

Reply to review egosphere-2025-6473 referee 1, review 2:

We thank the referee for their complete review of our work and their attention to details which enabled us to improve the quality of our manuscript.

Note that the lines referred to in the review correspond to the first version of the manuscript from 7th January 2026 and are therefore obsolete with respect to the version of the April review which we base our track changes on.

Lines 10-15: trend significance levels missing so it is difficult to assess if these trends could be interpreted as true influences or modifications.

We are mentioning in the text the significance of these trends. Adding the errors would imply that we exceed the fix number of character allowed for abstracts.

Line 21: replace "successfully detected" with "established"

Done

Line 22: replace with "over Antarctica during the Ozone Hole onset period in September" ... recovery over Antarctic has not been established outside of approximately the September month.

Done

Line 23: either delete "until recently" (Weber & WMO 2023 report no signs of Arctic recovery yet) or refer to a few recent papers indicating some changes consistent with Arctic recovery. WMO 2026 will report on this (not yet public) but is very careful in its wording, and it is not that relevant to refer to those recent papers here.

Done

Line 23: signs --> indications

Done

Line 24: observed --> reported

Done

Line 27: about --> approximately ... note that the authors are kindly requested to check the entire for the use of the word "about" (which is more informal) and replace it with "approximately" (which is more formal).

Done

Line 29: "point toward an explanation of" --> "suggest these negative trends could be due to ..."

Done

Line 31-32: change to “Ozone layer depletion as is common over Antarctica does occasionally occur over the Arctic as well - unlike mid-latitudes - due to colder ...”

Done

Line 33: “At springtime” --> “During springtime”

Done

Line 34-35: Suggestion to delete the sentence starting with “The stratospheric cooling ...”.

It is not necessarily true (for Antarctica this is (almost) not the case as it is always cold enough for PSC formation and a little bit colder does not change things that much while ozone depletion there is limited by other processes that saturate so that colder and more PSCs does not lead to more depletion. Furthermore, more cooling also affects atmospheric chemistry and reaction rates which also alter stratospheric ozone photochemistry and can lead to less depletion, not more depletion (see for example the phenomenon of super-recovery).

Done

Line 35: If the previous sentence is removed then "Moreover" should also be removed.

Done

Line 37: “expected stratospheric ozone trends”

Done

Line 40: Paralelly to the stratospheric ozone ... --> In addition to stratospheric ozone ...

Done

Line 43: tropospheric ozone trends in ppbv/decade. It would be helpful to try to translate that already here to a tropospheric column change in Dobson Units.

Done

Line 44: change to “... while Law et al. (2023) reported a strong seasonal cycle in tropospheric ozone trends, with ...”

Done

Line 46: Suggestion to start this section with one sentence describing the aim of the study. For example: “In this study we analyze different atmospheric ozone observational records over the Arctic for signals of recovery.”

Done

Line 65: “we focus on the period 2000-2024”. Add a justification on why this period is chosen (for example: it is well established that atmospheric concentrations of ozone depleting substances have decrease since approximately the year 2000. One or multiple recent Ozone Assessments could serve as reference if the phrase “well established” is used.)

Done

Line 94: change to “Appendix of Bjorklund et al. (2024)”

Done

Line 110: about --> approximately

Done

Line 112: delete the comma between active and correcting

Done

Line 122: delete "lower", if a reference is to be made to another part of the paper, use "next in Section 2.4".

Done

Line 124: It appears that what is meant here is that "If extreme outliers are found ... partial columns, they are removed ..." or "Extreme outliers that are found ... are removed ...". Otherwise it is a statement about a result that does not connect to the next sentence. Please clarify and/or correct.

Done

Line 129: change to “use sunlight in the UV range (approximately 350-340 nm) to ...”

Done

Line 136: delete “estimate”

This we don't agree with. The standard error is given by a statistical estimate in the mathematical sense.

Line 137: Recommendation to remove the focused-moon measurements from the record, they generally are not used due to questions about their quality and lack of assessment of their quality.

After discussing with the PI of each specific datasets, we have decided to include the focused-moon measurements with larger error-bars, which enables to compensate for their downgraded quality with respect to other measurements. This is in part motivated by the polar night above the Arctic circle where no sun measurement is available for several months.

Line 148: preferably make this claim explicit by adding a few references and also briefly describe per season what typical tropopause heights are.

We have extended the explanation of why we make this choice and we have added a suitable reference for the averaged tropopause limit in the Arctic (Zängl and Hoinka, 2001).

Figures 4+5+6: Suggestion to add the time series (as in Figure 2) for each station separately to the appendix. It is quite informative to have that information available visually as a time plot. And as this is a paper already rich in data and information, it is OK to have an extensive appendix.

We disagree with this proposition that would amount to more than 120 additional figures in the appendix. Note however that all the partial columns datasets are accessible on our website at <https://dora.aeronomie.be/index.php/data>

Line 215: Personally I would leave out referring to the 5 sigma level, a reader should be able to determine that from the numbers themselves. It also now gives rise to questions: 11.51/3.58 and assuming 3.58 is the 2-sigma error actually is > 6-sigma (6.43), so why the 5-sigma? There is no particular reason why 5-sigma is relevant. And one would round the sigma values to whole numbers, so in this case 6-sigma. Suggest to change to "... with a statistically highly significant drift ..." or even just leave it as it is and delete the "significant beyond the 5-sigma level". The "clearly" already signals there is very likely a problem with that time series.

Done

Figure 3: Resolution of the image is OK for a pre-print and a review but presumably will be improved if the paper is accepted.

This will be handled on a later stage of the review.

Lines 222-224: Which sonde type and is there literature supporting this notion that measurements from the two sonde types systematically differ?

Also, the papers by Stauffer et al. [2020, <https://doi.org/10.1029/2019GL086791>; 2022, <https://doi.org/10.1029/2022EA002459>] are not referenced. Both papers explore the long term stability of ozone sonde measurements due to a reported "sudden post-2013 TCO "dropoff" of ~3%–4%" and an "ozonesonde TCO drop is 3–7% compared to satellite and ground-based TCO, and 5–10% or more compared to satellite stratospheric O3 profiles". Scoresbysund is one of the stations for which a TCO dropoff is report (5.6%) based on a comparison with OMI.

The positive drift in sonde-IASI tropospheric ozone comparison for Scoresbysund should be discussed within the context of the findings of Stauffer et al. [2020, 2022].

We thank the referee for pointing this out. As a matter of fact, the ozonesonde type has not changed in 2016. In the early years of the Scoresbysund time series, they switched several times between the two ECC ozonesonde types, SPC and En-Sci, but since mid 2001 onwards, Scoresbysund is consistently using En-Sci. With an identical sensing solution strength, a 3-4 % bias in the ozone concentration readings exist between both ozonesonde types, as observed in dual soundings and simulation chamber intercomparison campaigns (Smit et al., 2007, 2014). Therefore, it is recommended to use a different sensing solution strength for both ozonesonde types (1% KI Full Buffer, also noted as SST1.0, for SPC and 0.5% KI Half Buffer or SST0.5 for En-Sci). The lower sensing solution strength gives again a 3-4 % lower ozone readings than the higher sensing solution strength, compensating for the higher En-Sci ozone readings than the SPC measurements. Throughout the entire Scoresbysund time series, the same sensing solution strength (SST1.0) has been used, which gives a 3-4 % high ozone bias for the En-Sci measurements with respect to the recommended combinations, which can be traced back to the reference ozone photometer in the simulation chamber in Jülich (the ozonesonde network reference). Transfer functions to those recommended combinations do exist (Deshler et al., 2017) and are included in the Standard Operating Procedures (SOPs) for ozonesonde measurements (Smit et al., 2014, 2021). In Scoresbysund, these transfer functions have been operationally implemented in the standard processing since 2016 onwards, but the historical time series (before 2015) have not been re-processed. As a consequence, the Scoresbysund non-homogenized data present in the data archives (NDACC, WOUDC) and used for the Stauffer et al. [2022] paper has a 3-4% low ozone bias introduced after 2016 with regard to the period therefore (e.g. from mid 2001 up to 2015), at least partially responsible for the so-called TCO En-Sci drop off of -5.6% mentioned in those papers.

However, the data used in the present manuscript has been homogenized, e.g. the transfer functions to one of the standard combinations (SPC SST 1.0 or En-Sci SST0.5) have been implemented for the entire time series if required, apart from some other corrections (pump temperature corrections, pump flow rate corrections, pump efficiency corrections, background current subtraction corrections), based on the available metadata. With those corrections, a large part of the ozonesonde low bias after 2016 has been resolved, but not entirely (see e.g. <https://ozone.meteo.be/uploads/media/66ab5646e95ba/qos2024-vanmalderen6-49.pdf?v20250829-1211>). The change in 2016 do more or less coincides with the En-Sci serial number change (S/N < 25250 vs. S/N >25250), but also with the other instrumental changes mentioned: radiosonde type, interface type (and hence pump temperature sensor type), pump battery type, etc. Especially the change in the pump temperature vertical profile after 2016 is remarkable, and the pump temperature measurement has a direct impact on the ozone partial pressure measurement of the ozonesonde. A recent paper on the OHP ozonesonde time series (Okamoto et al., 2026) struggles with the same feature (too low pump temperatures recorded during a certain

time period) and proposes a correction method, based on co-located Lidar vertical ozone profile measurements. Such an approach is not possible at sites lacking those co-located measurements.

We have modified the text in the manuscript according to this and added the relevant literature.

Line 252: change the first part of this sentence to "For assessing satellite drift in the Arctic zonal band, we average all ..."

Done

Line 252-253: remove the brackets and change to "... mean with weights given by ..."

Done

Line 260: It appears this provides a justification for not doing a more precise determination of the IASI-CDR tropospheric drift. If so, add that, for example by noting "which is beyond the scope of this paper which focuses on stratospheric ozone".

Not relevant anymore since the April review

Line 264-265: suggestion to have all brackets removed here, it makes it less readable. The text is fine without all the brackets. Add a comma before "respectively".

Done

Line 265: about --> approximately

Done

Line 268: Change to "The MEGRIDOP data shows a positive drift with ..."

Done

Lines 272-275, 282, 283, 288: similar to one of the previous comments, remove all the brackets.

Done

Line 288: add a comma before "respectively"

Done

Line 295: "square kilometers"

Done

Line 295-296: change to "small relative to the large ..."

Done

Line 301: “(2019) of ozone” --> space missing between bracket and “of”

Done

Line 313: change to "in the Arctic between 60N-90N" plus add degree signs

Done

Lines 314-316, equation (7): Delete this equation: the Pearson correlation coefficient is so well-known and widely used that there is no need to define it.

Done

Lines 328-329: change to " we define a group of stations as having CAMS correlations of member stations higher than $r > 0.8$

Done

Line 358: change to “Long term trends in Arctic stratospheric ozone due to changes in ozone depleting substances are expected ...”

Done

Line 370: change to “... during winter, following Brunner et al. (2006)”

Done

Line 376: change to “... compared to Vigouroux et al. (2015)”

Done

Line 390: note a big fan of this type of writing (“Let us now ...”). Rather keep it more formal, for example “Figure 13 shows the total coefficient of determination R^2 ...”.

Line 393: Similar here to the previous comment online 390. If there is a need to announce what will be done, announced that in the first paragraph of this section, for example “in this section we will explore this-and-that” or “in this section, we first ... then ...”.

We thank the referee for this idea, we have reformulated this whole section by first outlining the different steps in the intro and removing all less formal formulations, hopefully making it more readable.

Line 397: about --> approximately

Done

Line 414: this is not really the volume of PSCs but the potential volume, so please rephrase.

Done

Line 415: cloud --> clouds

Done

Line 417: about --> approximately

Done

Line 420: change to “form predominantly between ...”, they can occur up to 30 km altitude but occurrence above 25 km is rare, see Pitts et al., [2018], <https://doi.org/10.5194/acp-18-10881-2018>.

Done

Line 422: change start of the line to “For the total column”

Done

Line 423: change to “As with partial columns”

Done

Line 427: change to “much less well explained”

Done

Line 429: “during spring”

Done

Line 430: the use of “drives” sounds awkward, please rephrase

Done

Lines 432-433: change to “in particular the BDC”

Done

Line 433: if this is correct one then would expect some correlation between VPSC and BDC, so the suggestion is to add here the correlation value between both for the proxies used in this study.

Done

Lines 438-439: This begs for an explanation: why do results differ on this point from Bernet et al. (2023)?

In Bernet et al. 2023, the regression procedure did not use a stepwise regression. Therefore, and contrary to our model, they need to exclude by hand those proxies that are too highly correlated. In fact, they emphasize that the VPSC did improve R^2 in the case of Ny-Alesund, but they nevertheless decided to not use it. Note that they are also

working with only 3 stations and only total columns, therefore their appreciation of which proxies to prefer can naturally differ from ours. We have made this more explicit in the text.

Line 444: delete “due to climate forcing”, definitely not everyone is convinced about that interpretation

Done

Line 451: add which errors are referred to (presumably the errors associated with the trend calculation)

Done (error on each trends).

Lines 452-453: it is not fully clear what the "absolute difference" refers to; presumably the difference in ozone trends without detrending and with each proxy is detrended (the text mentions only "impact on effective ozone levels and conclusions"). Please clarify.

We thank the referee for pointing this as unclear, we have clarified in the text what we mean by absolute difference.

Lines 464-467: This should be moved to earlier in the section (maybe even to the start of the section at line 452)

We don't agree.

Line 501: The data in this study covers the period 2000-2021 so the effect of an additional 3 years as a possible explanation for the differences in trend significance can be assessed (and is not unimaginable, for a period of 21 years a few additional years can push trend significance levels over the threshold when significance levels are near 2-sigma).

Done

Line 505: change to "Trends for Lapland are, however, positive ...".

Done

Line 508: change to “indicates” (“means” is a too strong conclusion based on only a statistical analysis).

Done

Lines 515-516: suggestion to modify this, ("These results agree within the 2-sigma error margin.") ... the agreement within error bounds means that differences should not be considered relevant.

Done

Line 547: change to "especially during Spring."

Done

Line 556: change to "for the North Pole"

Done