Reply to Anonymous Referee #2 review of manuscript acp-2023-788

Trends in polar ozone loss since 1989: First signs of recovery in Arctic ozone column

Andrea Pazmino on behalf of all co-authors

We thank Anonymous Referee #2 for the time devoted to evaluate our work. We are particularly grateful for the warning on the wording using to explain the results related to metric 3 that we have taken particular care of. Your comments have helped us to improve our manuscript. Please find our answers below (in red)

***** More specifics regarding the more major issues/comments**

- L34 and related discussion in the manuscript: I would argue that a 1 sigma "detection" is not really a detection with enough significance; it is a likely detection as opposed to a robust detection (at 2 sigma or more), and some scientists in various disciplines would argue for even stronger significance levels, in addition to the fact that one often cannot or does not include all possible error bars in the analyses. In this case, the assumption of a linear relation between ozone loss amounts and VPSC is just that, an assumption taken as "truth" and any departures from this "truth" signify something related to ozone trends or recovery. In my view (and hopefully in the views of others in this field), this is just an indirect method at suggesting there may be "signs of recovery" (per the title of your manuscript), which is a better language than a "first quantitative detection". I understand that there is often a desire to show a "first" in research, but this can be overdone, and science progress is usually obtained via multiple analyses over time, especially for inferring trends, and the Arctic can change enough that adding or subtracting years can make substantial differences in the results. Rather than using a bold assessment like "first quantitative detection", I would urge the authors to use a more cautious statement. Error bars here are a lower limit, especially since the same sort of analysis for the Antarctic region yields error bars that are significantly smaller than other metrics results, so this is somewhat suspicious to me just on this basis, in addition to the fact that this method is more indirect than the other two metrics. Please counter this argument if you feel that there is a strong reason to declare victory on the Arctic recovery signal based on just one indirect metric and at the 1 sigma level (at best). I am not convinced, at this stage, and I feel that more metrics and years are needed for such a bold statement (including a broader community assessment, such as another WMO report, for example). I would not try to argue the validity of line 33 too strongly, as long as line 34 gets deleted, or replaced by something like "We argue that this points to the first signs of ozone recovery in the Arctic springtime lower stratosphere." [Although I personally would probably say this "may point", being a cautious person on such matters.] Alternatively, please make the case regarding such a bold statement by performing more error analyses - but this will typically increase the error bars, so the case will just become even weaker, I predict. Moreover, given that the ozone recovery path depends on both ODS and greenhouse gas effects, it is also difficult to provide a robust attribution of slightly positive trends to one effect or the other, without more detailed analyses; there are not enough model results to compare to, in terms of what a model would predict for one effect versus a combination of effects, in general, with comparisons to any of the observations shown in this work. I am therefore going to remain skeptical of broad sweeping conclusions for the Arctic, especially (although some caution is also recommended for Antarctic ozone studies). In fact, your own words at the end of the manuscript show more restraint and caution (with a pointer to another reference as well), so I imagine you actually agree with my words of caution. I think this shows nice results, whether one wishes to claim a "first" or not, and this is what should be the more important conclusion, a good set of analyses with hopefully reasonable error bars, and without overstating the possible conclusions.

As you have mentioned, the metric 3 uses an indirect but well-known assumption of the relationship between ozone loss and Polar Stratospheric Clouds due to the essential role of the latter in the heterogeneous chemical processes involved in the ozone depletion. We understand your concern about this indirect method and we took your warnings on wording into account. For example, the title has been changes as follows:

Trends in polar ozone loss since 1989: Potential sign of recovery in Arctic ozone column

To perform this study more robustly, we have decided to apply a multi-parameter regression model on the ozone loss dataset since 2000 using the Sunlit VPSC and a linear trend as proxies as shown in Equation 1:

$$MOLoss(t) = SunlitVPSC\_contr(t) + t1 * (year(t) - 2000) + \in(t) \quad (1)$$

where t is year since 2000, SunlitVPSC_contr(t) is the term corresponding to the linear (HN) or parabolic (HS) contribution of Sunlit VPSC, the variable t1 is the regression coefficient of the time proxy and $\in$ (t) is the fit residuals. The contribution of Sunlit VPSC is represented by Eq. 2 for the NH and Eq. 3 for the SH:

$$SunlitVPSC\_contrNH(t) = K_{0\_NH} + K_{1\_NH} * SunlitVPSC\_NH(t) \quad (2)$$

$$SunlitVPSC\_contrSH(t) = K_{0\_SH} + K_{1\_SH} * SunlitVPSC\_SH(t) + K_{2\_SH} * SunlitVPSC\_SH(t)^2 \quad (3)$$

where *K* are the regression coefficients of the respective proxies mentioned above.

All the regressions coefficients are significant at more than two standard deviations. The linear negative trends of ozone loss are significant and similar for both hemispheres, presenting slightly larger values than the previous results: 2.00±0.97 % dec^{-1} for the NH and 2.21±0.67 % dec^{-1} for the SH.

Those results confirm our previous ones but, as you mentioned, it is necessary to consider with caution the significance of the trends considering the error bars. The lines 33-34 were therefore changed as follows:

"Metric 3 provides a negative trend in Arctic ozone loss residuals with respect to the VPSC fit of **-2.00 ±0.97 (1σ) % dec^{-1}**, with **limited** significance at **2σ** level. **With such metric a potential** quantitative detection of ozone recovery in the Arctic springtime lower stratosphere **can be made**."

The Section 5.3 was changed introducing this new method and taking also into account the concerns of Referee 1. Please see the new Section 5.3 in the answer to Referee 1.

- As a related comment regarding Fig. 11, if one wants to claim enough robustness in the result and error bars, one should try to use 2 or 3 years less (or more) at the beginning or end of the series, to see how this affects the results and error bars. I think it is best, again, to be cautious in terms of a "robust detection" comment, unless the analysis is at least significant at the 2 sigma level, with enough sensitivity analyses as well.

The sub-section 5.3 was changed as explained above using a more robust method by applying a multi-parameter regression model. The low residuals values and the good correlation between observations and model (R~0.92) gives us confidence in the results.

- In addition, why not show what a polar-focused model would predict for such a metric, if one could add some credibility to the conclusions (in the Arctic especially) in terms of consistency with expectations.

Thank you for this recommendation and we have mentioned the results based on ozone loss from the SLIMCAT model in the last paragraph of the updated sub-section 5.3

- Also, of some interest, is there not a model-derived ozone loss onset date curve that can be compared to the Figure 9 results? What does this (or would this) show? Would this not be a useful comparison for what might be expected? This is not directly tied to ozone loss (but I do understand its connection to this). Any comments about this (in the text or at least as part of a reply) would be appreciated, since this might be worth considering as an added comparison (although not necessarily in this particular work).

This metric is only sensitive to large increases of ozone loss during the winter. This is then appropriate for the SH. The following Figure AR2.1 presents the onset days estimated from ozone loss values obtained from SLIMCAT simulations. The trend estimations were performed before and after 2000. All trends estimated by independently robust linear regression are significant at least 2σ. As for onset datasets from observations ozone loss, the lower trend values are observed for the threshold of 20% and the highest ones for 40% of ozone loss before and after 2000. The positive trends are slightly higher and vary between 4.4±1.0 and 6.1±1.8 day dec^{-1}. The ratio between the trends before and after 2000 of each ozone loss onset dataset is higher compared to the ones obtained from observations. It varies from -0.5 to -0.3. As expected, the model onset dataset presents a faster recovery than the observations (-0.3 to -0.2).

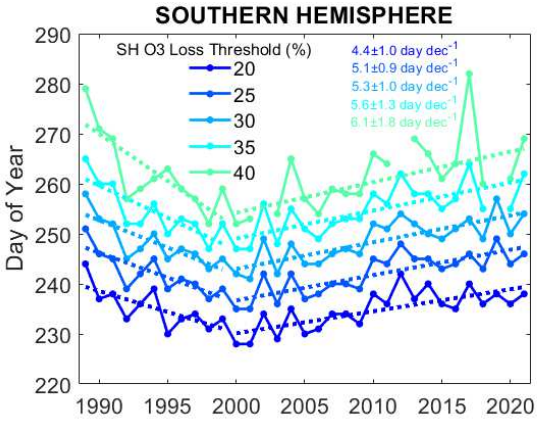


Figure AR2.1: Onset day when 10-day averaged ozone loss reach a particular ozone loss value: 20, 25, 30, 35 and 40% for the SH using SLIMCAT simulations. Robust linear fits before and after 2000 are also shown (dashed lines).

The following sentence was added in L362:

"The onset dataset obtained from SLIMCAT model simulations exhibits larger trends since 2000 that are significant at 2σ (not shown). The trends vary from 4.4±1.0 to 6.1±1.8 day dec^{-1}. The ratio between the trends before and after 2000 of each ozone loss onset dataset vary from -0.5 to -0.3 showing a faster recovery considering SLIMCAT simulations than using the SAOZ-MSR2 merged dataset, as already found using the ozone loss metric 1 (see Sect. 5.1)."

***** More minor corrections and suggestions**

- L22: Add "column" between "cumulative" and "loss" for clarity.

Done

- L23, Abstract: since there are somewhat complicated calculations that involve more than just linear trends, as stated on line 27 (parabolic), this seems somewhat simplified of a description, even if the Abstract has to be simplified and short.

Maybe consider the following wording: "Three metrics are used in trend analyses that aim to assess the ozone recovery rate over both polar regions: ..."

The sentence was modified as suggested. Thank you.

- L28: I think you mean (or should use) standard deviation as part of the error analysis (see other comments above), instead of standard error, or justify the use of this terminology better.

Thank you for this remark. Everything has been adjusted using standard deviation.

- L29: you should be less vague and specify what threshold refers to here, what quantity (ozone column), instead of making the reader guess (if/as the manuscript has not been fully read at this point).

The threshold values were added in the abstract as follows (L29)

"For metric 2, various thresholds were considered **at the total ozone loss values of 20, 25, 30, 35 and 40%**, …"

- L29: I would suggest "all of them showing a time delay as a function of year, in terms of when the threshold is reached."

Adopted. Thank you.

- L32/33: "the difficulty in finding a threshold value in enough of the winters."

Adopted. Thank you.

- L63: wildfires events --> wildfire events

Done

- L72: change "concentrations/columns(?)" to just "columns".

Done. Thank you.

- L120: SAOZ ozone data could be more specific SAOZ total column ozone data?

OK

- L133: delete the period after "used".

Done. Thank you

- L173: delete parenthesis after "merged dataset"; also change "bias" to "biases".

Done. Thank you to highlight all these errors.

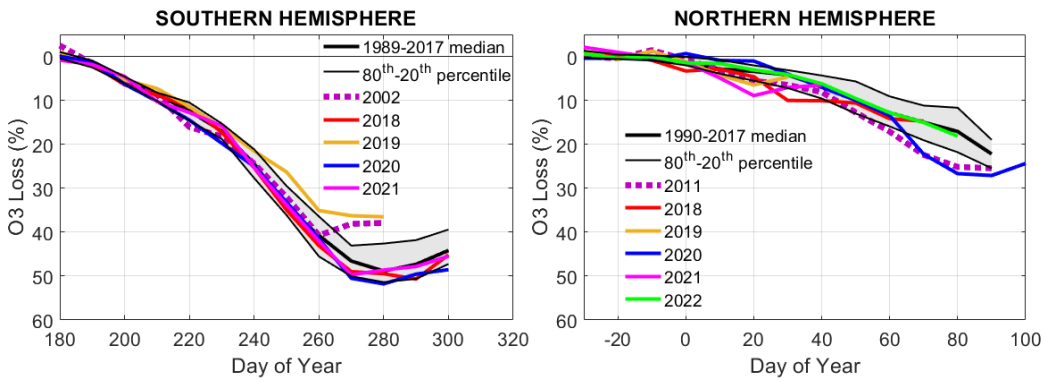
- L176, Figure 1 caption: please specify the year in the caption also (position of the 2021/2022 vortex edge...").

Done

- Figure 4: Please try to plot the thin black lines last, so they can be better seen on both panels. Consider making them slightly thicker as well. Should the y-axis title have a space between O3 and loss (maybe not, if you are referring to a specific variable instead of actual words). Also change "Day of the Year" in both panels (x-axis label) to "Day of Year" (as used in Figs. 5,6,7).

Thank you for your detailed work reviewing this paper. It is much appreciated.

You will find the new Figure 4 here below, following your recommendations.



- L235: For Figure 4, please specify in the text at which day of year the maximum amplitude of ozone loss between recent winters occurs (for both hemispheres). This will help the reader.

Thank you for your suggestion. The information was added in the text (L234)

“The maximum ozone loss is reached between mid-January and the end of March for the NH and between the end of September and mid-October for the SH.”

- Figure 7: vortex edge as defined in Pazmino, but does that follow from Nash et al. (could specify here as well, if so, in addition to the text). Also correct the typo in NH y-axis (Gradiant --> Gradient).

Indeed, the definition of the vortex was done using Nash et al. but here we prefer to specify the definition for the Gradient, as done in Pazmino et al., 2018 which is based partly on the Nash et al. definition.

- Figs. 5,6,7: please change the thin black lines so they are plotted on top, for better visibility (and consider making them slightly thicker as well).

The figures were updated as Fig. 4.

- L253: and so was the vortex stability (rather than "and as well as the vortex stability").

Adapted

- L254: August, linked to a wavenumber 1 event,...

Done

- L255: upper levels, with an associated decrease in size.

Done

- L262: heat flux increases rapidly at the end

Done

- L264, slowing down rapidly thereafter.

Done

- L268: event, of a magnitude similar to the anomalies related to the Calbuco...

Done

- L271: mean T anomaly of -10.1 +/- 4.5 K, as in 2018, but with a much larger sunlit VPSC than in 2018.

Done

- L272: The persistently low temperatures [or The persistently cold lower stratosphere...]

Thank you for the suggestion. The sentence was changed by “The persistently cold lower stratosphere...”

- L273: led to an acceleration of the October ozone loss...

Done

- L276: (Fig. 6), and the strength...

Done

- L280: The sunlit VPSC values are similar...

Done

- L283: Fig. 4), which lies within the climatology.

Done

- L293: Do you mean "The strength of the vortex edge exhibited values larger than climatology...?

Yes exactly. We have changed as suggested:

"The strength of the vortex edge exhibited values larger than climatology.."

- L294: vortex led to moderate ozone loss; also please specify again (if need be) where the ozone loss error bar values come from (or refer back to that discussion)., and if these represent one standard deviation (presumably not two).

The 1σ was added at the first time the ozone loss was mentioned in subsection 4.1 (L250) et 4.2 (L294)

- L300: ozone loss of the 2019 warm winter

Done

- L304: temperature anomalies at the 475 K level... and the mean anomaly value for the whole winter reached -5.3 K, as in 2018.

Done

- L306.307: final warming mode, also shown by the low values...

Done

- L308: persistent low temperatures less than ...

Done

- L319: a possible recovery path of total ozone... [or recovery rate]

The sentence was changed by "… a possible recovery **rate** of total ozone columns …"

- Figure 8: please make the grey legend stand out more; for example, use larger fonts for the legends and say % dec-1 to shorten the units and legend length, and use a bolder font if needed.

The figure was updated as suggested.

- L344: are positive (1.0 +/- 2.2 %dec-1)

Done

- L348: values, as we might expect a later onset... in relation to lower...

Done

- L350: what is the ozone loss time dataset? Is this not the same as the ozone loss onset days "are used instead of total ozone columns"... (?)

The sentence was changed as follows:

"In this study, the ozone loss **onset days** dataset is used instead of total ozone columns **onset days dataset** in order to consider chemical processes only."

- L371: You say a "3rd order polynomial..." and also mention a parabolic fit; to me, a parabolic fit is a 2nd order polynomial (i.e., a quadratic). Please clarify.

The subsection 5.3 was changed. Please see previous answers.

- L369-L371, I would say "dynamic range" or just "range" really; dynamical could appear to refer to something atmospheric...

Thank you for highlighting that. The word "dynamical" was removed.

- Figure 10: It might be interesting to try a linear fit after 2000 for the SH; not necessary for this paper, just a thought (how would that affect the results?).

A linear proxy was considered to represent the sunlit VPSC contribution on Ozone Loss in the SH as adopted in Sect. 5.3 for the NH. A larger negative trend of -2.8 ±0.8 % dec^{-1} was found compared to -2.2 ±0.7 % dec^{-1} using the quadratic relationship. Those values are within the standard deviation. Considering a linear relationship between O3 Loss and sunlit VPSC, the determination coefficient is weaker but still large with a value of 0.78 instead of 0.83.

- L419: was calculated, but it is not significant.

Done

- L420-422: It would flow better if the sentence "Regarding the SH..." was placed one sentence above, directly after the SH comment. Then one could just say "This metric appear sensitive..."

Done

- L433: Note that this trend is similar to the SH trend.

Done

- L436: add a comma after "datasets".

Done