Dear Editor,

Thank you very much for your attention to our paper. We thank the reviewers for their positive evaluation and comments on our paper. Their comments are taken into account in the revised manuscript.

Please find below the response letters to the reviewers' comments (the same letters are posted in the interactive discussion). The revised manuscript with the modifications marked by "tracked changes" is also uploaded.

As a corresponding author, I confirm that all co-authors concur with the submission in its revised form.

Yours sincerely,

Viktoria Sofieva, Dr., Adj. Prof.
Finnish Meteorological Institute,
Space and Earth Observation Centre
P.O. Box 503 (Erik Palmenin aukio, 1)
FIN-00101 Helsinki Finland
tel: +358 29 539 4698
fax: +358 29 539 3146
email: viktoria.sofieva@fmi.fi

Review #1

Dear Reviewer,

Thank you very much for your positive evaluation of our paper. We took your comments into account in the revised version of the manuscript. Please find below our detailed replies (black font) on your comments (blue font).

Reviewer#1 comments:

However, I think it would be even more useful if information about scientific contexts to the ozone changes are also included throughout this work. For instance, would atmospheric phenomena, such as large volcanic eruptions or wildfires, affect the ozone trends between 2020-2024? Was the ozone trend expected to be consistent between 2000-2020 and 2000-2024? What are the physical processes controlling the ozone trend in the upper/mid stratosphere and troposphere? What are the expectations for the ozone trends in the future? I think adding some of the scientific backgrounds of ozone trends would make this work even more valuable.

Thank you very much for the suggestion. In the revised version, we added general information about ozone recovery and what are expected estimates on ozone trends in different periods.

Specific Comments:

P2, Abstract: Adding some comments about the significance of this work in the context of the importance of long-term measurements of ozone, impact on climate, and potential improvements of atmospheric processes represented in global models would make the abstract more compelling.

Unfortunately, the abstract is limited to 250 words, so it is impossible to add details without cancelling other important information.

P2, L55: Antarctic ozone hole (Farman et al., 1985).

The reference is added.

P2, L56: References should be chronologically ordered. (?)

The references are now in chronological order.

P2, L57-60: Consider including citations supporting the information listed here.

We included the following references:

Keeble, J., Abraham, N. L., Archibald, A. T., Chipperfield, M. P., Dhomse, S., Griffiths, P. T., and Pyle, J. A.: Modelling the potential impacts of the recent, unexpected increase in CFC-11 emissions on

total column ozone recovery, Atmos. Chem. Phys., 20, 7153–7166, https://doi.org/10.5194/acp-20-7153-2020, 2020.

Solomon, S., Stone, K., Yu, P., Murphy, D. M., Kinnison, D., Ravishankara, A. R., and Wang, P.: Chlorine activation and enhanced ozone depletion induced by wildfire aerosol, Nature, 615, 259–264, https://doi.org/10.1038/s41586-022-05683-0, 2023.

Wohltmann, I., Santee, M. L., Manney, G. L., and Millán, L. F.: The Chemical Effect of Increased Water Vapor From the Hunga Tonga-Hunga Ha'apai Eruption on the Antarctic Ozone Hole, Geophysical Research Letters, 51, e2023GL106980, https://doi.org/10.1029/2023GL106980, 2024

Wang, P., Solomon, S., Santer, B.D. *et al.* Fingerprinting the recovery of Antarctic ozone. *Nature* **639**, 646–651 (2025). https://doi.org/10.1038/s41586-025-08640-9

P2, L63: 'Progressing' is vague. I recommend replacing it with a bit more quantitative statement.

Since we provide the quantitative estimates in the next paragraph, we decided to remove this sentence, as it duplicates this information.

P3, L80: In addition to updating the ozone trends, adding some context in terms of changes in the atmospheric circulation since the last publication would be desirable. For instance, did the Hunga-Tonga eruption play any role in ozone trends?

In the revised version, we added a discussion on the influence of circulation on ozone trends.

P3, L84: Adding some examples of the new datasets or citations would be helpful.

Since we describe in Sect.2 the changes in existing datasets and a new (SAGEII-OSIRIS-SAGEIII) satellite dataset, we added here (in Introduction) "(the details are provided in Sect. 2)".

P4, L101: Reference for the OMPS-LP would be useful.

This sentence describes the changes in SAGE-CCI-OMPS+ dataset, therefore we moved the reference (Sofieva et al., 2023), which describes all changes, to the end of the sentence.

P4, Table 1: I wonder why a reference for the SBUV-COH is not listed here.

The references for the SBUV-COH dataset are added.

P5, L117: Is there a reference for this information?

In the revised version, we added the link https://www.usgs.gov/volcanoes/mauna-loa/science/november-27-december-10-2022-eruption-mauna-loa

P6, Table 2: Is there a website where one can download the 'Alpine' ground-based ozone profiles?

The availability of the individual datasets is mentioned in the Data Availability section. Ozone trends are evaluated for each station and instrument type separately and then averaged by instrument type. We added this clarification into the revised version.

P7, L136: data -> outputs

Corrected

P7, L149: Compared to GB22... -> This sentence is vague. Please add more information, such as, how the GHG and ozone scenarios have changed? Is aerosol forcing only included in the refD2?

In the revised version, we added a more detailed description of refD2 simulations.

P8, L180: Was the 1985-2024 period also used for the ground-based observations?

Several ground-based datasets do not cover the full period 1985-2024. Therefore, the trend analysis for ground-based measurements was restricted to the period 2000–2024. This has been clarified in the revised manuscript.

P10, L221: Does this mean 8 models and one ensemble mean?

In the revised version, we clarified: " Final results from CCMI-2022 models (9 models in total) were averaged over the appropriate latitude bands…"

P11, L233-235: What could be the reason for the differences in trends diagnosed from these two datasets? Are these consistent with the previous results?

SAGEII-OSIRIS-SAGEIII is a new dataset. For SAGE-OSIRIS-OMPS, a similar feature was observed also in GB22. In the revised version, we added this note.

P11, L241: There could be some contributions from changes in atmospheric circulation or events that could have impacted ozone in the stratosphere.

Yes, these are included in (i). In the revised version, we mention this explicitly.

P11, L248: “changes in…” – Describe what this means or include a citation here.

We explain here the changes in trends shown in Figure S2. We made rephrasing, in order to avoid confusion.

P12, L276: “vary across the datasets,” – What could be the possible reason for this?

The variations in trends in the UTLS are due to two main reasons : (1) differences in the merged datasets due to different combinations of instruments and merging principles; (2) large variability in the UTLS, which results in large uncertainties in trend estimates. However, the second part of this sentence is ambiguous, and it does not contain important information. In the revised version, we remove it.

P16, L342: Can we explain why the CCMI models can reproduce the observed ozone trends in general?

In the revised version, we added: “…for 2σ uncertainties). Since CCMI models, by design, include the key physical and chemical processes controlling ozone evolution and are driven by observation-based forcings, they reproduce observed ozone trends reasonably well. In this work, the mean CCMI ozone trends …”

P18, L402: A strong…observed. –>A strong…in the NH extratropics (at latitudes…) and below 40 km is observed.

Corrected as suggested.

P20, L449: Please add a sentence explaining why trends are unchanged in the tropics and SH mid-latitudes but became less negative in NH mid-latitudes.

We added: "primarily due to high ozone levels in 2024”.

P21, L464: Please add some information about the importance of the ozone measurements network, comprehensive analyses of the data, and perspective of ozone changes in the future.

In the revised version, we added:

“Maintaining well-calibrated, merged ozone records and conducting periodic assessments of ozone trends are essential for reliably tracking the uneven recovery of the ozone layer and comparing observations with predictions from chemistry–climate models, especially in the face of emerging disruptions such as major volcanic eruptions, extreme wildfires, increasing greenhouse gas concentrations, and increased satellite debris. As key satellite missions approach the end of their lifetimes, the risk of observational gaps is growing, particularly for critical stratospheric species (e.g., water vapor, chlorine species, N_2O) that are needed to interpret changes in ozone. A reduced number of correlative satellite observations will undermine our ability to rapidly identify anomalies and instrumental drifts, leading to increased uncertainties in merged records and derived trends. In this context, the role of the ground-based network becomes even more important, both for satellite validation and for independent assessments of regional ozone variability and long-term change."

Review #2

Dear Reviewer,

Thank you very much for your positive evaluation of our paper. We took your comments into account in the revised version of the manuscript. Please find below our detailed replies (black font) on your comments (blue font).

Reviewer#2 comments:

Major comments:

While I like the fact that the paper is relatively brief, and I especially like the fact that the paper is so well-written, the choice for brevity means the reader has to go elsewhere for important details. A lot of these comments, which together I deem to be "major", represent areas where the paper could be improved by providing information I think is important enough to warrant an explanation.

Line 111: Table 1 gives 10 so-called "merged satellite datasets". However, only 8 of these 10 are used for the bulk of the paper. I did no not recall reading any words about why 2 datasets appear in Table 1, and are then largely not used. Also, in Table 1, the SWOOSH dataset is one of three used for assessing ozone trends as a function of latitude and longitude. For the other two of these datasets, the longitudinal resolution of the data are given in the second column of this table. Therefore, the longitudinal resolution of SWOOSH should be added.

In revised Table1, we added the horizontal resolution of the gridded version of SWOOSH. In the revised version of the paper, we specified that long-term zonally average ozone climate data covering the whole period 1985-2024 are used in the main analyses.

Finally, the last column of Table 1 is very important, but lacking in detail. For instance, for SWOOSH, the table states "Updated version of MLS". For SAGEII-OSIRIS-OMPS, we see "New dataset". Nearly every entry in this last column is lacking vitally important detail. Here, for all datasets, need either the dataset version number, or if this is not possible to provide, some data marker for either when the data were last accessed or even better, when the retrievals were last updated. P.S. I know some of this info is in the text. Still, should repeat in the table in my opinion.

SAGEII-OSIRIS-SAGEIII dataset is created after WMO-2022, therefore we indicate in "changes after WMO-2022" that it is a new dataset.

In the revised version, we added the information about the previous versions of the incoming datasets to the last column of Table 1.

Lines 103 to 104: Sentences "In GOZCARDS … after May 31, 2024)." Need some work. No need to end a sentence with the word "this". I'd end the first sentence with "used prior to WMO 2022" and strike "In fact" from the start of the second sentence.

We rephrased these sentences.

Line 109 to 110: unclear which "ozone profiles" where examined by Arosio et al. (2024); please add more detail here.

In the revised version, we clarified that Arosio et al. (2024) analyzed the OMPS-LP ozone profiles retrieved by the University of Bremen processor.

Line 116": can strike "the" just before "access"

Corrected.

Line 160: So, the regression considers two time periods: before January 1997 and after January 2000. I know of course this is decision of the LOTUS team. Here, or elsewhere in the paper, I think the reader would like to know: a) what happens to ozone measurements obtained from February 1997 to December 1999, the so-called "gap years" ? Are these measurements ignored? Or, are they fit somehow in a manner that allows for better calculation of other regressors, perhaps with a flat line for the ODS term? b) is the end of the regression line for the first period of time forced to match the start of the regression line for the second period of time?

I think based on the nomenclature of equations (1) and (2) that data in the "gap years" are actually used to help estimate the other beta terms. A simple written explanation would be quite helpful.

The "gap" corresponds to the horizontal line, so that the regression lines are connected.

In the revised version, we added the following clarification: "The trend proxy function includes the linear terms in ozone declining (before 1997) and recovery (after 2000) period, and a constant in the "gap" from 1997-2000. Using the gap function avoids the prescription of the turnaround point, which depends on latitude and altitude (Laine et al., 2014)".

Lines 222 to 223: The phrase "uncertainties were calculated with error propagation", without even a citation, are too vague to be of any use to the reader who may one day attempt to reproduce these results. In my opinion, either more detail needs to be added here (my preference), or else a very specific citation such as "as described in Section X.Y of Author et al." should be added.

In the revised version, we explained the evaluation of uncertainties and added references.

Line 226: Connecting to the first major comment above, for Table 1 (line 111), I was surprised to read "the eight merged satellite datasets" because I had remembered reading that Table 1 showed ten datasets.

In the revised version, we added the explanation that these are 8 merged datasets, which cover the period from 1985 to 2024.

Lines 243 to 244: this one sentence paragraph contains important information that I cannot fully understand, since the sentence is so short. Is it the 4 additional years to the proxies (i.e., a temporal extension) that has a small impact? Or, perhaps, is it updates to proxies for ENSO, QBO, etc that has a small impact. More detail here, that is, an actual paragraph, would be helpful. Also, please name the proxies.

This sentence describes the effect of updated data sources for the QBO and F10.7 proxies, and the updated sAOD record. We clarified this in the revised version.

Line 295: Figure 3b is a very important result. Given the anomalous behavior of ozone in year 2024. I suggest either for main or supplement a figure showing how the trend in stratospheric column ozone, as a function of latitude, varies for different choices of the end year. Although this next point is "minor" and easy to address, I'll include here as this point is a natural follow-on.

According to your suggestion, we added a figure to Supplement, which is similar to Figure 3b, but for different choices of end year: 2020, 2023, and 2024.

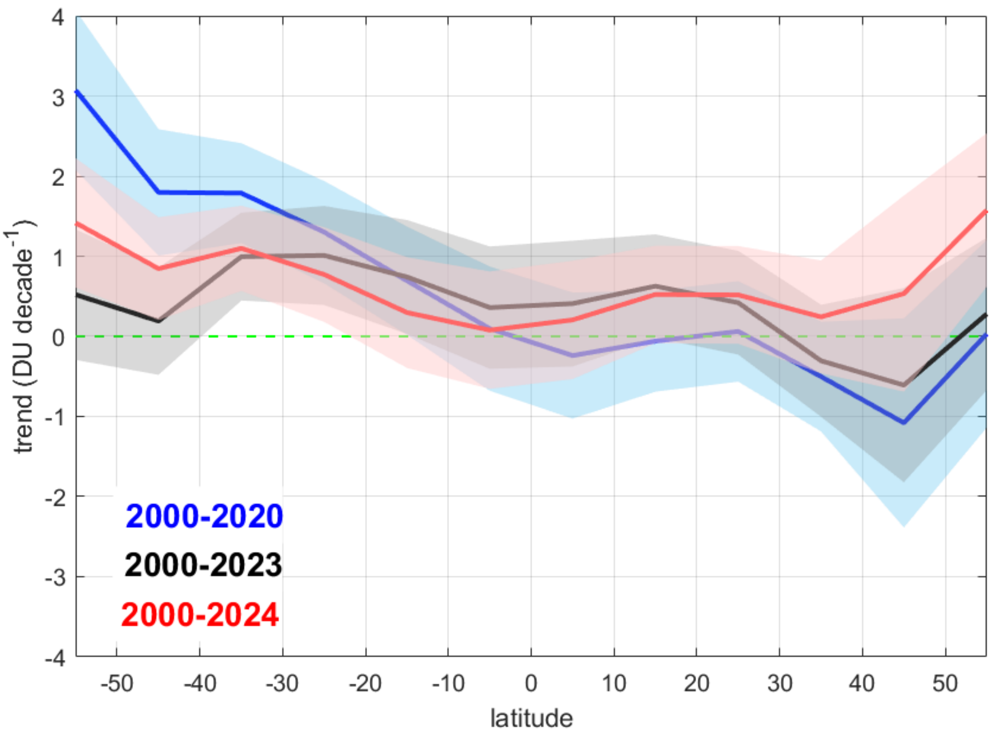

Figure S6. Post-2000 stratospheric column averaged trends (DU decade-1), solid lines, with 2σ uncertainty (shading), for different choices of the end year. Green dashed line highlights zero level. The trends for 2000-2024 period are shown also in Figure 3b.

For lines 304 and 305, where the latitudinal dependence of the ozone trend is compared to WMO-2022 and to Weber et al., 2022, would be great if the time period could be stated.

The information is added.

Lines 333 to 334: Following the same thread as just above, a more detailed explanation about the event that led to “high ozone levels observed in NH in 2024” (this phrasing could be improved) would be very helpful for the reader. Why force an interested reader who might not be familiar with Newman et al., 2024 to fetch this paper, to obtain a senses of the forcing that led to the high ozone?

In the revised version, we added that high ozone levels in NH 2024 were associated with planetary wave events that caused significant stratospheric warmings and forced poleward and downward ozone transport into the lower stratosphere.

Lines 364 to 369: Here, we have a discrepancy in ozone trends over Lauder, NZ found by various observing methods. Again, the text is terse that the reader has to do their own work to find out how often the ozonesondes are launched, whether there has been any attempt to compare ozone abundances reported by the ozonesondes to the other measurements for time periods overlap, etc. More detail here would be quite helpful. For the record, I have thought long and hard about whether to make such requests for more detail. I feel fine making this request, given the large number of co-authors on this paper with enormous expertise, including two colleagues who work at the Lauder station who undoubtedly can take on this task.

In the revised version, we added: “Zeng et al. (2024) reported that homogenization of ozonesonde data at Lauder has a small impact on post-2000 trends. The analyses of drifts between climate data records from different instruments at Lauder by Björklund et al. (2024) have explained roughly half of them by the different sampling, vertical sensitivity, or time periods and gaps. Especially, in the lower stratosphere the drifts between FTIR, sondes and Lidar partial columns (14-22km) were found below 1% decade^{-1}.”

Lines 397 to 399: Are the three merged datasets used here the only ones with longitudinal resolution? If so, this detail would be nice to add to the paper. If not, then some reason for the selection of these three would be helpful. Finally, why 2003-2024 here, when prior sections have focused on 2000-2024? Again, a simple explanation would be quite helpful. It seems this question is answered on line 433; if so, this detail should be appear earlier, and can be repeated here.

Yes, these are all currently available merged high-vertical-resolution datasets with resolved longitudinal structure. MEGRIGOP is from late 2001, SCIA-OMPS is from autumn 2002, pure MLS is from autumn 2004. Similarly to (Sofieva et al., 2021), we selected the years after 2003 for trend analyses, in order to avoid the influence of a major sudden stratospheric warming in September 2002 on ozone trends at Southern Hemisphere middle and high latitudes. We explained this in the revised version.

Minor comments:

Lines 41 and 88: suggest “chemistry-climate models” rather than “climate models”

Corrected

Lines 45 to 46: this sentence does not represent the discrepancy seen over the Lauder station. Suggest a new sentence or two here, that: names the stations (or region of the stations for the 7 used for the "Apline" region), describes the types of measurements, and makes at least passing mention of the one discrepancy that emerges (that is, ozonesondes over Lauder).

Unfortunately, the abstract is limited to 250 words, so it is impossible to add details without cancelling other important information.

The stations and types of measurements are named in Table 2, and the discrepancies are detailed in the dedicated section.

Line 142: it is a shame that the models used an ODS scenario from WMO (2018), since the ODS scenario in WMO (2022) is so different. I suggest adding a sentence or two describing these differences, either here or in Summary section. This paper by Lickley et al. https://acp.copernicus.org/articles/24/13081/2024/ would be a good source of information, for these differences.

In the revised version, we added that ODSs from WMO (2018) were recommended at the time when CCMI-2022 simulations were performed.

Line 145: suggest "stratospheric aerosol forcings"

Changed as suggested

Lines 149 to 150: the sentence "Compared to ... forcing" is quite vague. A sentence or two providing more information on these "changes" would be helpful to the reader. Also, perhaps again "stratospheric aerosol forcing".

The revised version contains a more detailed description of the changes.

Line 164: suggest "fit" rather than "fitted"; terribly minor suggestion and wow, the paper is so well written I even hesitate with this suggestion.

Corrected

Line 202: I think the text is referring to the first term on the right-hand side of Eq. (4). Otherwise, I am confused. If so, please clarify here as well as when the phrase "second term" is used, a few lines down.

Yes, the text refers to the right-hand size of Eq.(4). We corrected the text.

Line 232: Suggest “trends of …” and “upper stratosphere, that are …”

Corrected.

Line 277: I find the phrase “reporting slightly stronger positive trends in the SH middle stratosphere” to be confusing. Stronger in the middle stratosphere compared to the lower stratosphere? Or, stronger within the “several merged datasets” that is the subject of this phrase, compared to other datasets?

We agree that the second part of the sentence is ambiguous. In the revised version, we remove it, because it does not contain important information.

Line 347: if the word “assessment” is maintained, then an assessment should also be cited. Otherwise, can go with “evaluation” since GB22 is not an assessment.

Changed as suggested

Line 349: suggest adding “over 2000 to 2024” somewhere at the start of this paragraph, even though this detail is also in the caption of Figure 6.

Changed as suggested

Line 364: here and other places where appropriate, I would write “-2.5 to -6 % decade^-1”.

ON line 373, we see “reaching +2.5-4 % decade^-1” I would write this as “reaching +2.5 to + 4 % decade^-1”, or perhaps “reaching 2.5 to 4 % decade^-1”. My point is sometimes “-“ is used to mean “to”, and other times “-“ is used as a minus sign,

Lines 370-386: Again, the “2000 to 2024” detail should be interspersed in a place or two

Corrected according to your suggestion.

Line 388: I can not distinguish the color used for Lidar from the color used for satellite mean. Also, “gray shading” should be added to the caption.

We added the legend for gray shading and improved colors.

Line 419: text says “seen especially in SCIAMACHY-OMPS” but to me, this small negative trend appears to be seen only in SCIAMACHY-OMPS. Mighty be a hint of a negative trend in one other dataset.

Small negative trends over Antarctica between ~50° W and ~50°E are seen also in MEGRIDOP and MLS. In the revised version, we changed “especially” to “primarily”.

Review #3

Dear Reviewer,

Thank you very much for your positive evaluation of our paper. We took your comments into account in the revised version of the manuscript. Please find below our detailed replies (black font) on your comments (blue font).

Reviewer#3 comments:

General comments

The manuscript is clearly and carefully written and I have very few minor comments at all, but do have some overall comments to make.

The current work is clearly, and is stated to be, a follow-up to previous LOTUS work, but is also a direct descendant of the preceding work in Harris et al. 2015 and Steinbrecht et al. 2017 (and these papers should be cited), and I would say goes back to Harris et al. 1998.

The authors therefore have apparently made a decision to reduce the scope to that of an update. This is quite understandable but in one or two places I think some things have been missed that needed to have been included (if only briefly). In my view the paper does need to be able to stand alone to some extent. It certainly doesn't need to repeat everything from previous work but the fundamental points should be covered.

In the revised version, we added these references and briefly summarized the results of these studies.

The first example is the lack of discussion of the decision made by the LOTUS team not to include one or more proxies to represent the strength of the Brewer-Dobson circulation in the regression. This is an idiosyncratic choice because such a term is usually considered very important if the overall goal is to assess how ozone has changed due to chemical changes. If the goal is just to see how ozone has changed then you don't need any proxies at all of course. The QBO, ENSO and Solar Cycle terms will approximately even out over a long time period (although I note they are not shown in the current work) in contrast to circulation changes which potentially have a major influence on trends. To be clear I am not requesting that the analysis be re-performed with dynamical proxies but I do think there needs to be some discussion of this fundamental point of why they were not included, as it is closely tied to the purpose of the trend analysis, that is, what exactly are you trying to show?

We added the following clarification of the goal of this study in Sect3.1 of the revised version:

"Various approaches to trend analyses are discussed in detail in P19. The analysis presented in this paper does not include attribution of ozone trends. This study is aimed at examining changes in ozone trends (a combined effect of chemistry and dynamics), derived from observations and chemistry-climate models, between two time periods, 2000–2020 and 2000–2024, and therefore we choose similar proxies and regression model settings to those used in WMO-2022."

However, there are alternative approaches to ozone trend analysis where regression can be performed using EESC (Equivalent Effective Stratospheric Chlorine) and dynamical proxies. Inclusion of these proxies instead of a linear trend can help to attribute ozone changes to the evolution of ozone-depleting substances (ODS) and to changes in the Brewer–Dobson circulation (BDC). However, several studies have revealed some limitations of this approach. A single EESC curve might not be representative for describing ozone changes related to ODS at all latitudes (e.g., Kuttippurath et al., 2015; Weber et al., 2018). Dynamical proxies are usually estimated using data from models/reanalyses, which have uncertainties of their own and might differ across models (e.g., Rao et al., 2019; Šácha et al., 2024), especially in their trends. To minimize uncertainties in trends of dynamical proxies, they can be de-trended and used for characterization of dynamical variability (e.g., Li et al., 2023) and thus improving the fit. Recently, Petropavlovskikh et al. (2025) studied the influence of dynamical proxies on stratospheric ozone trends and found that inclusion of them has very small influence on trends from zonally averaged records in the middle and upper stratosphere but generally reduces trends uncertainties.

The LOTUS report (Sect4.1 of P19) discusses various trend and non-trend proxies, as well as methods for trend analysis (multiple linear regression and dynamical linear modelling (Laine et al., 2014), therefore we simply refer to the P19 discussion in Sect. 3.1 of our paper. However, we mention in our paper the importance of considering dynamical proxies, especially for trend analyses at high latitudes, and propose dedicated analyses in the future (Sect. 4.3). Additionally, we added in Sect.4.1: "Inclusion of de-trended dynamical proxies (e.g., Weber et al., 2022; Li et al., 2023; Petropavlovskikh et al., 2025) might reduce the sensitivity of trend estimates to the selection of the endpoint year. This can be investigated in future works" .

Previously (e.g., the supplement for Godin-Beekmann et al. 2022) the pre-1997 trends were shown in the latitude-height plane which is an important indication of the usefulness of the whole approach.

Since pre-1997 data, the regression model and pre-1997 trends are the same as reported by GB22, they are not shown in this paper. We added this note to the revised version.

Secondly there doesn't seem to be any indication given of how well the regression model performs. This is particularly pertinent with the recent unusual ozone years in both the northern and southern mid-latitudes. I would expect the QBO and ENSO proxies would have very different effects at different latitudes and altitudes but this is not represented anywhere.

One thing very helpful about the Weber et al. 2022 style of presentation for total ozone is that the reader can see at a glance how well the regression model has performed over the whole period and can visually assess the significance of the post 2000 trends in the different latitude ranges. This is a lot simpler to do for total ozone of course, but you could do it for a small number of selected height and latitude ranges. This would be very interesting for the NH lower stratosphere, for instance.

The performance of the LOTUS regression model is discussed in detail in P19 and also in GB22, so we believe it is not needed to repeat the full assessment of contribution of proxies in our paper.

The quality of the regression fit is usually characterized by fit residuals. Since trends uncertainty estimates are derived using fit residuals, comparison of trend uncertainties when using different time periods indicates also changes in the fit quality. The Supplement Figure S7 (Figure S6 in the original version) shows different components of uncertainty estimates. Dashed lines show the propagation of errors from fit residuals. Overall, they are smaller for 2000-2024 than for 2000-2020, as expected (details are presented in the manuscript).

Thirdly, table 1 lists no less than ten different merged satellite datasets (only eight are shown in figures 1 and 2) but I can't find any discussion of what the scientific value really is in having different groups create these merged datasets from different combinations of the same set of instruments. You should give the reader some brief indication of why this is worth doing.

In the revised version, we added:

"While some merged datasets include similar sets of satellite observations (for example, SBUV MOD and COH are created from the SBUV data, SWOOSH and GOZCARDS primarily use SAGE II and MLS data, SAGE II data are used in six merged datasets), they still differ in the methods used to merge the original measurements as well as the exact selection of original measurements (time period, version, etc.). Comparing trend results across different merged datasets enhances confidence in the derived trend estimates."

In lines 233-235 it is stated that two of these datasets give noticeably larger trends in the northern hemisphere upper stratosphere than the others, and in lines 275-277 you also say there are some differences in the southern hemisphere middle stratosphere. It would be important to investigate these further but I can understand that you might consider it to be out of scope for the present work.

We fully agree, these features should be investigated further. However, these studies are out of scope of the present work.

The discussion of the longitudinally resolved trends (in particular over the Antarctic) and the comparisons of ground-based instruments are also treated fairly cursorily, and I hope will be further investigated in more detail in other work. There is a statement (line 363) that discrepancies between the instruments at Lauder is greater than at the other two ground-based locations, but there has been some recent work specifically done on this topic motivated by SG 2022 which you don't consider or cite (Björklund et al. 2024 , Zhang et al. 2024).

Thank you very much for your suggestions and references. We include them and their conclusions in the revised version of the manuscript (see also below).

I think the set-up of the trends with the flat "gap" period in the regression model is very appropriate, and should lead to realistic fittings of the proxy terms (that is, it will give better results than trying to fit ozone to a V-shape).

Figures 3b and 3c are excellent, for many readers these will really be the key findings, especially that the southern hemisphere stratospheric ozone column shows a positive trend all the way from 20° S to 60° S.

Including the previous results in Figure 5 is very helpful for those interested in how the new results differ from the old ones, such as for the Ozone Assessment, along with the discussion in lines 239 to 260 and the supplement.

I was also very pleased to see Figure 7 with the representation of latitude dependent changes.

Thank you very much.

Specific comments

Line 41 "good agreement". I am not sure this is a justifiable statement but perhaps it is. There are some important differences between the model trends and the observations and the uncertainty ranges for both observations and models are still frustratingly large.

We removed "good".

Line 84 "new datasets have become available" – explain what this means (new instruments, newly reprocessed versions of the data, new combinations of data, etc)?

Here, in introduction, we added "(the details are provided in Sect.2)".

Lines 1-89 Somewhere in here you should mention the previous work on vertically resolved trends (at a minimum: Harris et al. 1998, Harris et al. 2015 , Steinbrecht et al. 2017).

In the revised version, we added these references and briefly summarized the results of these studies

Table 1 – I am not sure saying "present" is the best way to express this. For one thing, this paper will be published in 2026 but the datasets are only analyzed up to 2024. Secondly, people reading the paper in future years would probably prefer to know the actual year that was included.

We changed "present " to "2024".

Table 2 Isn't there also an ozone Lidar at Lauder?

The Lauder lidar is mentioned in Table 2.

Lines 115-132 There is no discussion of ozonesondes here even though we know the homogenization is a major issue and a lot of work has been done. (I suspect this might have been an accident).

This was accidentally missing. We added: " Ozonesonde data have been homogenized, so that their stability is expected to be improved (Ancellet et al., 2022; Björklund et al., 2024; Van Malderen et al., 2025)"

Line 135 Should "CCMI" be "CCMI-2022"? I am a little bit confused on this point because I wasn't clear if you were using the newer model runs.

Corrected.

Line 139 "newly developed" - is this old text copied from an old document?

We replaced "newly developed" with "recent".

Line 148 "a reference model dataset for comparison with observations" – this is what you're doing in the current work, isn't it? You should say that. Has there been other work to assess the models against observations?

Yes, we added "presented here", and slightly rephrased this sentence.

Lines 149-150 This is an important point, to list what has changed since GB22, however I think you need a few more words to flesh out each of the four items listed.

In the revised version, we added a more detailed description of refD2 simulations.

Line 164 – were the other datasets deseasonalized with respect to 1998-2008 as well?

The deseasonalization in the datasets, which are created based on deseasonalized anomalies, is usually more complicated. First, the anomalies are computed using a reference time period (corresponding to a reference instrument or majority of the instruments), and then anomalies from earlier (like SAGE II) or later (like SAGE III/ISS) operating instruments are offset to this reference anomaly. Therefore, we used deseasonalized anomalies provided in these datasets.

Equation 1 – I don't understand why the beta terms are said to be a function of time? Wouldn't they be a function of height and latitude but not time?

The LOTUS regression model allows seasonal dependence for all coefficients. However, in this paper the seasonal variability is fitted only for three first coefficients, therefore we remove "t" from other β terms

Lines 222-223 "the uncertainties calculated with error propagation" – please clarify this statement

In the revised version, we explained the evaluation of uncertainties and added references.

Line 341 "CCMI" – again this should be "CCMI 2022", shouldn't it be?

Corrected.

Lines 361-362 Unfortunately, for me this statement "opens a can of worms". If the sampling is different, and that difference affects the trends, then doesn't that mean this comparison shouldn't be done at all unless the sampling effects are investigated and taken into account in some way? (Eg by using satellite data only on the day of each ozonesonde flight).

The climate data records from ground-based instruments are affected by so-called representativeness issues. For detection of drifts, analyses should be on collocated measurements, for both ozonesonde and satellite. And even in this case local profiles from ozonesondes measured during ~1-2h will not be fully compatible with nearly instantaneous but spatially (~300-400 km) averaged profiles from limb instruments.

Here we compare climate data records and the derived trends using data from different measurement systems, with their inherent features. From our point of view, the information on how the derived trends agree or differ is important.

In the revised version, we added the references on the papers, which discuss the problem of representativeness of ground-based climate data records.

Line 363 As mentioned above, you should at least mention Björklund et al. 2024 and Zhang et al. 2024 here.

In the revised version, we added: “Zeng et al. (2024) reported that homogenization of ozonesonde data at Lauder has a small impact on post-2000 trends. The analyses of drifts between climate data records from different instruments at Lauder by Björklund et al. (2024) have explained roughly half of them by the different sampling, vertical sensitivity, or time periods and gaps.”

Lines 376-377 Isn't the agreement better at the European sites because multiple records have been combined (2 or 3 of everything) over dispersed locations and so smoothed out rather than one each at single points?

In the revised version, we added this as a possible reason.

Line 419 In the figure I can only see a negative trend in SCIAMACHY-OMPS, not "especially".

Small negative trends over Antarctica between ~50° W and ~50°E are seen also in MEGRIDOP and MLS. In the revised version, we changed “especially” to “primarily”.

Minor comments

Line 88 "... the climate models" replace with "coupled climate chemistry models"

Line 104 Improve the wording of "(But Aura MLS v4 ...) "

Line 110 "is done" – replace with "was found by" or similar words.

Line 157 Insert "and" before "stratospheric"

Line 188 Delete "the" before "zenith sky".

Line 232 Replace "they are" with "which are"

Line 236 Insert "the" before "majority"

Line 237 Insert "the" before "uncertainties"

Line 399 Insert "the" or "these" before "aforementioned"

Line 401 Insert "the" before "shorter"

Line 402 Insert "of" after "poleward"

Line 404 Insert "the" before "2003-2018".

All are corrected.

References

Ancellet, G., Godin-Beekmann, S., Smit, H. G. J., Stauffer, R. M., Van Malderen, R., Bodichon, R., and Pazmiño, A.: Homogenization of the Observatoire de Haute Provence electrochemical concentration cell (ECC) ozonesonde data record: comparison with lidar and satellite observations, Atmospheric Meas. Tech., 15, 3105–3120, https://doi.org/10.5194/amt-15-3105-2022, 2022.

Björklund, R., Vigouroux, C., Effertz, P., García, O. E., Geddes, A., Hannigan, J., Miyagawa, K., Kotkamp, M., Langerock, B., Nedoluha, G., Ortega, I., Petropavlovskikh, I., Poyraz, D., Querel, R., Robinson, J., Shiona, H., Smale, D., Smale, P., Van Malderen, R., and De Mazière, M.: Intercomparison of long-term ground-based measurements of total, tropospheric, and stratospheric ozone at Lauder, New Zealand, Atmospheric Meas. Tech., 17, 6819–6849, https://doi.org/10.5194/amt-17-6819-2024, 2024.

Kuttippurath, J., Bodeker, G. E., Roscoe, H. K., and Nair, P. J.: A cautionary note on the use of EESC-based regression analysis for ozone trend studies, Geophys. Res. Lett., 42, 162–168, https://doi.org/10.1002/2014GL062142, 2015.

Laine, M., Latva-Pukkila, N., and Kyrölä, E.: Analysing time-varying trends in stratospheric ozone time series using the state space approach, Atmos. Chem. Phys., 14, 9707–9725, https://doi.org/10.5194/acp-14-9707-2014, 2014.

Li, Y., Dhomse, S. S., Chipperfield, M. P., Feng, W., Bian, J., Xia, Y., and Guo, D.: Quantifying stratospheric ozone trends over 1984–2020: a comparison of ordinary and regularized multivariate regression models, Atmospheric Chem. Phys., 23, 13029–13047, https://doi.org/10.5194/acp-23-13029-2023, 2023.

Rao, J., Yu, Y., Guo, D., Shi, C., Chen, D., & Hu, D. (2019). Evaluating the Brewer–Dobson circulation and its responses to ENSO, QBO, and the solar cycle in different reanalyses. *Earth Planet. Phys.*, 3, 166–181. https://doi.org/10.26464/epp2019012

Šácha, P., Zajíček, R., Kuchař, A., Eichinger, R., Pišoft, P., and Rieder, H. E.: Disentangling the Advective Brewer-Dobson Circulation Change, Geophys. Res. Lett., 51, e2023GL105919, https://doi.org/10.1029/2023GL105919, 2024.

Van Malderen, R., Thompson, A. M., Kollonige, D. E., Stauffer, R. M., Smit, H. G. J., Maillard Barras, E., Vigouroux, C., Petropavlovskikh, I., Leblanc, T., Thouret, V., Wolff, P., Effertz, P., Tarasick, D. W., Poyraz, D., Ancellet, G., De Backer, M.-R., Evan, S., Flood, V., Frey, M. M., Hannigan, J. W., Hernandez, J. L., Iarlori, M., Johnson, B. J., Jones, N., Kivi, R., Mahieu, E., McConville, G., Müller, K., Nagahama, T., Notholt, J., Piters, A., Prats, N., Querel, R., Smale, D., Steinbrecht, W., Strong, K., and Sussmann, R.: Global ground-based tropospheric ozone measurements: reference data and individual site trends (2000–2022) from the TOAR-II/HEGIFTOM project, Atmospheric Chem. Phys., 25, 7187–7225, https://doi.org/10.5194/acp-25-7187-2025, 2025.

Petropavlovskikh, I., Wild, J. D., Abromitis, K., Effertz, P., Miyagawa, K., Flynn, L. E., Maillard Barras, E., Damadeo, R., McConville, G., Johnson, B., Cullis, P., Godin-Beekmann, S., Ancellet, G., Querel, R., Van

Malderen, R., and Zawada, D.: Ozone trends in homogenized Umkehr, ozonesonde, and COH overpass records, Atmospheric Chem. Phys., 25, 2895–2936, https://doi.org/10.5194/acp-25-2895-2025, 2025.

Weber, M., Coldewey-Egbers, M., Fioletov, V. E., Frith, S. M., Wild, J. D., Burrows, J. P., Long, C. S., and Loyola, D.: Total ozone trends from 1979 to 2016 derived from five merged observational datasets – the emergence into ozone recovery, Atmospheric Chem. Phys., 18, 2097–2117, https://doi.org/10.5194/acp-18-2097-2018, 2018.

Weber, M., Arosio, C., Coldewey-Egbers, M., Fioletov, V. E., Frith, S. M., Wild, J. D., Tourpali, K., Burrows, J. P., and Loyola, D.: Global total ozone recovery trends attributed to ozone-depleting substance (ODS) changes derived from five merged ozone datasets, Atmospheric Chem. Phys., 22, 6843–6859, https://doi.org/10.5194/acp-22-6843-2022, 2022.

Harris, N. R. P., Hudson, R., and Phillips, C. (Eds.): Assessment of trends in the vertical distribution of ozone, Stratospheric Processes and Their Role in Climate/International Ozone Commission/Global Atmospheric Watch (SPARC/IO3C/ GAW) Report 1, World Meteorol. Organ. Ozone Res. and Monit. Proj. Rep. 43, 289 pp., Geneva, 1998.

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.

Steinbrecht, W., Froidevaux, L., Fuller, R., Wang, R., Anderson, J., Roth, C., Bourassa, A., Degenstein, D., Damadeo, R., Zawodny, J., Frith, S., McPeters, R., Bhartia, P., Wild, J., Long, C., Davis, S., Rosenlof, K., Sofieva, V., Walker, K., Rahpoe, N., Rozanov, A., Weber, M., Laeng, A., von Clarmann, T., Stiller, G., Kramarova, N., Godin-Beekmann, S., Leblanc, T., Querel, R., Swart, D., Boyd, I., Hocke, K., Kämpfer, N., Maillard Barras, E., Moreira, L., Nedoluha, G., Vigouroux, C., Blumenstock, T., Schneider, M., García, O., Jones, N., Mahieu, E., Smale, D., Kotkamp, M., Robinson, J., Petropavlovskikh, I., Harris, N., Hassler, B., Hubert, D., and Tummon, F.: An update on ozone profile trends for the period 2000 to 2016, Atmos. Chem. Phys., 17, 10675–10690, https://doi.org/10.5194/acp-17-10675-2017, 2017.

Zeng, G., Querel, R., Shiona, H., Poyraz, D., Van Malderen, R., Geddes, A., Smale, P., Smale, D., Robinson, J., and Morgenstern, O.: Analysis of a newly homogenised ozonesonde dataset from Lauder, New Zealand, Atmos. Chem. Phys., 24, 6413–6432, https://doi.org/10.5194/acp-24-6413-2024, 2024.