Review#2 *of Froidevaux et al. about ozone and carbon monoxide trends in the tropical upper troposphere*

This second version of the manuscript is far much better than the previous one, resolving (by updating the text or by answering my comments) all the issues I have pointed out in my first review. The manuscript now reads well and with clear motivations. I have three new general comments that I ask the authors to address before publishing this study.

General comments:

• This study is based on the 2005-2020 while three additional years are available. For a paper about trends, it is difficult to understand why the period is not the longest available? Extending up to 2023 would include three additional years which would potentially improve the significance of the trends, especially for CO. This would mean to recompute all the figures and update the text accordingly – so an important effort – but it would make this paper stronger. If not, is there any justification to keep the 2005-2020 period?

Answer: The main issue here is not the MLS data but the setting up and running of the models; to keep it short, updates like this would not happen in less time than several months from now, besides the other "important work" of updating just about everything in this already long paper (with a lot of coding behind this). Also, we had already extended the times (years used) for the model and MLS data analyses beyond the original project's study dates. Just as importantly, the original funds for this project have run out (and the first author is now working half-time overall).

However, we have now run the MLS trends through 2023 and we agree to provide the zonal mean trends as an addition to the current main focus through 2020, to show how this affects the MLS UT O_3 and CO trends. The error bars are reduced by about 23%, but the overall results versus 2005–2020 agree within the error bars.

Therefore, the O_3 conclusion section now has the following added text: "We note that the MLS tropical UT zonal mean O_3 trends for 2005–2023 are 0.34 ± 0.22 %yr^{-1}, so these trends have only changed by a small amount versus the 2005–2020 results; it is useful that the trend error bars are reduced by about 23% for the analysis using 3 more years (the same holds for the CO 2005–2023 trends mentioned below). However, we cannot readily update any of the model simulations (and related trend comparisons) with more analysis years at the time of this writing."

Similarly, the CO conclusion section now has the following added text:
"We note that the MLS tropical UT zonal mean CO trends for 2005–2023 are -0.09 ± 0.23 %yr^{-1}, so these trends have changed by somewhat more than the ozone trend results and are closer to zero than the 2005–2020 MLS CO UT trends. Unfortunately, the coming end of the MLS data record will soon make such MLS updates impossible. As we noted for O_3, we cannot readily update any of the model simulations (and related trend comparisons) with more analysis years."

We also note, in this response (but not in the manuscript), that the TOAR-II assessment (update) is expected to cover a range of years not extending beyond 2020. Studies well beyond the current scope of this manuscript will be needed to assess how the model simulations used here will change regarding UT O_3 and CO trends, after additional years are carefully added in.

• *The discussion sections 2.2.4 and 3.2.4 are still difficult to read (in particular the first one) because citing to a lot of numbers from different publications. In order to improve the readability, I suggest to create figures with these numbers to allow the reader to visualize the comparison of your study with the findings from litterature. I also suggest to create a table with these numbers where to cite the papers in the supplement. This will lighten these sections and largely improve the message of the paper.*

Answer: This refers to the current discussion sections 3.1.4 (not 2.2.4) for O_3 and 3.2.4 for CO.

We have left the CO section mostly as is, since it is indeed shorter, and there are really no comparable UT trend studies for CO, so we point to some column results but in a more general way; we do not see that making a Table (or Figure) for CO is necessary or cleaner than our few "comparison sentences" in Sec. 3.2.4, where past references and process studies/results are also mentioned, for context. Each paragraph in that section has its purpose, and not all of the section is about past comparisons (for the very few that do exist).

For the ozone discussion section, we have added a Table (Table 2) to help summarize the most relevant (recent) studies of UT ozone trends in the tropics; in fact, there are really only two of these references, with time periods most comparable to the MLS period, and/or with enough information that includes ranges or error bars clearly enough. The trends are largely based on IAGOS and ozonesonde results (Thompson et al., 2021, and Gaudel et al., 2024); the results from Gaudel et al. (2020) and Wang et al. (2022) are basically based on the same datasets (IAGOS, sondes) but start in the mid-1990s, so we have relied on Gaudel et al. (2024), which is actually consistent with the above two references, but also shows 2004–2019 results (best for comparing to the MLS time period we use, 2005–2020).

As for O_3 column trend results, we have pointed the reader to the comprehensive work by Gaudel et al. (2024), and we stick to the focused analysis we have in our manuscript for such a comparison (also not readily summarized in a Table since it deals with mapped/gridded trends). We do include trend values showing a range in Table 2, rather than just average trends, so the range of mapped tropical UT O_3 trends from MLS is represented in that way.

We have also added some comments regarding the recent Ma et al. (2024) manuscript (in press), so this was an addition to the discussion section for O_3. We also done some clean-up in some of the text in that discussion section, which has several paragraphs not devoted to trend comparisons, but without much reduction in the total page number. There is really only a page and a half on the trend comparisons, but we feel that showing knowledge of past work is useful, and we could be criticized in a different way if we did not do that enough. In our opinion, just showing a Table without context cannot replace all (or much of) the text.

We do think that Table 2 adds a new dimension to the ozone results, so, despite the work involved in creating that, we thank the referee for that "push". At this point, however, we have probably reached a difference of writing style or opinion rather than a question of scientific accuracy or even clarity, if there is any further significant concern or request for more revisions along these lines; if need be, we will invoke the editor's assistance on such matters.

In summary, we feel fairly strongly that this has converged well enough now, per all the comments, including some new ones; we certainly do appreciate the feedback that has helped to improve this manuscript in the end.

• Contour plots showing differences, sensitivities or trends (Fig. 1, 4, 5, 7, 8, 14 and 16) should use diverging colormap (see below or https://matplotlib.org/stable/gallery/color/colormap_reference.html) centered over 0. For example, in the current state of the manuscript, MLS O_3 trend at 178 hPa (in Fig. 5) seems to be strongly negative on the Northern Pacific Ocean which is not the case. Using a diverging colormap centered over 0 would alleviate that.

Answer: We have certainly alleviated the issue mentioned above regarding Fig. 5 by using a slight change to the color map, which we find divergent enough; also, the choice of min. and max. values plays a role in such visualization issues, so some changes were made there as well; if one reads the color bar carefully, however, this should not really have been a very significant issue.

The same goes for other Figs. (4 and 14, as well as 7 and 16, for O_3 and CO), where we have made some adjustments as well. For Figs. 1 and 8, however, the number of colors is small and adequate enough and/or the values are basically all positive or negative, so issues of divergence and "optimum color bars" go away; we have added some missing latitude tick mark values as a minor improvement (as a response to another minor comment from another referee).

Spending any more time on changing color schemes will never satisfy everyone perfectly, as there is subjectivity in this. We believe that we have addressed such minor comments sufficiently well that no further esthetic (but also time-consuming) changes should really be needed.

Specific comments:

L43: What do you mean by "first-order" correlation? Please, clarify.

Answer: We have decided to eliminate this part of the sentence and to keep the (less vague) details in the main body of the manuscript.

L92: "...showed underestimates of satellite-derived…". Which satellite? Does MLS included?

Answer: We clarified this by adjusting the text as follows: "This model produced underestimates in CO comparisons versus tropical upper tropospheric CO from MOPITT, as well as versus MLS and ACE-FTS CO data in this region; those retrievals compared well with MOPITT CO." See L89–91.

L416: "…the O_3 and CO…". Please, remove "CO" since the section is about O_3.

Answer: Yes, done – thank you (at the start of section 3.1.3).

L447: "…MLS and TCO resolution…". Do you mean vertical resolution? Please, clarify.

Answer: We mean horizontal resolution (we also state "arrived at from appropriate horizontal smoothing" just before this part); we have also now added "horizontal" in front of "resolutions", to be sure this is clear. See L446–447.

L450: "…pattern of UT O_3 and…". Do you mean "UT MLS O_3"?

Answer: Yes, we have now specified "UT MLS O_3". See L451.

L494: "…at least at 147 hPa, there are two strong…". I see only one negative minima above the Pacific Ocean. Please, clarify.
Answer: Adjustments to the color bar and ranges have made the two strong minima at 147 hPa more obvious. We also clarify now (L495) that this is "at 147 hPa" rather than "at least at 147 hPa".

L511: "…observations of O_3 and CO…". Please, remove CO.
Answer: Yes, done - thank you.

L513: Remove "For O_3,".
Answer: Yes, done - thank you.

L623-638: I don't see the point here. What message would you like to bring here? Please, clarify the text.
Answer: Section 3.2.1 is a brief discussion of overall differences between the UT CO climatologies from MLS and the model simulations, with some discussion of potential reasons for systematic differences; these differences are well within the ranges of systematic uncertainties (accuracy) of the MLS data, but there may also be some model issues/systematics (e.g., CO emission details, as pointed out in the reference by Park et al.). Model/MLS biases are of some interest, for the record. However, there are too many likely sources of uncertainty (in both data and models) to try to assign specific causes regarding such biases. We do not really see what to change here, since there is not a specific enough comment (like what sentence might be a potential issue, or unclear). We believe that we have tried, at least, to discuss various (known) error sources and values, even if the conclusion is not that satisfying (no exact causes can be provided).

To wrap that paragraph up for added clarity (we hope) on "what is the point?", we have added the following sentence at the end: "In summary, while we cannot pin down the exact causes for the mean biases between the UT CO climatologies from MLS and the models shown here, a combination of MLS and model systematic errors likely provides a reasonable explanation." (L643–645)

Fig. 4: Include vertical axis for the latitude as it is done for the longitude.
Answer: Yes, done - thank you; similarly, we did this for Fig. 14 (the same plot but for CO).

Reply to anonymous referee #3 review of revised manuscript

This paper by Froidevaux et al. explores a unique opportunity to evaluate upper tropospheric ozone and carbon monoxide trends over the past 15 years based on the comprehensive Aura-MLS satellite observations and offers in addition a comparison to model simulations.

The revisions, particularly the restructuring of the manuscript in my opinion has helped remedy much of the previous reviewers' concerns that the results are very complex and hard to grasp. I have only a few more comments the authors should consider, which I hope will further help to clarify the content before I would recommend publication of this manuscript in ACP.

Minor suggestions:

L21 Not sure why WACCM-CEDS is all of a sudden introduced. Without more explanation, this is confusing. Authors should decide on whether to use WACCM6, WACCM, or WACCM-CEDS in the abstract. I suggest just 'WACCM' for better readability and then to introduce and explain WACCM6 or WACCM-CEDS only in the methods section. Similarly, I suggest to delete the specification '(CAM-chem-CAMS simulation)' since it is non-necessary information at this point or simply say instead 'using emissions from the CAMS chemical reanalysis'.
Answer: Yes, we agree to clarify the Abstract portion of the manuscript.

We use "WACCM" and we have deleted the detail about "WACCM-CEDS" (L21). We also say (L38) "The CAM-chem simulation driven by CAMS-GLOB-ANTv5 anthropogenic emissions yields…" to clarify the main meaning of the "non-CEDS" run…

As explained in the table, WACCM and CAM-chem refer to the model and WACCM-CEDS, CAM-chem-CAMS and CAM-chem-CEDS refer to the simulations. It is also explained in the table that CAM-chem-CAMS uses CAMS-GLOB-ANTv5 anthropogenic emissions; we do not see any need to make any further changes here.

CAMS is an important European project that produces many datasets. This is why we avoid simply using CAMS and we define CAM-chem-CAMS as the CAM-chem simulation driven by CAMS-GLOB-ANTv5 anthropogenic emissions. Readers should then understand that we are <u>not</u> using one of the CAMS chemical reanalyses.

L35 The discussion of these results in terms of model-measurement differences still needs the caveat that these differences are not statistically significant. The trends could be considered consistent with each other within the uncertainty bounds given a statistical z-test, and are not excluding each other as the authors imply.
Answer: We can lengthen the Abstract even more, although we did not feel that this is wise, as the error bars and the main text provide more rigorous information; we are not* trying to imply that two estimates disagree - but one model's average trends do* agree more closely with the MLS results. We agree to add the following sentence in the <u>main text</u> (L895-897), in order to find a middle ground on this comment.
"These three average CO trend results agree within the limits of the (2σ) error bars provided above, although the model versus MLS agreement is more marginal when non-CEDS CO emissions are used." Also, on L933, we have changed the word "whereas" to "while", as this implies somewhat less of a contrast (or 'disagreement').

L40 Small wording issue. I don't think the way you phrase the comparison of MLS UT CO trends to previously published total column CO trends is correct. It factually is like comparing apples to oranges. It would need some caveat that these measures are implying 'consistent, but muted behaviour' between total column and UT trends or something alike.
Answer: While this is technically correct, this is in an Abstract which is already too long; we think that one should not get bogged down in the details in the Abstract, but we have changed the main text. Also, it should not be that far-fetched to check comparisons of UT CO trends to whatever exists in terms of other CO tropospheric trends, given that such data sources and/or analyses are rather limited, and since we know that a significant amount of CO makes it to the upper troposphere from below. We do agree that this is a fairly small wording issue… Therefore, please see the main text (L817–818) where we have added the caveat as follows: "however, these [column CO] trends are not necessarily expected to agree with UT CO trends, since they represent two different altitude regimes."

L76 Suggest to use CO for Carbon monoxide, since you introduced it already in L64.
Answer: Yes, done.

L278 This description is still hard to follow… The CMIP6 emissions are not relevant, I suggest not even mentioning them. Then simply explain. Are both CAM-chem simulations run with CAMS-GLOB-ANT_v5.1, but one of them with CO emissions from CEDS? Is WACCM run with CAMS-GLOB-ANT_v5.3 and also CO emissions from CEDS?
Answer: Yes, we have deleted reference to CMIP6 and we have updated the text (see L277) as follows to clarify: "The emissions used here are taken from CAMS-GLOB-ANT_v5.1…". See also our comments regarding the Abstract above.

L284 I don't understand why you can 'exclude a change in CO secondary formation or sink between these two simulations…'. If the interpretation of the emissions used in the different simulations (previous comment) is correct, then differences may be due to using once v5.3 and once v5.1 that may affect the CO lifetime indirectly through chemical processing.
Answer: The reason is that versions 5.3 and 5.1 are almost identical; v5.3 only includes updates to shipping emissions for years after 2017. We have added this (and further) clarification, so the text (L283–288) now reads as follows: "The CO anthropogenic emissions from the aforementioned versions 5.3 and 5.1 are almost identical; version 5.3 only includes updates to shipping emissions for years after 2017. As there are no differences in the NO_x or volatile organic compound (VOC) emissions in all three simulations, we can exclude a change in CO trends between these simulations as a result of differences in CO chemical formation or sink."

L298 It would be good to provide a link or reference to the CMIP6 aircraft emissions source.
Answer: Yes, we have added the appropriate reference (L300).

L345 here you say 24S-24N, in the caption to figure 1 it is 26S-26N. Please correct. Correspondingly, it is not good practice to not have all axes labelled. Please add the titles (latitude) and tickmarks on the x-axes for rows 1 and 3 in Fig. 1, and all rows in Fig. 4.
Answer: Yes, we have added tick mark labels for latitudes in the Figures (1, 4, and also 14). The Fig. 1 comment has to do with bin centers in the text (and 24 degrees is correct for this); we have emphasized in the Fig. caption that that the full latitude range is mentioned there (so 26S to 26N is correct there).

L361 I would have expressed this more positively for MLS, i.e., that the MLS-ozonesonde differences support the finding of the models' positive biases.
Answer: We think this is not a different enough statement to merit a change.

L383 I would say the difference is definitely 'not significant' and not 'not very significant'.
Answer: Yes, we removed the word "very", although this goes back to the issue of how strongly one should ascribe a significance criterion (of either "yes" or "no") and to the wording choice.

L394 ff This is a very long-winded justification and I do agree that significance is overrated in atmospheric journal publications. However, I would be happier to see this simplified and that you just would say that while the results are not statistically significant, we still can learn about the general directions in the trends and model-observation differences. <u>No need to address this comment</u>, but I would like to give you and the editor a choice to simplify this section again.
Answer: Since this was (partly) in answer to a separate (referee) general comment, we decided to discuss this topic, as the issue does come up and may come up again elsewhere. We have now shortened this somewhat, as shown in the tracked changes, with little loss of information. We feel that the 5 sentences that remain (which actually include the topic of wider latitude bin averages and error reduction) are informative and appropriate, without too much length.

L488 I really like this sensitivity analysis! What would help the reader is to add in one sentence what a positive/negative ozone-ENSO correlation means. i.e., does a positive correlation mean that during El Niño (la Niña), ozone levels increase (decrease), correct?
Answer: Thank you. Yes, for the section mentioned here (see L496–498), we have added the following sentence: "A positive change (or a negative change) in tropical Pacific sea surface temperatures during El Niño (La Niña) conditions will correlate with ozone increases (decreases) in the regions with positive (negative) coefficients." More information on this topic has been nicely discussed by Oman et al. (2013), as we mention at the end of the ozone discussion section.

L587 This result (and particulary hypothesis about the dynamical influence) is consistent with the results of a recent study by Ma et al (https://doi.org/10.5194/egusphere-2023-2411, 2024). These authors find similar trends and provide in addition a model-based quantification of its drivers. In particular, stratosphere-troposphere exchange explains around 10-40% and tropospheric precursor emissions around 60-90% of the observed ozone increases in the South-East Asian tropical region. A reference could be included here to substantiate the claims by Thompson et al. and to provide additional interpretation to your results.
Answer: Thank you. We have now added this interesting but very recent reference and a related sentence (L582-585) of relevance to the tropical region: "While model studies in a recent paper by Ma et al. (2024) also confirm that lower stratospheric ozone and dynamics can significantly

influence long-term UT O_3 trends, their results suggest that, for the tropics, the largest influence (of order 60–70% or more) on UT O_3 comes from the tropospheric components of the O_3 sources." This sort of statement reflects what seems most pertinent for our manuscript's ozone discussion section; we will avoid any more detailed (and non-obvious) comparisons of such issues versus other specific references, as we feel that more significant (comparison) statements regarding this topic are best left for other/future work, such as community assessments.

L645/676 Suggest to add specific years here, particularly 2016 which is an extreme example!
Answer: Yes, we have added more specific and previously published information regarding the largest CO peaks and connections to El Niño years, as follows: "The largest CO peaks in the MLS upper tropospheric tropical record have been correlated with significant El Niño events in late 2006, in 2009–2010, and especially from late 2015 into 2016 (see Park et al., 2021, and references therein, for further information)." See L655–657.

L742/3 please ask the typesetting people to make sure to not break between the minus sign and the trend number!
Answer: Thank you.

L841 delete extra space between 'to' and 'increased'
Answer: Yes, done.

*L877 again I take some issue in the wording that CO trends from UT level measurements should be agreeing with total column CO trends. They *are* consistent but perhaps muted. Please rephrase.*
Answer: Yes, we have now added the following to that sentence (L903-904): ", although one does not expect complete agreement between UT and column trends. See also L817–818.