The manuscript provides a validation of the SNPP OMPS-LP v2.6 ozone dataset. Validation is done primarily with comparisons to MLS/ACE/SAGE III/ISS, but also other data sources like ozonesondes, lidar, and OMPS-NP. It is an important paper to be in the literature since OMPS-LP is poised to become the backbone of vertically resolved global ozone measurements in the near future. Overall the flow is easy to follow, presents useful information, and should be published subject to some revisions. My main suggestion is that the information could be summarized better for data users, specifically the estimated drift levels for the dataset.

We would like to express our sincere gratitude to the referee for their thorough evaluation and valuable comments that help to improve the manuscript.

General Comments

The choice to not degrade SAGE III/ISS to a similar resolution I don't think has a justification other than convenience. While it doesn't change any of the main conclusions of the paper in my opinion, various statements in the discussion are likely because of this choice. For example, the significantly degraded correlation with SAGE at high altitudes, or some of the increased oscillations noted in the drift analysis. I would suggest the SAGE III/ISS data be smoothed to approximately the same resolution as OMPS-LP and the analysis repeated.

> In our analysis the SAGE III/ISS measurements were interpolated onto a 1 km grid for comparison with OMPS. We tested smoothing the SAGE III/ISS data by applying a 2 km boxcar average smoothing function to the profiles and the recalculated the difference profiles and drifts using the smoothed data. As can be seen from the plots below this has little impact on the observed biases between OMPS LP and SAGE III/ISS and the oscillations in the drifts are still present and are a result of the shorter time period available for SAGE III/ISS. Since these changes are minor and do not change any of the discussions or conclusions in the paper we have elected to retain the original 1 km analysis in the paper.

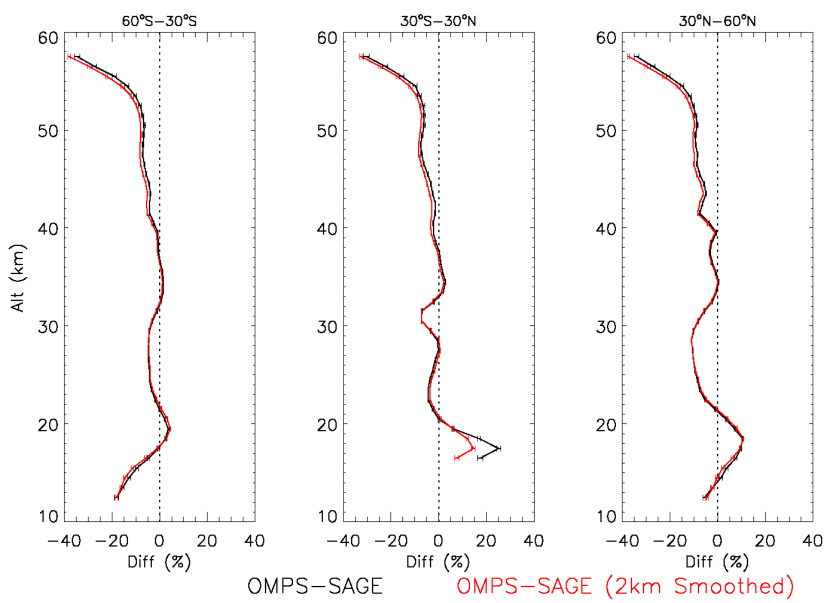

Figure RC2.1: OMPS LP vs SAGE III/ISS differences with (red) and without (black) 2 km boxcar average smoothing of SAGE III/ISS profiles for the period 2017-2024.

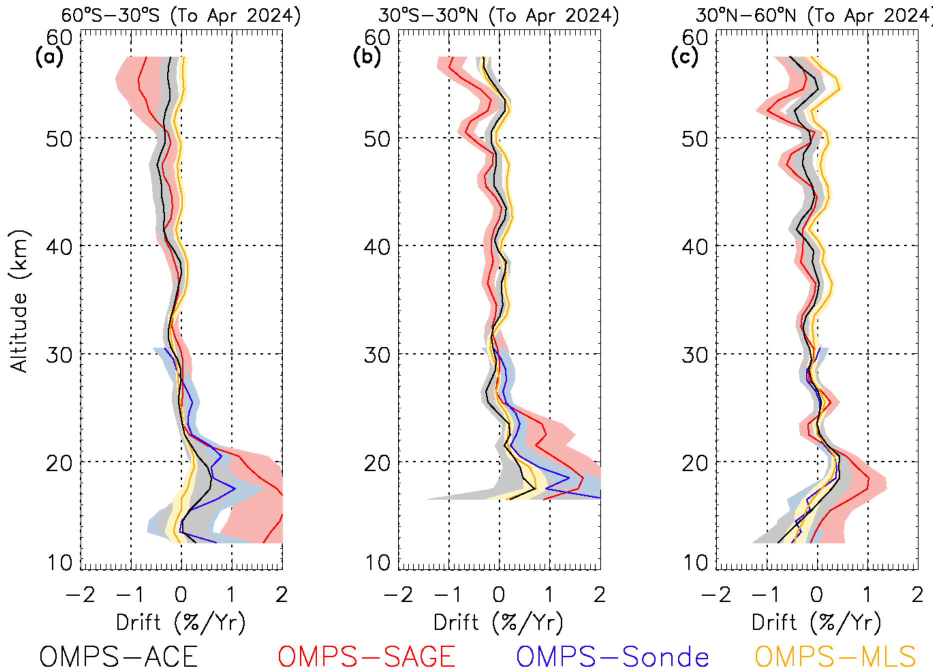

Figure RC2.2: OMPS LP vs SAGE III/ISS drifts with 2 km boxcar average smoothing applied to SAGE III/ISS profiles.

The paper is long, I don't believe it is too long, but there is a lot of information. What might be useful is to provide a summary table (something akin to the MLS data quality document tables) of observed drifts/estimated accuracy for the v2.6 data products based on the analysis done here to better provide easy to digest recommendations to users of the data.

The following summary table has been added in the conclusions section (section 9).

	60ºS to 30ºS		30ºS to 30ºN		30ºN to 60ºN	
Altitude (km)	Bias (%)	Drift (%/decade)	Bias (%)	Drift (%/decade)	Bias (%)	Drift (%/decade)
15.5	-5.84 (±2.84)	-4.74 (±0.89)	-	-	3.34 (±0.28)	-3.50 (±0.62)
20.5	4.32 (±1.08)	1.87 (±3.02)	2.32 (±3.88)	1.14 (±2.57)	4.14 (±0.07)	3.44 (±1.07)
25.5	-2.31 (±1.81)	0.44 (±1.48)	-0.86 (±0.94)	-0.40 (±2.15)	-8.15 (±1.66)	0.93 (±0.17)
30.5	0.65 (±9.54)	-2.38 (±0.47)	-3.90 (±6.14)	-1.41 (±0.08)	-5.29 (±6.24)	-0.54 (±0.69)
35.5	2.61 (±0.50)	-0.22 (±0.59)	2.58 (±0.38)	0.91 (±0.19)	-0.43 (±0.58)	1.24 (±1.55)
40.5	-0.21 (±1.29)	-1.75 (±1.24)	-0.35 (±1.29)	-0.22 (±0.55)	-2.02 (±1.79)	-1.06 (±1.37)
45.5	-4.57 (±0.33)	-2.12 (±3.30)	-4.07 (±0.11)	0.22 (±1.50)	-6.76 (±0.32)	0.54 (±2.24)
50.5	-6.48 (±0.78)	-2.11 (±1.44)	-6.61 (±0.71)	-1.16 (±0.25)	-8.41 (±0.86)	0.11 (±2.84)
55.5	-10.59 (±25.79)	-1.21 (±2.26)	-9.43 (±15.23)	-2.32 (±0.12)	-13.18 (±21.80)	1.59 (±4.49)

I find the stability discussion not to be wrong, but maybe is too optimistic in some places. Specifically the statement "We therefore conclude that there is no significant systematic drift in OMPS LP version 2.6 ozone for the period 2012 to 2024 and that it is suitable for use in ozone trend studies." which may not technically be incorrect, but the numbers presented are 2%/decade from 20-50 km and becoming larger outside. Statements like "We find small drifts ... of less than 2%/decade ..." may not that be useful to potential users of the data trying to attribute changes of 1%/decade, which is the level that most trend studies would be looking at. I think this is solved if a summary table is included with estimated drift levels so the user can decide if it is useful for their application, as well as maybe a softening of the blanket statement that the dataset is suitable for ozone trend studies.

A summary table has been added as mentioned above. We have also added an estimation of the threshold needed to confidently detect trends in the stratosphere as suggested by referee 1 and changed the statement mentioned above to read “We therefore conclude that there is no significant systematic drift in OMPS LP version 2.6 ozone for the period 2012 to 2024 and that OMPS LP data meets current WMO requirements for long-term stratospheric ozone trend studies.” We have also removed the subjective word “small” from the discussion.

Specific Comments

l. 98. There should be a more detailed explanation of how this study differs from the v2.6 validation already done by Kramarova et al. (2024)

We have added the following sentence to section 3:

“This study builds on Kramarova et al. (2024) which compared OMPS LP version 2.6 to MLS for the period 2012-2021 to include other sources of correlative data and extend the evaluation period to April 2024.”

l. 102. The motivation/wording here seems like a repeat of what is at the start of Sec. 6, and the Sec. 6 wording does a better job of motivating how OMPS-LP could be validated going forward. On the same topic, I know it is outside the scope of the paper, but there is no mention anywhere (that I could find) of the potential of validating/cross-calibrating OMPS-LP based on the overlap between NPP and N21 or future satellites.

Although the wording is indeed similar, in this section we are advocating for the use of solar occultation instruments whilst they are available and ozonesondes, whereas in section 6 we are looking forward to a time when there are no other high vertical resolution satellite observations available for which to validate against.

It is true that there should be some overlap between successive OMPS LP instruments and we will use this to cross-calibrate/validate them. We have added the following sentence to the conclusions section:

“There will also be some overlap between successive OMPS LP instruments which we can exploit in order to cross-calibrate/validate them, this will enable us to determine any bias offsets between them.”

l. 112. Here the evaluation period is stated as ending June 2024, but earlier it was April 2024?

This has been corrected to say April 2024

l. 130. At this point the co-location criteria has not been stated, but there are figures showing coincidences which is slightly confusing.

We have added "We only use observations co-located with OMPS measurements (see Section 4)" to this sentence to direct the reader to the co-location criteria.

l. 150. I believe SAGE III/ISS does extend to +/- 70 degrees latitude? albeit for a very limited time of the year

This has been changed to say "approximately 70 degrees"

l. 180. I know in previous studies since Aura and NPP are in similar orbits you can get almost perfect coincidences between MLS and OMPS-LP. Is that is what is done here? Or is the same criteria applied to MLS?

The same coincidence criteria is applied to MLS, however, we always take the closest matching profile within that criteria, so for MLS there are times where there are almost perfect coincidences and these are used.

l. 184. "The only time criterion is that the profiles be on the same day..." is this actually what is done or is it a +/- 12 hour window around the observation time?

In this study we take comparison profiles from that same day, and not from a +/- 12 hour window.

l. 186. "... do not account for the small differences in the vertical resolution..." this is true for MLS and ACE, but the SAGE vertical resolution is significantly better than OMPS-LP

As described above, we tested smoothing the SAGE III/ISS profiles but this had little effect on the biases and drifts and our overall conclusions.

l. 203. "To the 1:30 pm local solar time ..." the measurement time is only 1:30 pm at the equator, presumably you mean to the actual measurement time of OMPS-LP?

That is correct, the reference to 1:30 pm has been removed and the sentence now reads "to adjust both ACE-FTS and SAGE III/ISS observations to the measurement time of OMPS LP"

l. 206. When I read this paragraph I thought it meant that we would only be seeing results from the wide latitude bands, but that is not the case. I would reword it or move it after the 5 degree zonal means comparisons are done

The paragraph has been reworded as below:

"Initially matched profiles were averaged into 5 degree zonal means for comparison. In addition, owing to limited data coverage from correlative solar occultation satellite observations (see Figs. 2 & 3), the data were further averaged into 3 wide latitude bands to increase the number of observations in each bin for comparison statistics. These bands are 30°S-60°S, 30°S-30°N and 30°N-60°N and exclude the polar regions. "

Sec 5.1: I find the systematic difference between MLS and SAGE interesting here, specifically in Figure 4 you could interpret it as MLS being ~5% higher than SAGE almost everywhere. This is quite different than say what is reported in the Wang et al. validation of SAGE. I know it's different data versions etc, and doesn't have anything to do with OMPS-LP, but I'm wondering if you noticed this and is it expected.

The differences seen in Figure 4 indicate that SAGE is higher than MLS by ~5%, since here we are plotting OMPS-MLS and OMPS-SAGE. If we do the same analysis with SAGE version 5.3 (see figure below) then the comparison is closer to MLS, and is consistent with Wang et al. (2020). However, when we switch to SAGE version 6, the SAGE/OMPS/MLS difference increases, and this is consistent with what is stated in the data users guide for SAGE III/ISS version 6 which says that ozone is increased by ~2% in this version.

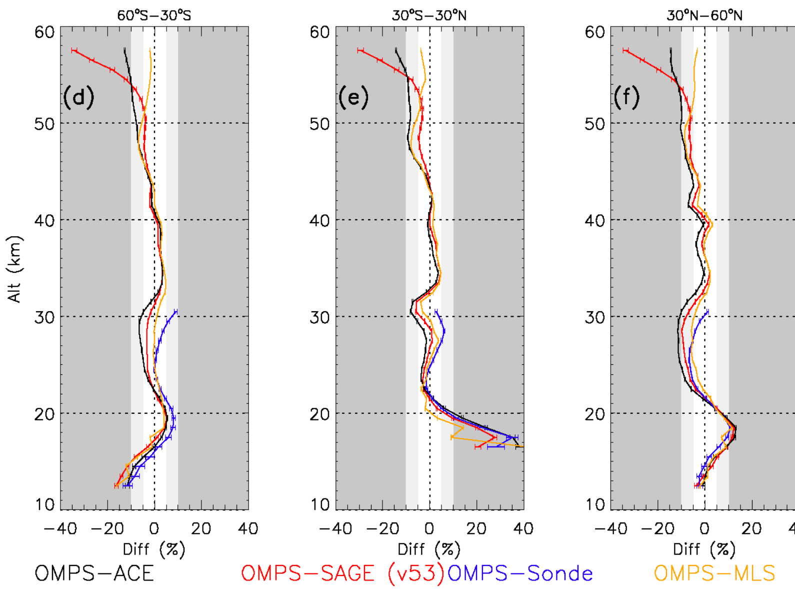

Figure RC2.3: Differences between OMPS LP and correlative measurements using SAGE III/ISS version 5.3 data (red) instead of version 6.

Fig. 4: I find it hard to quantify the differences from the color plots since the scale is so large. Perhaps some black contour lines could be added to the for example +/- 10% difference levels to guide the reader?

The color scale for Figure 4 has been updated to ±30% as suggested by referee 1:

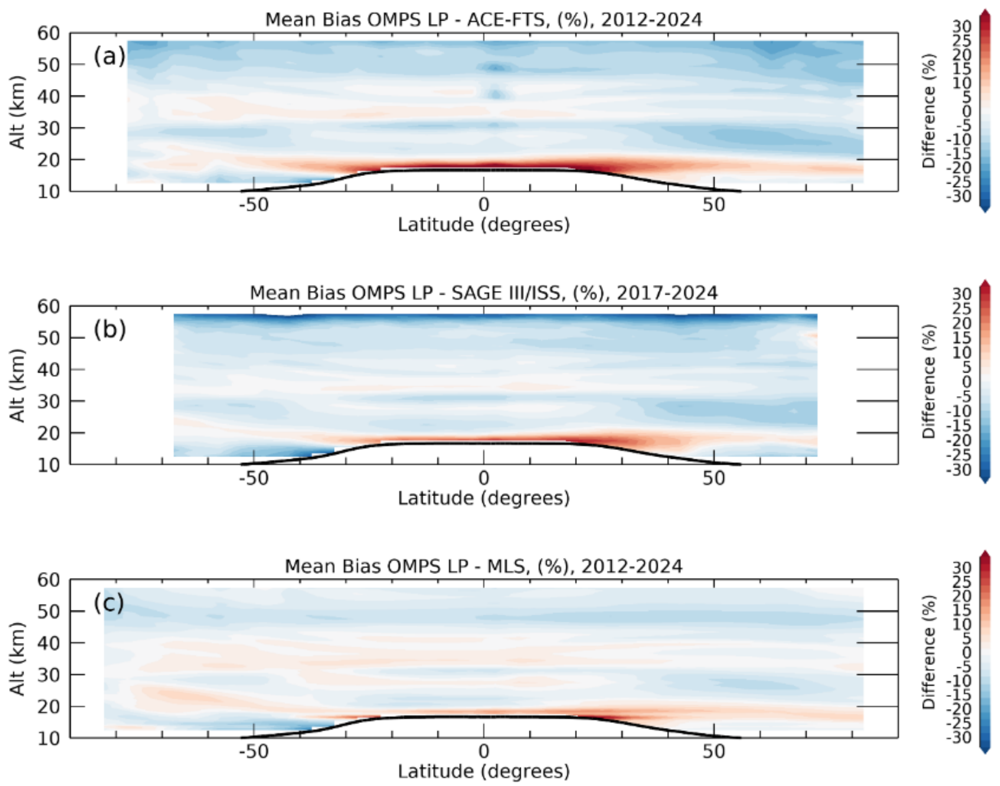

Figure RC2.4: Updated Fig. 4 panels a-c

l. 237. Does the altitude biases correlate with the changing wavelengths used in the retrieval in altitude?

> In version 2.6, the number of UV pairs has been increased to 6, compared with 3 UV pairs used in v2.5. The vertical range where the algorithm suppresses the contributions from shorter wavelengths is now dynamical in v2.6, whereas it was static in v2.5. As a result, we do not see clear changes in bias structures that can be directly attributed to a wavelength switch, except in the tropical middle stratosphere ~ 30-32 km, where the algorithm transitions from the longest UV pair (322/356 nm) to the VIS triplet (510/606/675 nm). However, in this part of the paper we describe vertical pattern of biases in the middle and lower stratosphere, where retrievals solely relay on the only VIS triplet.

l. 250. "degraded precision and increased noise for SAGE III/ISS measurements...": yes, but a lot of this is probably because of using the raw 0.5 km SAGE measurements instead of degrading it down to the resolution of the other measurements

> In our analysis the SAGE III/ISS measurements were interpolated onto a 1 km grid for comparison with OMPS, Wang et al. also noted increased noise and biases in SAGE III/ISS ozone at higher altitudes, however, we agree that there may be a contribution from the differences in vertical resolution, and we have now added the statement “it should be noted that, although we interpolate SAGE III/ISS observations from a 0.5 km to a 1 km vertical grid, we have not degraded the SAGE III/ISS profiles down to the resolution of OMPS LP, and this may also contribute to the lower correlations at higher altitudes” after this line.

Fig. 5: Is there some motivation for the line at 0.75?

This was to be consistent with Figure 11 in Kramarova et al. (2018) which showed correlations for OMPS LP version 2.5, to allow the reader to more easily compare the two figures

l. 307. "However, these biases are not seen when compared to MLS." is a possible explanation uncertainties in the diurnal scaling?

The biases are present both before and after diurnal scaling (see figure below) and so we don't believe that these biases are related to the diurnal correction.

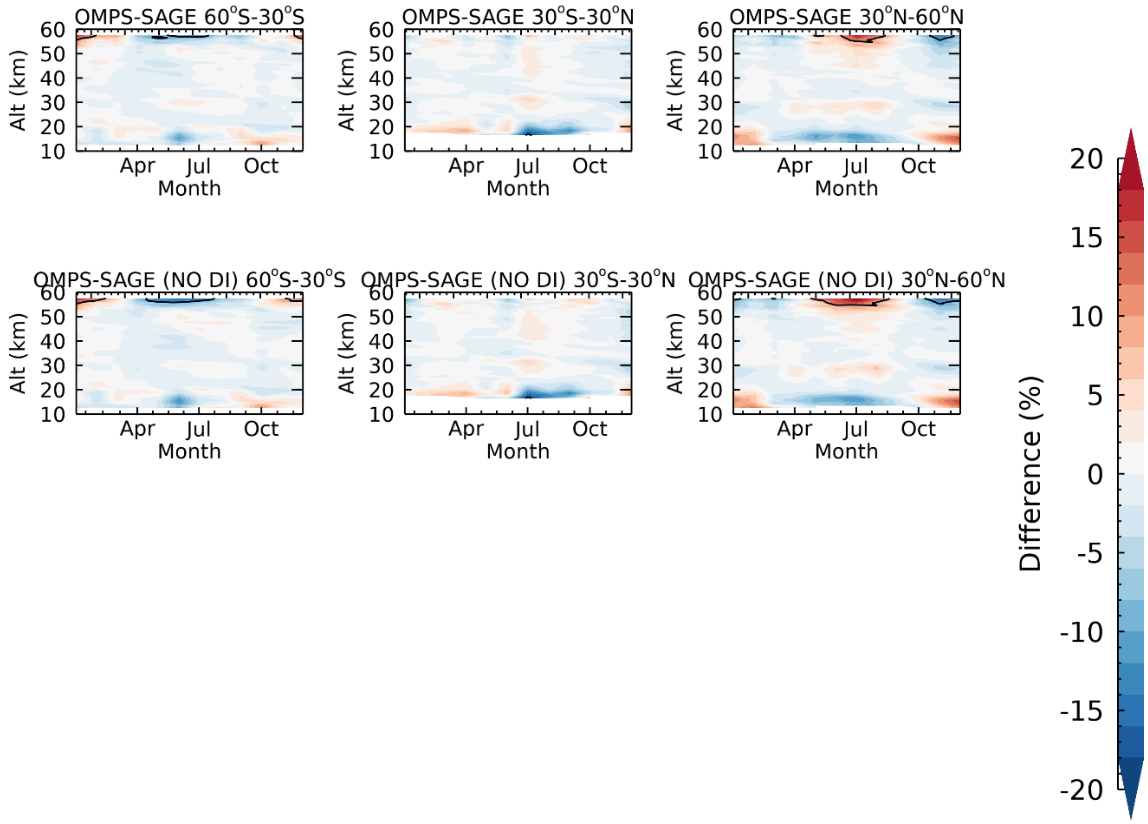

Figure RC2.5: Seasonal cycle biases between OMPS LP and SAGE III/ISS with (top) and without (bottom) diurnal correction (DI).

Figure 8: The scales here are really quite large, here we are extending to +/- 20%/decade, when observable ozone trends are ~1%/decade

The scale in Fig. 8 was optimized to illustrate drifts and associated error bars over the large vertical range. The figure has been updated (see below) to add vertical dashed-dotted lines indicating 0.3%year drifts, which is the WMO stability threshold value for stratospheric ozone trend studies.

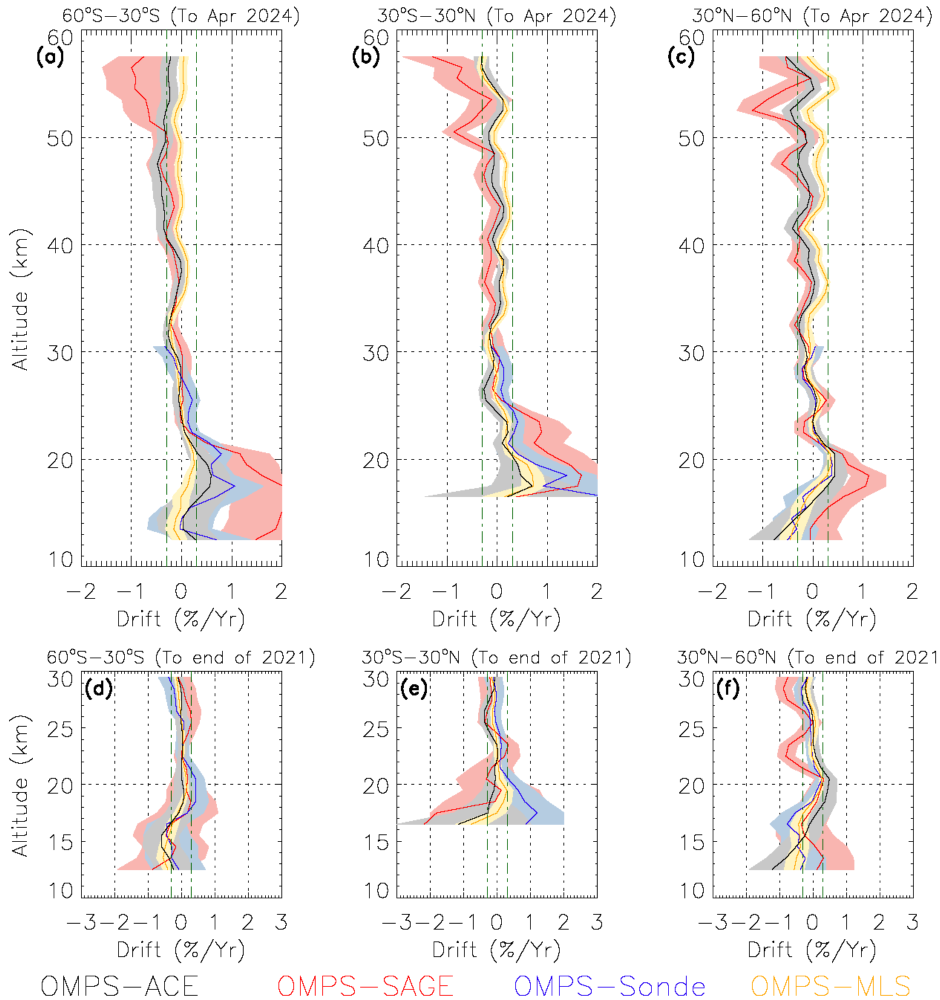

Figure RC2.6: Updated Figure 8 with 0.3%/year lines added.

l. 497. Here the drift is calculated until the end of 2021 at low altitudes, presumably because of the Hunga influence. But this wasn't done in comparisons with the other instruments, so it is an issue with OMPS NP? It seems odd.

The same was done for the other instruments. Panels d-f in Fig. 8 show the drifts calculated to the end of 2021 at low altitudes. This was indeed done because of the Hunga influence and is described in the text in the paragraph beginning at line 348 in the original manuscript.

l. 522. "Previously, we limited our comparisons geographically to exclude polar regions" I think some of the comparisons with MLS previously did extend to 80?

Whilst it is true that some MLS/OMPS differences extended to 80 degrees for the 5 degree zonal mean plot in Fig. 4 panel c, all of the analysis was limited to 60S to 60N.

Figure 12. It is odd here to use two significantly different scales, 10% for the NP comparisons and 50% for the MLS comparisons. Also error bars are only shown for the NP comparisons?

The plot has been updated to rescale the MLS comparisons to be consistent with the NP comparisons (see below). The error bars are there for MLS but they are very small.

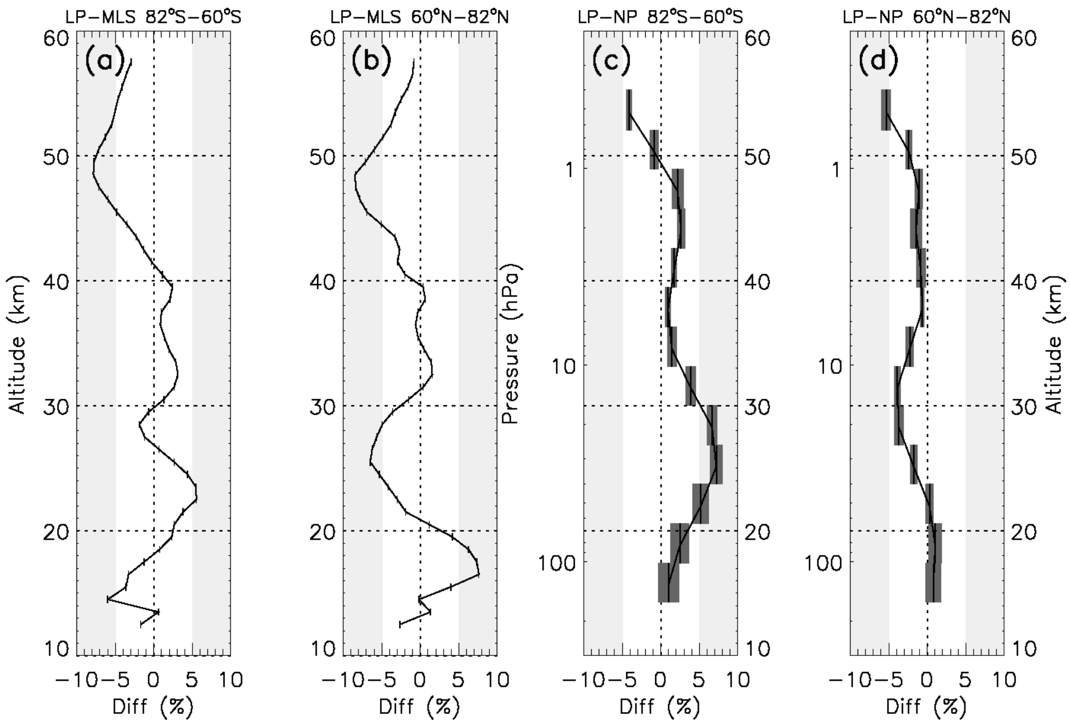

Figure RC2.7: Updated Figure 12.

Figure 13. Here we are also back to using the full time period with NP drift calculations instead of just to 2021

We used the full time period here as there was very little or no effect of the Hunga eruption in the polar regions for comparisons with MLS and NP.

Sec 8. This section is quite short, and all of the information is in the other sections in some shape or form. I would ask the authors to consider if it could be removed.

The section was removed and incorporated into the conclusions section.

Technical Corrections

l. 484. "OMP SNP" -> "OMPS NP" - Corrected