that the Asian contribution to the O_3 trends in WUS is possibly underestimated in this study. We also found that the model did not capture the significant increase in summertime O_3 levels in China in recent years, which could explain the low contribution from Asian sources. Additionally, international shipping can have a disproportionately high influence on tropospheric O_3 due to the dispersed nature of NO_x emissions (Butler et al., 2020; Kasibhatla et al., 2000; von Glasow et al., 2003), together with the weakened NO_x titration, resulting in the overestimation of O_3 trends. The fixed CH_4 mixing ratio during simulations also biased the modeled O_3 trends, which deserves further investigation with the varying CH_4 levels in future studies. The coarse model resolution also contributed to the biases. The overestimate of O_3 trend over EUS in winter, likely related to the bias in NO_x titration, implies the overestimate of source contributions to the trends in magnitude.”

5) I suggest that information on how you calculated the trends (least square linear trend) and your definition of “significance” should be an essential part of the method section.

Response:

Thank you for the suggestion. We have now added following sentences in the *Emissions and Observation* section: “Trends in this study are calculated based on the linear least-squares regressions and the significance is identified through the F test with the 95% confidence level.”

6) The authors replied that “the results for the tags “STR”, “LGT”, “AIR”, and “SOIL” should be similar between sector and regional run”.

I don’t understand this answer. If there are two simulations – one with regional and one with sectoral attribution – with the same atmospheric dynamics and the same emissions, the contributions of identical tags (air, str etc.) should be the same between the two simulations? As example, the trend of STR in Fig 5e

is 0.64 ppb/decade and the trend of STR in Fig. 7e is 0.70 ppb/decade.
I would expect that the two tags show the same trend in both runs. Most likely I get something wrong here. Could you please explain this in more detail? Is the atmospheric dynamics different? This should also be explained in the manuscript.

Response:

The differences between the sector and regional simulations could be due to the slight difference in the atmospheric dynamics related to the nudging of the wind fields. The zonal and meridional wind fields are nudged to the reanalysis data at a 6-hourly relaxation timescale, rather than completely driven by the reanalysis data as in the chemical transport models. Also, other meteorological factors like temperature and humidity were not nudged to the reanalysis. However, the slight differences would not affect the main conclusion of this study. We have added the explanation in the manuscript.

7) The changed Sect 3.4 jumps between the subfigures (8a, b followed by 8e). Some information are doubled (trends stratosphere). Please rephrase the section completely and think about splitting Fig 8 into two figures."

Response:

We have now split the Fig. 8 into two figures and rephrased the section as the following:

"Many studies have reported that O_3 spatial distribution is strongly modulated by changes in large-scale circulations (e.g., Shen and Mickley, 2017; Yang et al., 2014, 2022). Based on the MET experiments with anthropogenic emissions kept unchanged, the changes in large-scale circulations show a weak influence on the U.S. O_3 trends in summer (Fig. 9a) but cause a significant O_3 rise in the central U.S. in winter (Fig. 9b). Averaged over the U.S., the near-surface O_3 mixing ratio in winter increases at a rate of 0.7±0.3 ppb/decade during 1995–2019 in MET experiments. It suggests that the variation in the large-scale circulation is responsible for 15% of the increasing trend in wintertime O_3 mixing ratio by 4.7±0.3 ppb/decade in the U.S. during 1995–2019 simulated in BASE experiment.

The changes in atmospheric circulation pattern support the above finding. Compared to 1995–1999, anomalous northerly winds locate over high latitudes of North America in 2015–2019 (Fig. 9c), strengthening the prevailing northerly winds in winter. In addition, an anomalous subsidence occurs over the central U.S. in 2015–2019, compared to 1995–1999 (Fig. 9d). The anomalous subsidence transport O_3 from high altitudes and even stratosphere to the surface and the strengthened winds transport O_3 from remote regions (e.g., O_3 produced by Asian NO_x emission) to the central U.S., both contributing to 0.2±0.1 ppb/decade of the O_3 increase over the U.S. (Fig. 10). The finding is consistent with Lin et al. (2015) that variations in the circulation facilitate O_3

transport from upper altitudes to the surface, as well as foreign contributions from Asia. The anomalous atmospheric circulation is likely linked to the location of the midlatitude jet stream, which is influenced by ENSO cycle."

8) Further, I am somewhat surprised by the statement "The mixing ratio is sometimes expressed as concentration in many studies, so we prefer to keep it as it is." Mixing ratios and concentrations are two complete different things. Please check for example this Eos article:
https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/00EO00007
Please correct your manuscript accordingly.

Response:
We have now corrected the descriptions as "mixing ratio" throughout the manuscript.

Some more detailed comments:

l78: Please do not use the term 'contribution' when talking about the perturbation method.

Response:
We revised it to "when being used to estimate the impacts of changes in multiple sources".

l107ff:As mentioned in the first review: there are methods on the global scale for sectoral attribution (e.g. Emmons et al., 2012, Grewe et al. 2017, Butler et al., 2018)

Response:
We have now added these references.

l291ff: I don't understand this sentence. Do you mean inland shipping?

Response:
We have revised it to "Due to a strong chemical sink associated with photolysis of O_3 with subsequent production of hydroxyl radical (OH) from water vapor in summer (Johnson et al., 1999), the effect of increased **international shipping emissions over the remote ocean regions** on the continental United States was blunted."

l307ff: Please reference Fig 5 b/d here

Response:
Added the reference.

l328f: Do you mean near shore shipping emissions or the actual activity data?

Response:

We mean near shore shipping emissions and modified this sentence as "The decrease in near-shore shipping emissions …"

l374f: Where can I see the trends 1.2 / 1.5 ppb/decade? Please explain (also in the manuscript).

Response:

We now revised it as "due to the weakened NO_x titration. Increases in aviation and shipping emissions **together** explain the 1.2±0.1 and 1.5±0.1 ppb/decade of O_3 trends in WUS and EUS, respectively".

396f: I can't follow the conclusion why changes in anthropogenic emissions are the main factor from what is written here. Please clarify and rephrase if needed.

Response:

Because the trends shown by the MET experiments with the variation in large-scale circulation alone only account for a small fraction of the trends in the BASE experiments with the combined effect of large-scale circulation and emissions.

However, considering the large-scale circulation contributes to 15% of the O_3 trend in winter, which is also an important factor driving the O_3 change, we have deleted this sentence.

l449f: What is the role of the emissions for the bias?

Response:

We have added it as "Large uncertainties existing in the emissions also result in the biases in the O_3 simulation." In this study, we applied the latest CEDS version 20210205, which has corrected several biases as compared to its previous version 20160726.

l462ff: Why is the part about the bias over China deleted?

Response:

We have added this sentence again.

l494ff : Please quantify the differences

Response:

Quantified as "In 2014, the global total anthropogenic emission of NO_x in

$CEDS_{2021}$ is about 10% lower than the $CEDS_{2016}$ estimate. This difference is mainly reflected in the NO_x emissions in China and India. $CEDS_{2021}$ has a lower estimate of the global NMVOCs emission than $CEDS_{2016}$ by more than 10% during the recent decades, attributed to lower emissions from Africa, Central and South America, the Middle East and India."

l500f: The sentence about EDGAR seems misplaced here. Either include a proper comparison with EDGAR (btw. V 5.x is available) or delete this sentence.

Response:

We have deleted it.

Reference:

Hoesly, R., O'Rourke, P., Braun, C., Feng, L., Smith, S. J., Pitkanen, T., Siebert, J., Vu, L., Presley, M., Bolt, R., Goldstein, B., and Kholod, N.: CEDS: Community Emissions Data System (Version Dec-23-2019), Zenodo, https://doi.org/10.5281/zenodo.3592073, 2019.

Cheng, J., Tong, D., Liu, Y., Yu, S., Yan, L., Zheng, B., Geng, G., He, K., and Zhang, Q.: Comparison of current and future PM2.5 air quality in China under CMIP6 and DPEC emission scenarios, Geophys. Res. Lett., 48, e2021GL093197, https://doi.org/10.1029/2021GL093197, 2021.