Thank you to both of the Reviewers for the thoughtful suggestions on our manuscript! We have updated the paper accordingly, and replies to specific questions are provided here in blue text.

RC1

Review of "Stratopause trends observed by satellite limb instruments" by Dubé et al.

This manuscript presents trends in stratopause height and temperature for 2005-2021 from the OSIRIS instrument and compares them with results from SABER. Overall, the manuscript is well written, and the results are sound and supported by the data and methods. The results presented are relevant to the global context of the study of the impacts of climate change in the middle atmosphere, where data scarcity is often a problem. Therefore, I recommend the publication of the manuscript in ACP after some revisions.

My main concern is:

- In subsection 2.5, the authors explain their method to compute the stratopause location based on the SPARC/IO3/GAW report. They acknowledge that this is based on the properties of ozone in the middle atmosphere. However, they do not clarify why this method is necessary, or whether it is better or worse than the simple determination based on temperature, which is the method proposed by the WMO. If the authors prefer, for whatever reason, the method based on the MLR equation presented, the differences in both methods should be discussed in the context of their needs and benefits for the computation of stratopause trends presented (not related to ozone), and a comparison of the results with both methods should be presented.

We believe there is some confusion here. The stratopause location itself is determined only based on the location of the temperature maxima in observations, as described in Section 2.4. The method from the SPARC report is only used to calculate the trends in the observed stratopause height and temperature.

Other issues to be addressed:

- In line 5 of the Introduction, the authors state that "little attention has been given to the stratosphere". Well, it can be true compared to the study of surface climate, but

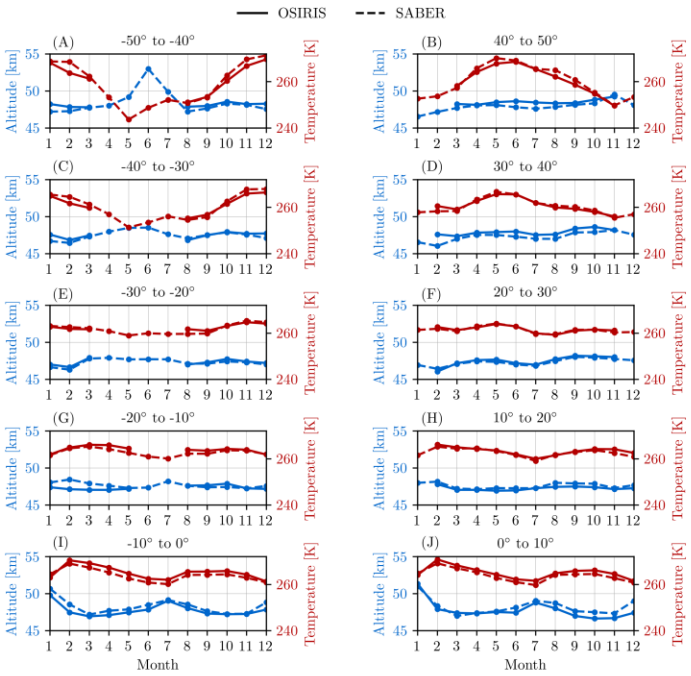
considerable research on the stratosphere has been developed over the last decades, with substantial efforts such as all the work developed in the framework of the APARC, something that some of the authors are very aware of. The statement is made after referring to the changes in the limiting layers, the troposphere, and the mesosphere, and can translate to the wrong view that nothing has been done on the stratosphere, which is not true. I think that for a balanced discussion, it is necessary to make a specific mention here to the work on stratospheric contraction by Pisoft et al. (2021), which the authors cite later in the text.

Thank you for the comment. We did not mean to refer to the stratosphere as a whole, but only to observational studies of stratopause height trends. This statement has been removed to avoid any confusion, and the Pisoft et al. (2021) study is now cited. The updated sentence is "There is observational evidence of both tropospheric expansion (e.g., Santer et al., 2003; Seidel and Randel, 2006; Meng et al., 2021) and mesospheric contraction (e.g., Zhao et al., 2020; Mlynczak et al., 2022; Liu et al., 2024), and models project that stratospheric contraction should also occur (Pisoft et al., 2021)."

- The caption of Fig. 1 should better explain the contents presented in the plots. For example, what color corresponds to altitude, and which one to temperature? Also, I recommend keeping the vertical axis at a fixed range for altitude across all plots. In this way, the plots would directly translate at a glance if the stratopause is higher or lower.

The figure caption now mentions that red corresponds to temperature and blue to altitude.

While a consistent vertical axis is helpful for comparing the panels quickly, the temperature and height variability is much smaller in some of the latitude bands than in others. If the y-axis is set to the range needed by the 50S-40S panel, it is very difficult to see the seasonal cycle and the difference between SABER and OSIRIS at e.g. 20S-10S (see the figure included below for reference). For this reason, we find it more helpful to use varying axis limits, despite the fact that the plot may take more time to understand.



- The first paragraph on page 8, after Fig. 4, is badly explained. The authors could try to rewrite it to better explain that they are comparing SABER data across different periods, the rationale for doing so, and why they are comparing them to OSIRIS.

Thank you for the suggestion. The OSIRIS temperature observations are more limited than those from SABER, both with regards to the latitudinal sampling pattern and the time period that has useful observations for trend studies. Because of this, we use SABER to get an idea of (1) how the OSIRIS sampling pattern, and particularly the lack of measurements at mid-latitudes during the winter, affects the trends, and (2) how the choice of time period used for the trend calculation impacts the trends.

To look at the sampling effect, the SABER trends in Fig.4 were changed from those in Fig 3. by including months when OSIRIS did not measure (these months were excluded from SABER before calculating the trends in Fig. 3). It is this comparison that we attempted to explain in the first paragraph on page 8. To help clarify things, we have updated Fig. 4 to include an additional line, showing the same SABER trends from Fig.3, which are impacted by the OSIRIS sampling. The text discussing the figure has also been rewritten.

- Could you elaborate on the reason for the positive trend of altitude observed in Fig. 4 at 50ºN? Is it instrumental? Real climate variability?

By comparing the two time periods in the figure, the trend for 2005-2021 becomes less positive (closer to zero) when the analysis is extended to 2002-2024. The slightly positive

trend is therefore likely a feature of the interannual variability, and we expect the trend at this latitude will become negative in the future. This is now mentioned in the manuscript.

- Page 9, the last paragraph in section 3: It would be good to state how the obtained trend values compare to previous ones in the literature.

Unfortunately, our ability to compare directly with earlier results is limited as the Pisoft et al (2021) study focuses on global trends from 1960-1999 and 2000-2080. We have nonetheless added a statement that says “The trends from the CCMI-2022 models are also comparable to the global trends from Pisoft et al. (2021), who found a global decrease in the stratopause height of ~60 m/decade between 1960 and 2080 in the older CCMI-1 models.”

- In the conclusions section, the authors claim to present "the first comprehensive analysis of stratopause height trends between 60ºS and 60ºN". I think this is an unnecessary (it adds nothing scientifically relevant) and bold statement, and it should be removed. First, I do not think that an analysis limited to the tropics and extratropics, and covering only a 15-year period, can be considered "comprehensive". Also, in the manuscript, the authors refer to previous works that include results on trends in the stratopause.

The statement has been edited to just say “We present an analysis of stratopause height trends 60ºS and 60ºN…”

- The Code and Data Availability section includes a link to GitHub for the LOTUS regression code. GitHub is not a frozen repository; it is not reliable for long-term availability, and it is recommended that its users create Zenodo repositories to store code used for scientific purposes and publication. Therefore, although ACP does not currently enforce a strict policy on it, I recommend that the authors deposit the LOTUS code in a Zenodo repository, which, among other benefits, will provide a DOI and a proper way to cite it.

While it is important to have accessible code, the LOTUS regression model was not developed for this paper and was only the package used to calculate the trends. The developers of the code package would have to decide to upload it themselves. We did add a reference to the version of the code package that was used for the trend calculations to the manuscript, which should help with reproducibility.

RC2

The manuscript Stratopause trends observed by satellite limb instruments by Dubé et al. examines trends in temperature and stratopause height over the 2005-2021 period using

OSIRIS observations. The results are compared with SABER measurements and chemistry-climate model simulations. The authors apply a newly developed OSIRIS temperature retrieval, allowing a latitude-resolved observational assessment of stratopause height variability and trends. Comparison of these trends with chemistry-climate models provides useful context for model representations of the middle atmosphere. Together, these aspects make the study a valuable contribution to middle atmosphere research. I recommend publication in Atmospheric Chemistry and Physics after the comments below are addressed.

Page 1, line 15: The statement "little attention has been given to the stratosphere" is bit misleading, as it is immediately followed by a discussion of the importance of stratospheric temperature trends, which have been investigated in several earlier studies. The authors should clarify that the lack of attention refers specifically to observational studies of stratospheric height rather than to stratospheric trends more generally.

Yes, we agree that this can be misleading. At your suggestion, and that of Reviewer 1, the statement has been removed completely to avoid causing any confusion.

Page 4, line 101: The stratopause is identified as the local maximum using cubic interpolation of monthly zonal means. Although cubic interpolation is reasonable it can introduce small oscillations, especially near boundaries or when gradients are weak. Have the authors tested whether the diagnosed stratopause height and trends are sensitive to the choice of interpolation method? Brief discussion or comparison with different method would help demonstrate robustness.

We also tried calculating the OSIRIS stratopause height using 2^{nd}-5^{th} order spline interpolation. This choice of order does change the stratopause height of some of the monthly zonal mean profiles compared to a cubic spline interpolation (30% for 2^{nd} order, 11% for 5^{th} order), but the corresponding trends are unaffected. We therefore believe that cubic interpolation is adequate for our purposes.

Page 4, Section 2.5:

Please state in Section 2.5 that LOTUS model was used (and version), and list the source of proxies implemented.

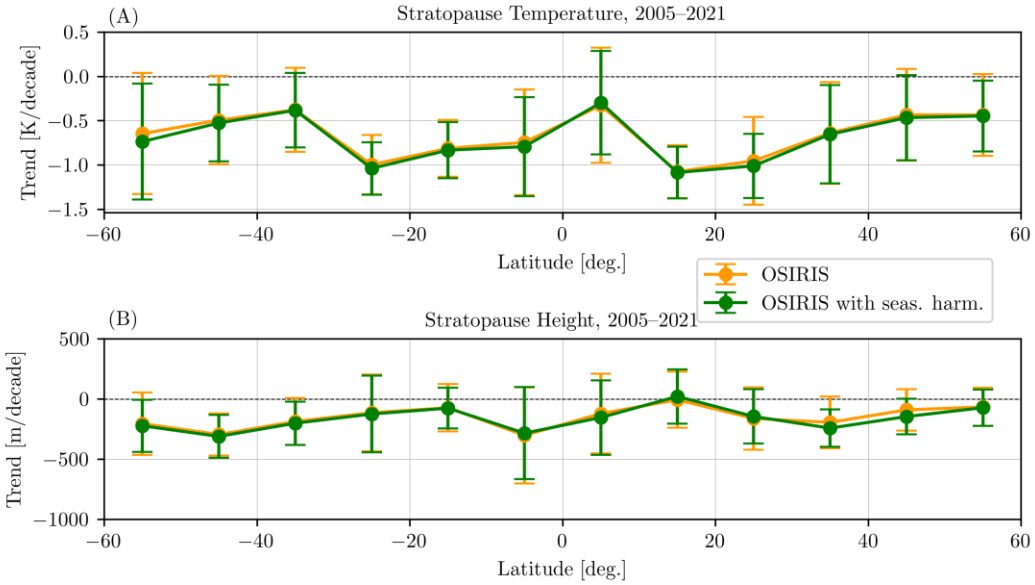
This information is now provided in the text.

It would be helpful to state explicitly that the MLR is applied to the deseasonalized stratopause height and temperature anomalies described in Section 3.1.

This is now mentioned in the text.

The MLR equation includes QBO and ENSO as regression terms with constant coefficients which means they are treated as season-independent. Since the impacts of QBO and ENSO are seasonally modulated could the authors clarify this choice and comment on how sensitive the stratopause height and temperature trends are to this assumption?

We recalculated the OSIRIS trends including additional proxies for the seasonal harmonics of the QBO and ENSO. The figure below shows that this has a very minor effect on the temperature and height trends, so we do not feel that it is necessary to include the seasonal modulation in our study.



Page 8, section 3.3: The model results would benefit from a short comparison with the stratopause height trends reported in previous modelling studies cited in the manuscript, i.e., Pisoft et al. 2021.

This was also mentioned by Reviewer 1. We added a statement to the paper that says “The trends from the CCMI-2022 models are also comparable to the global trends from Pisoft et al. (2021), who found a global decrease in the stratopause height of ~60 m/decade between 1960 and 2080 in the older CCMI-1 models.”

Page 7&8, Figure 3&4: The authors note that the OSIRIS sampling pattern affects the SABER stratopause height trends, particularly near 50°. As differences at other latitudes are relatively small, could the authors provide additional insight into why this latitude appears to be especially sensitive to the sampling choice?

It is because the coverage of OSIRIS depends on the amount of sunlight. The higher latitudes have the fewest number of days when OSIRIS can take measurements. This is now mentioned in the manuscript.

Page 9, Figure 5: In section 2.3 authors states that 11 CCMI-2022 models are used, but Figure 5 shows distributions based on only seven models. Please clarify the discrepancy.

This was a mistake, thank you! We intended to use all 11 available models, but a few of them were missing from the figure. Both figures 5 and 6 have been updated to use all available models. This does not change any of our main results, but it does mean that the CCMI-2022 model trends have a larger spread.

Page 9, line 186: Figure 5 presents a multimodel distribution of stratopause trends but each model contributes only a single free-running ensemble member. As a result, the spread reflects both structural model differences and internal variability based on single realization. One sample only is not representative, especially with the large variability in the tropics. This represents a limitation that should be mentioned by the authors.

Thank you for pointing this out, it is now mentioned in the manuscript.

Page 9, line 197: The authors describe this study as the first comprehensive observational analysis of stratopause height trends, while also noting in the Introduction that some observational studies have previously examined stratopause heights. These statements needs some modification (e.g., emphasizing the latitude-resolved nature of the analysis or the use of the new OSIRIS temperature retrieval) to avoid overstating complete lack of work in this area.

You are correct; this statement has been changed. It now says “We present an analysis of stratopause height trends between 60S and 60N using observations from two satellite instruments, OSIRIS and SABER. This is the first comparison of the inter-annual variability in the stratopause height and stratopause temperature and of the latitudinally-resolved stratopause trends from these two instruments.”