1. **We thank Referee 3 for the careful and thorough review of our manuscript and for the valuable comments and suggestions. In the following, we provide a point-by-point response to each comment raised by the reviewer.**

1. In general the work is very carefully carried out and well explained.

3. The use of dynamic proxies in the MLR model can be argued for or against but in my opinion is a good choice if the aim is to identify a trend due to EESC. Results such as those shown in figure 7 and figure 12 illustrate the significance of this choice.

4. I am slightly surprised you don't highlight more that southern hemisphere mid latitude total ozone has a significantly positive trend from about 15-40 degrees even without dynamic proxies, this seems like a very important finding to me.

This is now mentioned in the manuscript text. See comment 18 for the specific changes.

5. I have only one serious concern, which is the use of the three different SAOD proxies. I must say that I am not at all convinced that this is valid.

6. The purpose of this proxy should be to quantitatively account for the *chemical* effect of extra aerosol surface area in the stratosphere. The dynamical effects should be captured by your dynamical proxies. GLOSSAC is latitudinally resolved so there should be no issue there. Figure 6 panels (i) (j) and (k) seem completely implausible to me. If the SAOD proxies are intended to capture both chemical and dynamical consequences of particular eruptions in a totally unspecified way then there is a large danger of introducing "fudge factors" that don't have a clear physical meaning. The approach taken here seems to me to somewhat defeat the purpose of MLR.

7. Therefore I would suggest either clearing up the roles of these proxies in much greater detail (by which I mean the physical explanation and not just the statistical effect on trends), or keeping just one SAOD proxy for all periods based on GLOASSAC, or deleting periods from the record that MLR doesn't work properly for.

We clearly indicate that the SAOD proxy does not stand alone for the chemical effect due to sulphate aerosol, but also covers the dynamical and radiative response from the enhanced aerosols (Line 273). For a given aerosol loading, the dynamical response can be highly variable and this is the main reason to use different SAOD proxies. Another reason would be that the aerosol surface area density (SAD) is responsible for the ozone depletion rather than the extinction coefficient and the relationship of SAD to SAOD might be different for different eruptions/conditions. We added some additional text to the paper describing, why we use three separate proxies for our regression.
Line 269 (track changes version):
"**Volcanic eruptions have different impacts on stratospheric circulation, chemistry, and ozone distribution, depending on magnitude, season, and latitude of the event (Stolarski et al., 2006; Chehade et al., 2014; Weber et al., 2018, 2022). This variable response is illustrated in panels i-k of Fig. 6. The Australian wildfire (end of 2019) and HTHH had a stronger initial impact in the SH than in the NH because the volcanic aerosols (and stratospheric water vapour) were**

transported mostly into the SH resulting in an ozone loss (band of negative values in panel i). The Pinatubo eruption had a stronger impact in the NH (boreal winter) as illustrated by the band of negative coefficients from the tropics to the polar region (panel j, Aquila et al. (2014)). The dynamic response to the El Chichón eruption showed the opposite effect. Due to changes in the SH circulation the dynamical ozone increase exceeded the chemical ozone loss caused by the eruption (see panel k, Zerefos et al. (1994); Angell (1997)). However, nearly all El Chichon coefficients are not statistically significant at 95% confidence level.."

8. It is very important to note that the APARC Hunga report found only quite modest chemical effect on ozone.

9. *Østerstrøm, F. F., M. L. Santee et al. (2025): Effects of the Hunga eruption on stratospheric ozone and related trace gases. In APARC, 2025: The Hunga Eruption Atmospheric Impacts Report [Yunqian Zhu, Graham Mann, Paul A. Newman, William Randel (Eds.)]. APARC Report No. 11, WCRP Report No. 10/2025, DOI: 10.34734/FZJ2025-05242, available at https://aparc-climate.org/publications/aparc-report-no-1*

10. The finding that tropospheric trends are not making a significant contribution to total ozone trends is also very important.

The chapter of Oesterstrom and Santee (2025) is already listed. In the discussion section the following is stated in the manuscript (Line 482 in the track change PDF): "The HTHH-SAOD proxy accounts to some extent also for the effect from the enhanced water vapor following the underwater eruption of HTHH (Vömel et al., 2022; Millán et al., 2022). Changes in transport/circulation had a larger influence on ozone than altered chemistry due to enhanced chlorine activation following the eruption (Oesterstroem and Santee, 2025)."

2. Line 21 It's not quite right to say "ie, the ozone hole", because you were talking about both hemispheres.

   We slightly changed the sentences to remove this incorrectness.

3. Figure 1 – The caption should state what the colors are marking.

   The colors indicate the viewing geometry of observation of each instrument. This was added to the caption of the figure.

4. Figure 1 – As a comment, the different datasets are not very independent with heavy reliance on SAGE II in the first part of the record. This would influence the regression results because the fit is over the whole period.

   This is really important, the datasets are not independent in the first period at all. This reflects the limited availability of high-quality global ozone profile measurements

during that period. However, despite using the same instrument data, the merged datasets differ in their retrieval versions, bias corrections, merging methodologies, sampling, and quality control procedures.

5. Lines 103-105 The satellite total ozone products have been extensively validated against each other and against ground based data so they are not very independent.

   This is true and we removed "independent" from the sentence.

6. Figure 2 - I really like this figure which sets things out very clearly, but the caption should state what the colors are marking.

   The caption was extended, indicating that the colors highlight different instrument families.

7. Figure 2 – There is a minor inconsistency in the labels of OMPS-NP, where NOAA SBUV specifies Suomi-NPP and NOAA-20 but SBUV NASA doesn't.

   Fixed

8. Line 108 "we need all ozone profiles in units" – please re-word slightly – it's more of a choice you've made.

   We changed the sentence as follows:
   In this study, all ozone profile datasets are expressed in units of number densities [molec cm^{-3}] as a function of geometrical altitude [km].

9. Lines 121- 140 It's great that you looked into this effect, which shows care and attention to detail.

   Thank you!

10. Lines 139-140 You should include a short comment about what would be the general cause of why this might happen, is it departures from the zonal mean?

    We changed the text to better address the comment, as follows:
    "**The significance of these trends decreases with altitude, showing the most significant trends at 25 km. This is possibly because of stronger vertical ozone or temperature gradients near this altitude, which increase the uncertainties from the non-linear ozone unit conversion and vertical axis transformation."**

11. Lines 158-159 Again, I think a short comment is needed here about what would cause such effects. Does the fact the blue line in Figure 5 is slightly higher than the red line suggest that the more satellite instruments are included, the differences become greater?

We agree that an increasing number of contributing satellite instruments could potentially lead to larger differences between merged datasets due to differences in long-term stability. However, the comparison shown in Fig. 5 does not indicate a clear dependence on the number of instruments alone, because the 1996 to 2024 line covers a longer time period than the 2000 to 2024, which includes more instruments (Fig. 2), but the 1996 to 2024 line shows lower drift uncertainties.

12. Line 165 You should include here a reference to Table 2 so the reader can check what each of the abbreviated proxy names refer to.

We added a references for Table 2.

13. Figure 6 comments:

1. I very much like this figure which I wish appeared in all papers which use MLR, in order to show that the regression results appear physically plausible and prevent the process turning into a 'black box'.
2. You need to say in the caption that the proxies have been normalized to 0-1 etc

   We added the information about the normalization to the caption:
   All proxies were normalized to -0.5 to 0.5 apart from the SAOD proxies (i), (j) and (k), which were normalized to 0 to 1.

3. What is the horizontal bar next to each proxy name representing?

   This is where SWOOSH does not have enough data in the high altitudes and the MLR does not work well anymore.

4. The solar response looks weak given how prominent the solar cycle is in long-term global ozone? (eg Figure 1 of Weber et al. 2022)

   We compared our coefficients with Weber et al. 2022 Fig. 5. We see that in Weber et al. the QBO response is stronger than the solar response by a factor of two The same relationship we see in our results. We see the maximum response from our solar proxy at approximately same latitudes as in Weber et al. 2022.

5. What is the grid cell in the tropical UTLS that shows up in panels a, b, c, g, h and j? Is it real?

Those grid cells are at around 9 km in the tropics and with that below the tropopause. Because of this, we don't think they are real, but rather result from fewer observations as well as cells in which the MLR does not work well anymore. We added the tropopause in this figure to highlight this.

6. A comparison of panels i, j and k show very little resemblance. Why is the effect of the aerosol so different?

The influence of SAOD from volcanic activity heavily depends on the strength, latitude and time of the year of the eruption. For this reason, we see very different results for the three different aerosol proxies and decided to split the SAOD into three separate SAODs. See also comments earlier.

14. Line 278 But wouldn't the long-lasting perturbations be caused by the aerosol remaining at elevated levels for longer, which would be captured by GLOSSAC? If this is not so then do you need to add a lag term?

It's not clear to which part of the manuscript this question has been raised.
A time lag was not used in recent studies Sofieva et al. (2026), Godin-Beekman et al. (2022), Weber et al. (2022). We use a latitudinal dependent SAOD, which accounts for the delay in the aerosol loading with latitudes due to the meridional circulation.
looking at our time series we do not see any indications for a presence of a time lag, but it might be something that needs a more in-depth analysis (beyond the scope here).

15. Line 270 What do you mean by "absolute" here?

The sentence was changed so it only says "ozone timeseries", as a reminder that we use absolute units and not anomalies in our analysis.

16. Line 283 Is the fact that the ozone datasets agree well shown anywhere?

This is not shown in the paper. Weber et al. 2025 showed that there is good agreement between the satellite data after bias correction. It is only shown exemplary for the stratospheric columns (after bias correction) in Fig. 11 of our paper.

17. Lines 285-289 I find this discussion a bit confusing because the orange and blue lines agree very well in the SH from about 15 degrees to 35 degrees, and only then start to diverge, but in the NH they are different at 30 degrees, which is the opposite of the text. (It sounds like the text is referring only to the higher latitudes).

We added "40 poleward°" to the text to indicate which part of the southern

hemisphere we are talking about. In the northern hemisphere at 30° we indeed see a larger difference between both methods, but this is still well within the uncertainties.

18. Lines 283-289 Also in this discussion, I would have thought that the fact the SH trends are significantly positive at many latitudes would be worth mentioning?

We now mention the significant trends in the southern hemisphere in comparison to the non-significant trends for the second method:
“The trends in the SH 40° (poleward) increase on average to +3 to +4 DU per decade (+1.3 to +1.5% per decade) when using the full proxy setup, while also being significant. With the LOTUS settings, SH trends are lower remaining at about +1.5 DU per decade and insignificant above 45°S (see also Fig. A2 for the trends in % per decade).”

19. Line 291 "The more positive trend … in their and our … " – more positive than what?

We changed the text for more clarity:
“The full regression using additional dynamical proxies and three SAOD proxies shows above 40°S a stronger positive trend than the LOTUS regression (without dynamical proxies and a single SAOD proxy)”

20. Lines 308-312 I'm not clear what point you're making here. It is true that the LOTUS proxies were not designed for the poles, but as you say, they make a big difference in high mid-latitudes.

This part describes the trend model of the LOTUS regression, which does not work well in the polar regions compared to the Full Regression Model. We show that with the additional dynamical and the three SAOD proxies, the model works very well at high and mid-latitudes as well as in the polar regions. To improve on the statement, we changed the manuscript as follows:
“It is imperative to note, that the LOTUS setup (panel b) was developed for extra-polar latitudes and thus is not suited for polar regions, which is why it only serves as an illustration here.”

21. Line 314 Some minor re-wording is need here, because the "early phase" sounds like it might mean the early history of the ozone hole in the 1980s.

We added some text to refine the sentence:
“Polar ozone recovery was first identified in the early phase of the ozone hole season (September)”

22. Line 321 The influence of what, sorry?

We added: “influence of the dynamical proxies” to the sentence.

23. Lines 323-325 The effect on the ozone hole was very modest according to the APARC report.

This is mentioned in the discussion section

24. Lines 328-330 I can't see that you have demonstrated this, perhaps you could make this more explicit.

This is shown in Fig. 8 by the huge changes in ozone trends, when splitting the SAOD into three separate proxies. Our reply to comment 1. explains it in more detail.

25. Figure 8 The arrangement of the panels isn't very user-friendly. Panels a-b-c-d are natural, but not d-e-f which would be easier to compare if they were in the same row or column.

That is correct. We tried various setups for this figure and even split up the figure into two, but this way it was not clear that panel d) was the same in the next figure. This is why we decided to include the SAOD split into Fig. 8.

26. Figure 8 Comparing panels a and c, by eye, the extra proxies seem to improve the fit in some years (2012, 2021) but don't help much in others (1982, 2008, 2018, 2019). Can you comment on that?

We don’t expect the dynamical and split AOD proxies to catch every minima and maxima of the time series, but rather the most prominent ones. The years 2018 and 2019 seem to have improved but not perfectly.

27. Figure 8 There is a huge difference in the quality of the fit between (b) and (d), but not so much between (d), (e) and (f). Why does the use of the third proxy make the recovery trend so much greater? Are you sure this is not an artefact?

The important distinction here is that the third AOD proxy is the split between the Pinatubo and the HTHH period. The HTHH period shows much lower values in comparison to the Pinatubo period, splitting these periods and then normalizing them improves the fit for the HTHH proxy for the 2000 to 2024 time period by not being overshadowed by the Pinatubo period. The larger post-ODS peak trends seen in panel d (+9%/decade) is in close agreement with the result from Weber et al. (2022),

12%/decade, indicating that the years 2020-2022 are an episodic excursion from the long-term ozone recovery due to the ODS decline.

28. Lines 359-363 Why would this be?

We have not examined the smaller SH positive trends of GOZCARDS in detail, but we did reproduce the results of Sofieva et al. 2026, using deseasonalized ozone data. We think this difference comes from the seasonality that was either removed by the deseasonalization or not caught by the MLR.

29. Lines 370-371 Wouldn't the use of the climatological value mean your trends will be affected by a rising tropopause height over this period, and/or tropical expansion into higher latitudes? Do you allow for different seasons?

Since we use quite broad latitudes with 10° steps and the changes of the tropopause are rather small, we don't think that the use of the climatology will influence the calculation of the stratospheric column strongly. But we want to look into this in more details when using datasets with a better resolution (beyond of scope here).

30. Line 376 What is "bias corrected" here?

We added a short description of the bias-correction after the sentence:
"**The stratospheric ozone column was bias-corrected by subtracting the mean during the 1998–2008 period from each dataset (anomaly) and adding back the multi-instrument mean to obtain bias-corrected absolute ozone values for each dataset."**

31. Figure 12 – It is very encouraging to see southern hemisphere stratospheric ozone is recovering so strongly in the IUP regression, although equally concerning that this is not true in the NH. Can you comment on this?

This is most probably due to the asymmetrical ozone trends in the two hemispheres already evident in Fig. 9. Previous studies have shown that hemispheric differences in stratospheric circulation lead to an asymmetry in stratospheric HCl trends, with a faster decline in the Southern Hemisphere (Chrysanthou et al., 2025; Dubé et al., 2025). The associated decline in inorganic chlorine reduces ozone loss and therefore leads to a stronger ozone recovery in the SH. This aspect is already discussed in the manuscript.