Reply to Anonymous Referee #1 review of the manuscript acp-2025-3963

Evaluation of factors affecting TOC and its trend at three Antarctic stations in the years 2007–2023

David Tichopád on behalf of all co-authors

We sincerely thank Anonymous Referee #1 for the time dedicated to reviewing our work and for the constructive feedback. Your valuable comments have significantly contributed to the improvement of our manuscript.

Please find our answers below (in red)

In this study, the authors present the trends of total column ozone at three Antarctic ozone monitoring stations over the period 2007-2023, both annually and during the spring months. They also discuss the impact of several dynamical processes on Total Column Ozone variability at the three monitoring stations and over the Antarctic more broadly using the MERRA-2 reanalysis. They present trends derived using the LOTUS (Long-term Ozone Trends and Uncertainties in the Stratosphere model) for the first time on ground-based and satellite overpass datasets in Antarctica. Finally, the authors offer a detailed analysis of the TCO variability during the 2019 and 2020 ozone holes. I think this is a very wonderful and timely study, as complications surrounding the identification of ozone hole recovery in Antarctica due to dynamic variability are still being discovered and actively discussed. This study fits ACP very well, and this reviewer recommends publication after revisions.

We sincerely thank the reviewer for their positive and encouraging assessment of our work. We are glad that the study is considered timely and relevant, and we appreciate the acknowledgement of the importance of addressing the complexities associated with identifying ozone hole recovery in Antarctica.

General comment about the logical structure and flow of the paper:

The results section is a little hard to follow because it is organized by station (i.e., 4.2 Marambio, 4.3 Troll, 4.4 Concordia), but there are discussions in some sections that contain results from all three stations. For example, the statistical tests (Durbin-Watson test) on lines 219-22, the discussion of the standard coefficients in lines 236-241 and again in 291-298. An alternative suggestion is to organize this section not by station, but by topic: trends, statistical tests, and then proxy regression coefficients.

We thank the reviewer for this constructive suggestion. We agree that organizing the results section by topic rather than by station can improve clarity and readability. In the revised manuscript, we have restructured the section to present the results under the headings "Trends," "Statistical Tests," and "Proxy Regression Coefficients." We believe this new organization allows for a more coherent comparison across all three stations and makes the key findings easier to follow.

Along the same lines, including a single table for the trends would enable an easy comparison. The overall trend estimate is a major result of this paper, and it was difficult to find and compare at first glance. One single table reporting the trend, uncertainty, p-value, and adjusted R^2 would be very helpful.

We agree that a single table summarising the trends, uncertainties, p-values, and adjusted R^2 would greatly improve clarity and allow for easier comparison (Tab. R1). A new table has been added to the Results section in the revised manuscript to present these key results.

Tab. R1 Linear trends of TOC at the three Antarctic stations (Marambio, Troll, and Concordia) in 2007–2023. The table presents the estimated trend (DU/decade), the associated uncertainty, the p-value, and the adjusted R^2 for each station. A statistically significant trend is marked in bold ($p < 0.05$)

Station	Trend [DU/decade]	Uncertainty [DU/decade]	p-value	adjusted R^2
Marambio	**3.43**	±3.22	0.04	0.94
Troll	-1.09	±3.91	0.58	0.97
Concordia	1.15	±4.25	0.59	0.95

General comment about the conclusion:

The conclusion lacks a discussion about the broader implications and interpretation of the trends: what does it mean for monitoring for the ozone hole recovery identification and uncertainty? It is a good and interesting result, and the authors should discuss how it will impact the larger discussion of ozone hole identification. How should the reader interpret the trends? For example, the LOUTS model (and other MLR models) is often used as a method to ascribe dynamical and natural TCO variability so that the resulting trend can be interpreted as ozone-depleting gas (ODS) chemistry-related. Is that the goal in this study, and how should the reported trend be interpreted by the reader?

We agree that discussing the broader implications of the trends is essential. In the revised manuscript, we have expanded the Conclusion section to clarify how the reported trends can be interpreted in the context of ozone hole recovery and associated uncertainties.

General comment about limitations.

The authors could expand upon the limitations in this study and how they may impact the interpretation. First, what is the uncertainty of the results and conclusions due to the temperature and cross sections used in the instruments, and from the merging of the satellites with the ground-based instrumentation? How does this impact the seasonal trend? Second, what is the impact of not using the LOUTS model on the monthly spring trends (September, October, November)? The authors are recommended to discuss how 'simple' linear trends do or do not compare with studies that use dynamical proxies in their trend analysis. Finally, the authors start the regression analysis in 2007, which is different from the LOTUS (SPARC, etc) report and most other LOTUS studies, which report the post-2000 trend. When comparing this study's results to others, this is an important caveat.

Thank you for this comment, and we will respond point by point:

1) We calculated the LOTUS regression (Tab. R2) for, in addition to the compiled time-series (used in the study), separately for the OMI data and the MERRA-2 data (the same months were used). In the case of the Marambio station, the trend is statistically significant for all datasets, but the magnitudes of the trend are slightly different. Also, in the case of the other stations, where the trends are statistically insignificant in all cases, the magnitudes of the trends differ slightly. These differences can be caused by the bias between ground measurements and the OMI and MERRA-2 data, which in mean reaches 1-3%. However, it reaches values up to 10 % within months (most at the Concordia station).

Tab. R2 Linear trends of TOC at the three Antarctic stations (Marambio, Troll, and Concordia) in 2007–2023 compiled time series, OMI and MERRA-2 data. The table presents the estimated trend (DU/decade), the associated uncertainty, the p-value, and the adjusted R^2 for each station. A statistically significant trend is marked in bold ($p < 0.05$).

Station	Fit results	Trend [DU/decade]	Uncertainty [DU/decade]	p-value	Adjusted R^2
Marambio	Compiled	**3.43**	±3.22	0.04	0.94
	OMI	**4.58**	±2.97	0.00	0.95
	MERRA-2	**4.20**	±3.00	0.01	0.95
Troll	Compiled	-1.09	±3.91	0.58	0.97
	OMI	2.42	±3.30	0.15	0.98
	MERRA-2	1.59	±3.31	0.34	0.98
Concordia	Compiled	1.15	±4.25	0.59	0.95
	OMI	1.47	±4.61	0.53	0.95
	MERRA-2	0.47	±4.40	0.83	0.95

The uncertainty related to the temperature dependence of absorption coefficients and the choice of cross-section datasets for Brewer instruments is small, typically below about 1 %. The temperature effect is <0.01 % K^{-1} (Redondas et al., 2014), resulting in a negligible impact on the derived DS ozone trends.

For cloud-free conditions and SZA < 70 degrees the overall uncertainty in NILU-UV is estimated to be ±5% (Sztipanov et al., 2020). For SZA > 70 degrees the impact of cloudiness, the vertical profile of ozone and temperature, the imperfect cosine response of the instrument, and the absolute calibration error will reduce the accuracy of the TOC values (Sztipanov et al., 2020; Kazantzidis et al., 2009).

The SAOZ instrument uses Bogumil et al. (2003) ozone cross sections in the visible, where temperature dependency is practically negligible. The systematic uncertainties on the ozone absorption cross sections, considering slight dependence on temperature, are approximately 3% in the SAOZ spectral range (Orphal, 2003). A complete budget estimation can be found in Hendrick et al. (2011).

This limitation is, however, now noted in the manuscript.

2) The LOTUS regression could not be applied for the monthly spring trends (September, October, November) because the variance inflation factor (VIF) exceeded 10, indicating strong multicollinearity among the predictors. As a result, assessing individual predictor contributions using LOTUS was not feasible for these months.
3) The comparison of our linear trends with other studies using dynamical proxies is addressed by comparing the TOC trends derived from our analysis with those obtained from OMI overpass and MERRA-2 datasets using the LOTUS regression. The results of this comparison are presented in a table in the Supplement, providing context for how our post-2007 trends relate to other studies that report post-2000 trends.

Specific Comments:

Line 50, Johnson is spelled incorrectly in the citation.

Corrected

Line 58: Since this is the first introduction to the LOUTS model in the paper, I would recommend moving the citation (SPARC/LOTUS 2019) to here.

Corrected

Line 88-90: The authors could also include information about the ozone cross sections and effective temperatures for the other instruments in addition to the SAOZ instrument. I would also recommend a short discussion of the impact of the chosen cross-section in the TCO retrieval on the trends, especially for Antarctic springtime ozone trends. Please see Fragkos et. al. 2013 and Rodeondas et.al 2014. Redondas et al. (2014) point out that the Brewers (as of the writing of their paper), in particular, do not use a correction factor for the effective temperature where SAOZ does. If the authors think it appropriate, please discuss the impact the cross section and assumed effective temperatures may or may not have on the trends, particularly the seasonal (springtime) trends.

Thank you, as we said earlier in the general comment: The uncertainty related to the temperature dependence of absorption coefficients and the choice of cross-section datasets for Brewer instruments is small, typically below about 1 %. The temperature effect is <0.01 % K^{-1} (Redondas et al., 2014), resulting in a negligible impact on the derived DS ozone trends. This limitation is now noted in the manuscript.

Line 135, Did you mean 25 missing measurements (like in the Bernet paper)? Or was there a reason not to include months with more than 5 missing measurements?

We used a threshold of up to 5 missing daily measurements per month to ensure representative monthly averages. In contrast, Bernet et al. (2023) excluded months with fewer than 25 measurement days, which is conceptually similar. This choice was made to balance data coverage and representativity; months exceeding the threshold were excluded (April to August for all three stations). A clarification has been added to the manuscript.

Line 142: Please include that the primary objective of LOTUS in the first phase was to derive the trends using global satellite datasets. Now in the third phase, there is a focus on trends

derived from individual monitoring stations, regional variability, and representation of dynamic and physical processes.

Thank you, included.

Line 160: Please discuss how seasons with missing data and merging satellite with ground based instruments could bias the harmonic fitting in the LOTUS model.

Line 164-170: This section has some repetitive information about the construction of the QBO proxies. It repeats some of the same wording from the other studies (particularly the Bernet 2023 paper). I would suggest making this section more concise and citing the other papers instead of repeating their discussions here.

Corrected

Line 172: Can you clarify the statement that there are no statistically significant trends in the predictors? Your caption for figure S1 states: "Time series of predictors supplemented with a linear trend (red; except for the solar and QBO predictors), which was statistically insignificant at the significance level of $p = 0.05$.") This leads this reviewer to believe there are significant trends in all the predictors but the solar and QBO. This reviewer expected that there should be a trend in at least some proxies. Also, please clarify if the predictors were or were not detrended prior to the regression? It will have an impact on the interpretation of the trends.

The predictors were not detrended prior to regression because linear trends were not statistically significant in any of the predictors ($p > 0.05$). The exception is the solar cycle (Solar), which exhibits an ~11-year periodic cycle that is not linear, and removing it would eliminate any physically relevant signal. Therefore, detrending was not methodologically justified and could have biased the interpretation of the results.

Line 180: One suggestion, if you have time (or maybe this is for a separate paper), is to look at the trends in just the overpass dataset. The comparison between the trends derived from the overpass dataset alone and the trends derived from the ground-based supplemented with the overpass dataset would give an indication of the uncertainty in the trend due to any inhomogeneities between the GB and overpass datasets (up to 9% for some months in your own analysis).

We thank the reviewer for this suggestion. We are currently planning to prepare a separate manuscript, in which we will compare trends derived from the overpass dataset alone with those derived from the ground-based dataset supplemented with the overpass data.

Line 218-222: Since this section discusses all three stations, perhaps it's best to move it from the Marambio station section. Alternatively, move the discussions of each individual station to their respective sections. However, I would suggest the former.

We thank the reviewer for this suggestion. This issue has already been addressed based on our previous general comment regarding the logical structure and flow of the manuscript.

Lines 287-298: Please consider clarifying this section. The authors appear to be comparing their results to the results of Bernet (2023). However, it is a little confusing which parts were discussed by Bernet (2023) and which were the results from this study. It seems to be line 291, but please clarify.

Thank you, corrected.

Line 359: Johnson is spelled incorrectly

Corrected

Line 431; NOAA PSL, 2025, in the citation for the NASA OMI link should be removed.

Corrected

Line 437: Please give the link to the LOTUS model code here: https://usask-arg.github.io/lotus regression

The link to the LOTUS model code in R will be added to the manuscript during the revision process.

References:

Bernet, L., Svendby, T., Hansen, G., Orsolini, Y., Dahlback, A., Goutail, F., Pazmiño, A., Petkov, B., and Kylling, A.: Total ozone trends at three northern high-latitude stations, Atmos Chem Phys, 23, 4165–4184, https://doi.org/10.5194/acp-23-4165-2023, 2023.

Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., Vogel, A., Hartmann, M., Bovensmann, H., Frerik, J., and Burrows, J. P.: Measurements of molecular absorption spectra with the SCIAMACHY Pre-Flight Model: Instrument characterization and reference spectra for atmospheric remote sensing in the 230-2380 nm region, J. Photochem. Photobiol. A, 157, 167–184, 2003.

Hendrick, F., Pommereau, J.-P., Goutail, F., Evans, R. D., Ionov, D., Pazmino, A., Kyrö, E., Held, G., Eriksen, P., Dorokhov, V., Gil, M., and Van Roozendael, M.: NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations, Atmos. Chem. Phys., 11, 5975–5995, https://doi.org/10.5194/acp-11-5975-2011, 2011.

Kazantzidis, A. Bais, A. F., Zempila, M. M., Meleti, C., Eleftheratos, K., Zerefos, C. S.: "Evaluation of ozone column measurements over Greece with NILU-UV multi-channel radiometers," Int. J. Remote. Sens. 30, 4273-4281, 2009.

Orphal, J.: A critical review of the absorption cross-sections of O3 and NO2 in the 240-790 nm region, J. Photochem. Photobiol. A: Chemistry, 157, 185–209, 2003.

Redondas, A., Evans, R., Stuebi, R., Köhler, U., and Weber, M.: Evaluation of the use of five laboratory-determined ozone absorption cross sections in Brewer and Dobson retrieval algorithms, Atmos. Chem. Phys., 14, 1635–1648, https://doi.org/10.5194/acp-14-1635-2014, 2014.

Sztipanov, M., Tumeh, L., Li, W., Svendby, T., Kylling, A., Dahlback, A., Stamnes, J., Hansen, G.H., Stamnes, K.: Ground-based measurements of total ozone column amount with a multichannel moderate-bandwidth filter instrument at the Troll research station, Antarctica, Appl. Opt., 59, 97-106, doi: 10.1364/AO.59.000097, 2020.