

Response to the comments of Reviewer #1- Manuscript: *Evaluation of ozone trends in the mesosphere/lower thermosphere using a new merged dataset of ozone profiles* by Szelag et. al.

We would like to thank Reviewer #1 for the comments on our paper. Please see our response below.

Major comments

Reviewer: *Please indicate the ozone units used (volume mixing ratio) at the beginning of Section 2, if not before. Also, please indicate what density information was used in the conversion of GOMOS measured ozone density profiles to mixing ratio. Did this follow a similar procedure to that used for the altitude to pressure conversion? If not, are densities in the two conversion processes consistent?*

Authors: We now indicate clearly that we use ozone mixing ratio profiles. The GOMOS mixing ratio is computed consistently with the altitude-pressure-density conversion using measurements from MIPAS (see Sec. 2 in the Supplementary Information). In the revised version, we indicated this explicitly. Note that now we first describe the datasets, then introduce the summary Table 1.

Reviewer: *Why is the seasonal cycle calculated using only a portion of the time series for some instruments (page 12)? What criterion is used to select the range? Is the resulting cycle affected by high variability such as irregular occurrence of sudden stratospheric warmings?*

Authors: Since our merged ozone values are computed as the median of deseasonalized anomalies, we need them aligned. Therefore, we used the common operational period 2005-2011 for GOMOS, MIPAS and MLS. In addition, using the common period for the seasonal cycle does not introduce differences due to different periods for evaluation of seasonal cycle. For ACE-FTS, due to its coarser sampling, we extended the period for evaluation of the seasonal cycle to 2005-2018 (in order to decrease its uncertainty), and assume that the seasonal cycles 2005-2011 and 2005-2018 are nearly the same. For SOFIE and HALOE, we use the full operational period for computation of the seasonal cycle, and their deseasonalized anomalies are offset to 2005-2011 in the merging procedure.

In polar regions, sudden stratospheric warmings cause high variability, and all mean estimates are affected by them.

In the revised version, we added these notes.

Reviewer: *The MLR uses a piecewise linear term for representing trends. Can you provide some justification for this? How sensitive are the results to the precise year specified for the turnaround point? Please show or discuss the trend before 1997.*

Authors: The piecewise linear trend approach is motivated by the well-established ozone turnaround around 1997 following the peak in ozone-depleting substances, which marks the recovery period (Harris et al., 2008). Although trend estimates can be sensitive to the choice of regression period (Harris et al., 2015), the robustness of the trends was assessed by applying the MLR over different analysis windows (Fig. 11 in the revised version). The resulting trends are very similar in both spatial structure and magnitude, indicating only weak sensitivity to the selected starting year.

To assess sensitivity to the trend formulation, additional regressions were performed using the LOTUS regression model (<https://usask-arg.github.io/lotus-regression/>) with an independent linear term (ILT). The ILT includes separate linear trends for the periods before 1997 and after 2000, allowing the two segments to vary independently, in contrast to the PWLT, which enforces continuity at 1997. The resulting ozone trend patterns and magnitudes closely match those obtained with the original MLR, demonstrating that the results are also not sensitive to the chosen trend representation.

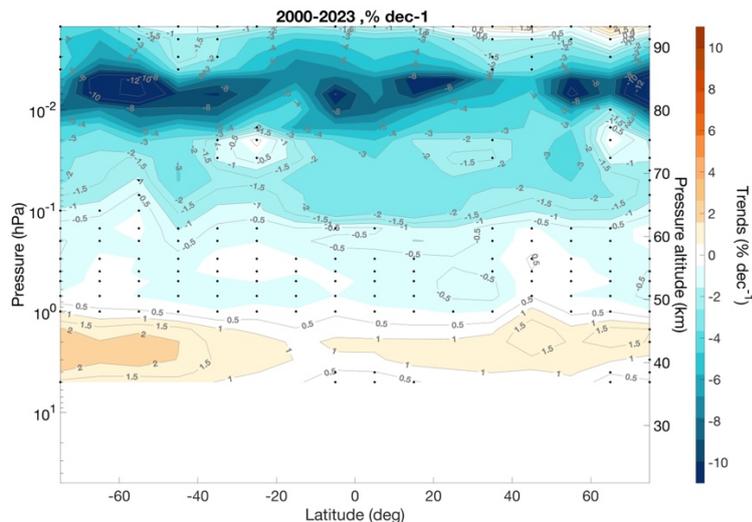


Figure. Latitude-altitude variation of ozone trends for 2000-2023 derived from the METEOR-O₃ dataset using LOTUS regression model with ILT. The black dots indicate trends that are not statistically significant at the 95% confidence level. Trends are given in % per decade.

The period prior to 1997 in the present dataset spans only 1991-1996, which is too short to robustly estimate or meaningfully visualize ozone trends, particularly given the strong interannual variability in mesospheric ozone. For this reason, the pre-1997 trend plot is not shown.

This discussion has now been added to the manuscript, together with an additional figure provided in the Supplement (Figure S6).

Minor comments

Reviewer: (Abstract) “near the mesopause” It’s probably better to give an altitude since the mesopause location is quite variable (around 80-85 km in summer high latitudes; around 95-100 km elsewhere).

Authors: Corrected as suggested by the Reviewer.

Reviewer: (p. 8; line 96-04) Information on SOFIE is absent.

Authors: Information about SOFIE is now added.

Reviewer: (p. 8; line 19, and elsewhere throughout the text) “wintertime (DJF)” It is recommended that you avoid a northern hemisphere bias by only using “wintertime” when presenting and discussing NH latitudes.

Authors: DJF is now indicated by *boreal winter* instead of wintertime.

Reviewer: (p. 13, line 297) what weighting is used?

Authors: The weights are inversely proportional to the estimated uncertainties of the deseasonalized anomalies. This is clarified in the next sentence. No changes to the text were made in response to this comment.

Reviewer: (p. 14, line 10) Please provide more justification for why HALOE and SOFIE are treated differently than the other datasets. The idea of adjusting for an offset is particularly a concern in the case of HALOE because its time series is so important for the calculation of the day trend.

Authors: HALOE and SOFIE did not measure in the period 2005-2011 (see also the explanation of the selection of the period for computation of deseasonalized anomalies above).

Reviewer: *I recommend moving Figure S4 from the Supplement to the main paper.*

Authors: We have now added Figure S4 to the main paper (Figure 12) following Figure 11 (previously Figure 13).

Note that we have removed the text in lines 91-98 discussing ozone-temperature relationships and temperature trends. Since temperature trends are not analyzed in this study and are not shown in the manuscript, this discussion is outside the scope of the present work. A combined analysis of ozone and temperature trends will be addressed in a future study.

Reviewer: *In the Summary section, please indicate that the trends are calculated on pressure levels.*

Authors: We now plot trends on the pressure levels and indicate this in the Summary. We also added how the altitudes are computed in the text: Pressure altitudes are computed from pressure as $z = 16 \log_{10}(1013/P)$, where P is pressure in hPa. This pressure-derived altitude coordinate is used throughout the manuscript. The colorbar range is set to -11 to 11 ($\% \text{ dec}^{-1}$).

References

Harris, N. R. P., Kyrö, E., Staehelin, J., Brunner, D., Andersen, S.-B., Godin-Beekmann, S., Dhomse, S., Hadjinicolaou, P., Hansen, G., Isaksen, I., Jrrar, A., Karpetchko, A., Kivi, R., Knudsen, B., Krizan, P., Lastovicka, J., Maeder, J., Orsolini, Y., Pyle, J. A., Rex, M., Vanicek, K., Weber, M., Wohltmann, I., Zanis, P., and Zerefos, C.: Ozone trends at northern mid- and high latitudes – a European perspective, *Ann. Geophys.*, 26, 1207–1220, <https://doi.org/10.5194/angeo-26-1207-2008>, 2008.

Harris, N. R. P., Hassler, B., Tummon, F., Bodeker, G. E., Hubert, D., Petropavlovskikh, I., Steinbrecht, W., Anderson, J., Bhartia, P. K., Boone, C. D., Bourassa, A., Davis, S. M., Degenstein, D., Delcloo, A., Frith, S. M., Froidevaux, L., Godin-Beekmann, S., Jones, N., Kurylo, M. J., Kyrölä, E., Laine, M., Leblanc, S. T., Lambert, J.-C., Liley, B., Mahieu, E., Maycock, A., de Mazière, M., Parrish, A., Querel, R., Rosenlof, K. H., Roth, C., Sioris, C., Staehelin, J., Stolarski, R. S., Stübi, R., Tamminen, J., Vigouroux, C., Walker, K. A., Wang, H. J., Wild, J., and Zawodny, J. M.: Past changes in the vertical distribution of ozone – Part 3: Analysis and interpretation of trends, *Atmos. Chem. Phys.*, 15, 9965–9982, <https://doi.org/10.5194/acp-15-9965-2015>, 2015.

Response to the comments of Reviewer #2- Manuscript: *Evaluation of ozone trends in the mesosphere/lower thermosphere using a new merged dataset of ozone profiles* by Szelag et. al.

We would like to thank Reviewer #2 for the comments and the effort that the reviewer has put into this paper. Please see our response below.

General comments

Overall, the methodology described appears to be consistent with that used for previous merged datasets. However, I have a few important concerns related to the methodology (see #1.1-1.3 below) which should be addressed. Besides those issues, there are several points where the presentation of the methodology needs major clarifications (see #1.4-1.8). I am also providing many minor comments and technical corrections (see #2).

Specific comments:

Reviewer: *Why start at 22 km if this is a MLT dataset? If the full range is provided, isn't this more of a stratosphere-to-LT dataset? Calling it a MLT dataset may be underselling its value, unless there are considerations when utilizing data from the stratosphere?*

Authors: The new dataset was originally designed with a primary focus on the mesosphere and lower thermosphere (MLT), which motivated the initial terminology. However, the quality and spatio-temporal coverage of the merged observations allowed us to extend the dataset into the stratosphere. As a result, the dataset now spans the middle atmosphere, from the stratosphere through the mesosphere and into the lower thermosphere.

We agree that referring to the dataset solely as an MLT dataset may underrepresent its value. Accordingly, the manuscript has been revised to describe the dataset as a middle-atmosphere dataset, while the title retains emphasis on the first estimate of mesospheric ozone trends, which remains a key novel aspect of this work.

Reviewer: *The merging procedure relies on median values. Given that medians are calculated from the deseasonalized anomalies, I would assume these are each effectively close to 0? However, medians are then matched across varying time periods. I think this procedure may be problematic given the region where the negative trends are found later in the text (e.g. Fig. 13) is at altitudes where the anomalies are also close to 0. To reduce artefacts of the merge, I would expect medians to be aligned over overlapping time periods only. Perhaps further clarity on the description of the procedure in Section 4.3 would address this choice of methodology.*

Authors: The anomalies are illustrated in Figures 6-9 (old version of the manuscript, in the revised version Fig.5-7 and S5); they are not necessarily close to zero. The anomalies from individual datasets are aligned. The median (over individual datasets) is computed for each latitude zone, pressure level and month. In the revised version, we added an equation for the merged anomaly.

Reviewer: *I would expect the ozone profiles in Section 5 to be calculated using all datasets, same as the merged anomalies. There is no clear reason why all datasets couldn't be used for the merged seasonal cycle.*

Authors: The original datasets correspond to different illumination conditions, and the sampling density is substantially different. To restore the seasonal cycle, we use the data from two dense samplers, MLS and MIPAS, which provide the best estimation of the seasonal cycle. A similar approach is applied for the merged stratospheric ozone datasets (Sofieva et al., 2017, 2021, 2023).

Reviewer: *Perhaps related, the regions of significance in Fig. S4 which provide ozone trends in ppmv vary notably from those in Fig. 13. This suggests that the ppmv profiles and their uncertainties are impacted by the methodology choices.*

Authors: The difference between the regions of significance in Fig. S4 (we moved S4 to the main manuscript as Figure 12) and Fig. 13 (Figure 11 in the revised manuscript) was caused by an inconsistency in the implementation of the trend calculation for the ppmv profiles, which has now been corrected. After fixing this issue and rerunning the analysis, the regions of significance are consistent. In addition, we have extended the altitude range of the ppmv trend analysis to match that used in the %/decade.

Note that we have removed the text in lines 91-98 discussing ozone-temperature relationships and temperature trends. Since temperature trends are not analyzed in this study and are not shown in the manuscript, this discussion is outside the scope of the present work. A combined analysis of ozone and temperature trends will be addressed in a future study.

We now plot trends on the pressure levels and indicate this in the Summary. We also added how the altitudes are computed in the text: Pressure altitudes are computed from pressure as $z = 16 \log_{10}(1013/P)$, where P is pressure in hPa. This pressure-derived altitude coordinate is used throughout the manuscript. The colorbar range for the trends is set to -11 to 11 (% dec⁻¹).

Reviewer: *Table 1: The caption in Table 1 states “Information about the datasets used in the merged dataset”, however, this table only lists information about the instruments. It is important to distinguish between full instrument dates, vertical range, etc. and those selected for inclusion when creating the datasets in this study. The reader is likely to be most interested to know the ranges used for METEOR-O3. A good example is MIPAS, which appears to be split into two separate time ranges (with only the latter used) and two different modes.*

Authors: We agree that the original caption and structure could be misleading because it mixed general instrument characteristics with the subsets used in the merged product. We have now revised Table 1 to include the information relevant to this study, i.e. used in the METEOR-O3 dataset.

Reviewer: *Table 1: Similarly, I suggest starting Table 1 with a column for dataset name, as these are never defined (i.e. MLS, MIPAS_NOM, MIPAS_MA_UA, ...). Following this, the term “dataset” should be used throughout the text, rather than “instrument” or “sensors”, when discussing the data used for the merge.*

Authors: The manuscript has been revised to introduce dataset names and descriptions and definitions prior to Table 1. In addition, the terminology has been harmonized throughout the manuscript, and the term “dataset” is now used consistently in place of “instrument” or “sensor” where appropriate.

Reviewer: *Table 1: The column called “Local time/equator crossing time” is hard to reconcile with the illumination conditions referred to later. I would recommend changing this column to something like, “Illumination conditions” and listing “daytime, nighttime, sunset, sunrise” accordingly. This would clarify later text such as L. 208-209 where GOMOS is not listed.*

Authors: The column labeled “Local time / equator crossing time” has been changed to “Illumination conditions”, and the entries now explicitly indicate the relevant observing conditions (daytime, nighttime, sunrise, sunset) for each dataset. The equator crossing time is now added to the general data description.

Reviewer: *Table 1: There are numerous discrepancies between the values listed in Table 1 and those in the text from Section 2.*

Authors: Table 1 has been revised to summarize the datasets specifically used in METEOR-O3, whereas the text provides a general description.

Reviewer: Fig. 2-4: *The purpose of the text starting at L. 219 (and Figs 2-4) should be clarified. These appear to be a series of validations to ensure the datasets generally match with each other. But there is also a large portion of text providing background on the general distribution of ozone. Is the goal of Fig. 2 and 3 to ensure that each dataset shows the “key patterns” in ozone?*

Reviewer: Fig. 2-4: *I would recommend adding a separate section header to indicate the goals of this text/figures or add more clarity in the text.*

Authors: Yes, exactly. The primary goal of Figures 2-4 (in the revised manuscript Fig. 2-3 and S4) and the accompanying text is to demonstrate the consistency of the individual datasets prior to merging, by verifying that each reproduces the well-established vertical structure, diurnal behaviour, and seasonal characteristics. We have revised the text to clarify the purpose of Figures 2-3 to clearly state the objectives of this analysis. We add the following sentence to the manuscript: *To assess dataset consistency prior to merging, Figures 2-3 demonstrate how well the various individual datasets reproduce the vertical structure and diurnal behaviour of ozone in the stratosphere and MLT.* Figure 4 has been moved to the Supplement as Fig. S4.

Reviewer: Fig. 2-3: *Similarly, the datasets within Fig. 2 are limited to the same time period, but this is not the case for Fig. 3 (understandable, given the time period of ACE). It is unclear whether these choices are made for intercomparison purposes, or to match the years selected for the merge, or a mix of both. This makes it hard to know what the reader is meant to get out of these figures.*

Authors: For Figure 2, the period 2005-2011 was selected for intercomparison purposes, as it represents a time window with overlapping observations from MLS, MIPAS_NOM, MIPAS_MA_UA, and GOMOS. For Figure 3, the same approach is not possible because the time spans of the solar occultation instruments differ. Only ACE-FTS overlaps with the 2005-2011 period, whereas HALOE and SOFIE operate largely outside this window. Therefore, climatological distributions for HALOE and SOFIE are shown using their full available periods, while ACE-FTS is shown for the overlapping years, consistent with Figure 2. The text has been revised to explicitly state these choices.

Reviewer: *Many variables used are not explicitly defined in the text. This is crucial for readers to understand the methodology. Please provide all definitions across each equation (e.g. $\rho(z)$, k , z , $\Delta(t)$, $\sigma\Delta$, ...)*

Authors: We have carefully revised the manuscript to ensure that all variables used in the equations are now explicitly defined in the text.

Reviewer: *Eq. 2: s appears to be defined two different ways. This should be clarified. If the robust estimator is used, a citation should be provided.*

Authors: We now clarify this section as follows: where s^2 is the sample variance. In Eq. (2), we used a robust estimator for the sample variance (similarly to the approach taken by Sofieva et al., 2017), i. e., $s = 0.5(P_{84} - P_{16})$, where P_{84} and P_{16} are the 84th and 16th percentiles of the distribution, respectively.

Reviewer: *Eq. 2: Is $\sigma\rho$ as written here used anywhere?*

Authors: Yes, it is used for uncertainty estimates, see Eq. (4).

Reviewer: *Eq. 3: $\rho(z)$ is the monthly zonal average term from Eq. 1 – is this not the same as ρm ? And if they are the same, why is ρm defined differently?*

Authors: Yes, $\rho(z)$ is not the same as ρm , as stated in the manuscript (see Eq.(3) and explanation of variables after it). In the revised version, we added that ρm is the *climatological* mean. Furthermore, the equation for the seasonal cycle now has a number. We now first provide formulae for the seasonal cycle and its uncertainty, and then introduce deseasonalized anomalies.

Reviewer: *Eq. 3: Is it correct to interpret t_i as time in years? If so, this should be stated. Also, what does i mean here, and is it necessary?*

Authors: The time variable in Eq. (3) refers to monthly time, not time expressed in years. The index i was originally used only to enumerate individual monthly mean values. To avoid ambiguity, we have revised the notation by removing the index i and now denote time simply as t , which is explicitly defined as time in months.

Reviewer: *Eq. 4: The uncertainty of ρm should be the same as σ_{ρ^2} right? If so, why is it defined differently in Eq. 4, with σ_{ρ^2} now corresponding to the uncertainty for a single year's monthly mean?*

Authors: No, the uncertainty of ρm is not the same as σ_{ρ^2} . The σ_{ρ^2} describes the variance of individual monthly mean values, while the uncertainty ρm is the uncertainty of the climatological monthly mean (seasonal cycle), which is estimated using error propagation. This section has now been reorganized (see previous comment).

Reviewer: *Generally, the introduction of ρm and $\Delta(t_i)$ is confusing. Is ρm what is shown in Fig. 5? And this is the same as $\rho(z)$? It may be helpful to show Fig. 5 along with the definition of ρm and its uncertainty and afterward define $\Delta(t_i)$ (i.e. reverse the order of Eq. 3 and 4 and more explicitly indicate that Fig. 5 shows ρm)*

Authors: $\rho(z)$ is not the same as ρm (see the detailed explanation above). Yes, ρm is shown in Fig. 5 (in the revised version it is Figure 4). As already mentioned above, in the revised version we first provide the equations for the seasonal cycle and its uncertainty, discuss Fig. 5 (now Figure 4), and then introduce deseasonalized anomalies. The revised text explicitly states that Fig. 5 (now Figure 4) shows ρm .

Reviewer: *Unless I'm missing something, the data shown in Fig. 5 is essentially the same as that from Fig. 2-4, with different axes/selections? I suggest reorganizing all sections between L. 210 and L. 292 to make this section clearer and easier to follow.*

Authors: Yes, the figures are related, but they serve different and complementary purposes (see previous comment). The primary goal of Figs. 2-4 (in the old version of the manuscript) and the accompanying text is to demonstrate the consistency of the individual datasets prior to merging. Specifically, Figs. 2 and 3 show multi-year averaged boreal wintertime (DJF) distributions. Figure 5 (in the old version for the manuscript), in contrast, represents the annual cycle of climatological mean ozone. Unlike Figs. 2-4, which display results from individual instruments separately, Fig. 5 allows direct comparison across latitude bands and altitudes. We have revised the text to better clarify the purpose of Figures 2-4 to clearly describe the goals of this analysis (see previous comment). Figure 4 was moved to Supplement, so what was previously Figure 5 is Figure 4 in the revised manuscript.

Reviewer: *Eq. 6 Extra “)” in several terms*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *According to Table 1, Section 2.6, and Figures 1 and 3, SOFIE data runs from 2007-2023. However, in Figures 8, 9, and 10, SOFIE data appears to run from 2007- 2016 within the merged dataset. This is not explained anywhere in the text.*

Authors: As shown in Supplementary Figure S1, the sampling pattern by sunset/sunrise measurements changes around 2016. So, for the Southern Hemisphere high-latitudes, the SOFIE anomaly data extend until 2016, while they extend until 2023 in the Northern Hemisphere. In the revised version, we added an explanation of the change in the SOFIE sampling pattern in Section 2.

Reviewer: *For the seasonal cycle calculations, why is 2005-2011 used for MLS, when 2005-2018 is used for ACE-FTS? When taking the median of pre-merged datasets (Section 4.3), are the full time periods used? (i.e. for MLS, is 2005-2011 used, or 2005-2023?) The seasonal cycle for ACE_FTS is also calculated “using a different time period” - why? Even with Fig. 8, it’s quite difficult to fully understand the procedure here.*

Authors: Since our merged ozone values are computed as the median of deseasonalized anomalies, we need them aligned. Therefore, we used the common operational period 2005-2011 for GOMOS, MIPAS and MLS. In addition, using the common period for the seasonal cycle does not introduce differences due to different periods for evaluation of seasonal cycle. For ACE-FTS, due to its coarser sampling, we extended the period for evaluation of the seasonal cycle to 2005-2018 (in order to decrease its uncertainty), and we assume that the seasonal cycles 2005-2011 and 2005-2018 are nearly the same. For SOFIE and HALOE, we use the full operational period for computation of the seasonal cycle, and their deseasonalized anomalies are offset to 2005-2011 in the merging procedure. In the revised version, we added these notes.

Reviewer: *As far as I can tell, regions of negative trend occur where ozone concentrations are very low and represent very small changes. However, these results do generally appear significant and seem useful to report. Including S4 for context following Fig. 13 would help the reader gauge the magnitude of the detected changes in ppmv.*

Authors: To better place the percentage trends in context, we have moved Figure S4 into the main manuscript (Figure 12 in the revised manuscript) immediately following Figure 11 (Figure 13 in the old version of the manuscript).

Reviewer: *I expected to see trends at either the secondary or tertiary maxima. However, the strongest region of negative trend in Fig. 13 appears to be just below the secondary maximum, although this is unclear. Given that these maxima and their drivers were introduced on page 9, I would expect the discussion of the trends in Fig. 13 to refer to those locations specifically.*

Authors: Because trends are expressed in % in Fig. 13 (Figure 11 in the revised manuscript), regions with lower background ozone can exhibit larger relative changes. In the discussion of the trends in Fig. 13 (Figure 11 in the revised manuscript) we now refer to those locations.

Reviewer: *METEOR_O3 (subscript) vs. METEOR-O3 inconsistency throughout*

Authors: We now use METEOR-O₃ consistently throughout the manuscript.

Reviewer: *Why does METEOR-O3 end in 2023, when some instruments are still operational?*

Authors: The METEOR-O₃ dataset currently extends to 2023 because this study was carried out within the framework of the METEOR project, which covered data processing up to that year. While some contributing instruments remain operational beyond 2023, data after this period were not included in the present version of the merged dataset to ensure consistency and completeness within the project scope. We plan to extend METEOR-O₃ to include more recent observations in future updates.

Reviewer: *Paragraph 1: I suggest defining the altitude range of the MLT in the first sentence. More citations would support the statements in the first two sentences. “Over the last years” → “Over the last few years”?*

Authors: The altitude range of the MLT has been defined in the revised manuscript. The first sentences we now support by two review articles that synthesize the relevant literature, therefore, we added (Laštovička, 2017, 2023, and references therein) to make this explicit.

Reviewer: *Paragraph 2: Statements about “cooling” are very general. Where is this cooling found? And should this be discussed as a “trend” rather than a “signal”?*

Authors: We now correct the statement as follows: The middle and upper atmosphere is experiencing a global mean negative temperature trend observed across datasets and models caused by the increasing concentration of CO₂.

Reviewer: *Paragraph 4: More citations are needed to support the statements in the sentences beginning on L. 54 and L. 59.*

Authors: The references are added to support our statements as follows:

Mesospheric ozone plays an important role in atmospheric chemistry, dynamics, and energy balance, yet it remains one of the least-investigated parameters (Laštovička, 2023). Variations in its concentration can substantially influence the composition and behavior of the upper and middle atmosphere, thereby affecting the coupling between different atmospheric layers (Sinnhuber et al., 2012; Seppälä et al., 2009, 2024). In recent years, there has been a growing scientific interest in understanding solar-driven changes in ozone and temperature in the mesosphere and upper stratosphere (Salminen et al., 2019; Szélag et al., 2022; Seppälä et al. 2024).

Reviewer: *L. 79: “we used the limb and occultation instruments that provide data in the mesosphere and the lower thermosphere” → Limb and occultation instruments that observe ozone specifically*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *L. 83: “Note that for some instruments, the selected period is shorter than their full operational period” → with Table 1 as-is, it is hard to differentiate between operational and selected periods, and why these vary.*

Authors: We have now revised Table 1 to include the information relevant to this study.

Reviewer: *L. 84-86: What is the relevance of instruments with different illumination conditions having comparable numbers of profiles?*

Authors: This is just the fact about these datasets. This indicates that the datasets provide comparable temporal sampling, minimizing potential sampling biases.

Reviewer: *Table 1: MA UA and NOM not defined in Table 1 caption. Neither is the notation for “20(40)”*

Authors: The manuscript has been revised so that the dataset descriptions and definitions are introduced first - before Table 1 (see previous comment). That way all terms used in summary Table 1 are defined.

Reviewer: *L. 116: Please define UTLS*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *“On-board” vs. “on board” inconsistency throughout*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *L. 164: “0-19 %” and the citation for this statement should be verified*

Authors: We corrected the statement as follows: A recent validation study (Sheese et al., 2022) found that v4.1 ACE-FTS ozone (the latest version for which validation has been performed) exhibits a bias relative to other instruments of approximately -1% to $+5\%$ in the lower stratosphere, $+2$ to $+9\%$ in the middle stratosphere, and up to about $+15\%$ in the upper stratosphere. The estimated precision of v4.1 ozone is $\sim 6\text{--}10\%$ between 20 and 45 km and $\sim 5\text{--}10\%$ above 45 km.

Reviewer: *Section 2.5: Discussing uncertainty for MLS as % rather than ppmv would be more consistent with how uncertainty is discussed in the other datasets.*

Authors: We now provide uncertainty for MLS as % in brackets as follows: The systematic uncertainty of version 5 ozone data is in the range 0.1 to 0.3 ppmv over most of the stratosphere and mesosphere (on average 7-25%) but increases to 0.9 ppmv ($\sim 30\%$) at 0.001 hPa (~ 95 km).

Reviewer: *L. 191: What are “the instrument effects”?*

Authors: Most retrieval algorithms use radiative transfer simulations and incorporate various instrumental characteristics. We removed this sentence and added the reference Gordley et al., 2009b to the previous sentence, as the referenced paper already contains this information.

Reviewer: *L. 197: MLS is not the “longest dataset”, ACE-FTS is*

Authors: We agree that MLS is not the longest dataset in terms of temporal coverage. The text has been revised to clarify that MLS is the longest dataset with dense sampling, rather than the longest dataset overall.

Reviewer: *DJF should be indicated as “boreal wintertime” rather than “wintertime” throughout, as this dataset covers both hemisphere*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *All figures: The datasets labelled in the legend should match the dataset names throughout (e.g. MIPAS_MA_UA)*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *L. 225-227: Citation needed*

Authors: We now add the citation: Brasseur and Solomon (2005).

Reviewer: *L. 252: “mid-latitudes” → “Northern Hemisphere mid-latitudes”?*

Authors: Corrected as suggested to “at northern mid-latitudes”.

Reviewer: *L. 291: “earlier validation and intercomparison results” → specifically which?*

Authors: We added references (Smith et al., 2013; Lopez-Puertas et al., 2023; Das et al., 2023)

Smith, A. K., Harvey, V. L., Mlynczak, M. G., Funke, B., Garcia-Comas, M., Hervig, M., Kaufmann, M., Kyrölä, E., López-Puertas, M., McDade, I., Randall, C. E., Russell, J. M., Sheese, P. E., Shiotani, M., Skinner, W. R., Suzuki, M., and Walker, K. A.: Satellite observations of ozone in the upper mesosphere, *J. Geophys. Res. Atmospheres*, 118, 5803–5821, <https://doi.org/10.1002/jgrd.50445>, 2013.

Manuel López-Puertas, Maya García-Comas, Bernd Funke, Thomas von Clarmann, Norbert Glatthor, Udo Grabowski, Sylvia Kellmann, Michael Kiefer, Alexandra Laeng, Andrea Linden, Gabriele P. Stiller, MIPAS ozone retrieval version 8: middle-atmosphere measurements, *Atmospheric Measurement Techniques*, 10.5194/amt-16-5609-2023, **16**, 22, (5609-5645), (2023).

S. Das, S. M. Bailey, M. E. Hervig, B. Thuraiajah, B. T. Marshall, Validation of Version 1.3 Ozone Measured by the SOFIE Instrument, *Earth and Space Science*, 10.1029/2022EA002649, **10**, 10, (2023).

Reviewer: *L. 291-292: Regarding the explanation for GOMOS's bias, it may be helpful to point the reader back to Fig. 1 and Section 2.2.*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *Fig. 6: Legends should say "daytime" and "nighttime" for consistency. The vertical range on the y-axis could be reduced in the bottom panel.*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *Section 4.2: Fig. 6 and 7 are not formally introduced (especially Fig. 6 which is first referenced regarding ozone's diurnal cycle in L. 295) which makes it hard to tell the goals of these figures. Is it a true statement these show an example of the pre-merge from MIPAS_MA_UA from a couple of viewpoints?*

Reviewer: *And can it be assumed that this was a validation procedure performed across all datasets and the same conclusions held? A statement to this effect would be important to add.*

Authors: Figures 6 and 7 (old version of the manuscript) illustrate the pre-merging procedure using MIPAS_MA_UA as an example from two complementary perspectives. Figure 6 (Figure 5 in the revised manuscript) shows a selected latitude, while Figure 7 (Figure S5 in the revised manuscript) presents the deseasonalized anomalies and the corresponding weighted-mean anomaly at each altitude level. Because Figure 7 was an expansion of Figure 6, it has now been moved to the Supplement (Fig. S5).

We now introduce Figure 6 (now Figure 5) properly. The paragraph 4.2 was revised as follows:

Ozone in the mesosphere exhibits strong diurnal variations, as discussed above and illustrated in Figures 2 and 3. This behaviour is also evident in the top panel of Figure 5, which shows examples of absolute ozone mixing ratios from MIPAS_MA_UA at 56 km and 5° latitude under different illumination conditions. While the absolute ozone mixing ratios differ substantially between illumination conditions, the deseasonalized anomalies (bottom panel of Figure 5) are nearly identical for instruments with dense sampling. The daytime, nighttime, and weighted-mean deseasonalized anomalies are highly consistent across the full altitude range of the MIPAS_MA_UA measurements, as shown in Supplementary Fig. S5. This validation procedure was performed across all datasets. This allows us to consider the weighted mean of daytime and nighttime deseasonalized anomalies as the aggregated (we will call them "pre-merged" hereafter) deseasonalized anomaly from each instrument.

This was a validation procedure performed across all datasets. We now state it clearly in the manuscript.

Reviewer: *L. 312: "illustrated" → "illustrates"*

Authors: Corrected as suggested by the Reviewer.

Reviewer: *Fig. 9, S3, S4: Missing unit on the colorbars*

Authors: Corrected as suggested by the Reviewer.

Reviewer: L. 346-347: “merged deseasonalized anomalies (Eq. 3)” → *I’m not sure that Eq. 3 is the correct reference, as the deseasonalized anomalies have not been merged in that equation.*

Authors: We have revised these lines as follows: For this purpose, we computed the merged seasonal cycle $\rho_m(\text{merged})$ as the mean of the seasonal cycles from MLS, MIPAS_NOM, and MIPAS_MA_UA for daytime and nighttime illumination conditions. These datasets were selected because of their dense sampling and to ensure maximum spatial coverage. This merged seasonal cycle is then applied to the merged deseasonalized anomalies.

Reviewer: L. 374: “60-80 km” → “60-90 km”?

Authors: Here, we distinguish between two altitude regions: smaller negative trends are observed below 80 km, whereas larger negative trends occur above this altitude. No change was made to the text in response to this comment.

Reviewer: L. 401: “2923” → “2023”

Authors: Corrected as suggested by the Reviewer.

Reviewer: *The “recovery period” is defined both 1997-2023 and 2000-2023. This should be consistent, also with a reference for the choice of year.*

Authors: In the revised manuscript, the recovery period is defined as starting in 1997, based on the well-established ozone turnaround following the peak in ozone-depleting substances. The period 2004-2023 is not treated as an alternative definition of the recovery period, but is used solely as a sensitivity test, corresponding to the time interval with the best spatio-temporal data coverage. The text has been revised to clearly distinguish between them.

Reviewer: *Supplement L. 36: “MIPAS_MA_-UA” → “MIPAS_MA_UA”*

Authors: Corrected as suggested by the Reviewer.