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.

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.

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.

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?

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.

Specific Comments:

Line 50, Johnson is spelled incorrectly in the citation.

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.

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.

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?

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.

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.

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.

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).

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.

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.

Line 359: Johnson is spelled incorrectly

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

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