Reply to Reviewer 1

We thank the Reviewer for the careful reading and the constructive criticism of our manuscript. The comments and remarks have helped to improve the paper. We understand two main concerns raised by the Reviewer, regarding (i) the focus on one single reanalysis and (ii) equating pressure and height trends. We have extended our analysis and revised the manuscript accordingly. In the following, we address all comments in further detail (Reviewer's comments in italics, quotations of the corresponding revised text passages in blue).

Based on the suggestions of all reviewers, we have repeated our analyses of LMS structural changes with four modern reanalyses, namely ERA-Interim, MERRA-2, JRA-55 and JRA3Q, in addition to ERA5. The structure of the paper has essentially remained the same, although the individual sections have been expanded to include a comparison of the results between the reanalyses. We have also adapted the title accordingly: "Long-term changes in the thermodynamic structure of the lowermost stratosphere inferred from reanalysis data"

Major comments

(i) Comparison of reanalysis data

I find the authors' decision to focus only on ERA5 without including other reanalyses in any way unfortunate. Although modern reanalyses are improving at a substantial rate, the differences between them, especially in the stratosphere, are well known. For example, for BDC, which the authors themselves mention as a very important factor for trends in LMS, a substantial spread in climatology and trends between reanalyses was found (e.g., Abalos et al. 2015 or Šácha et al., 2024). The inclusion of other reanalyses could make the conclusions substantially stronger in the case of agreement between reanalyses, or reveal discrepancies between them. In this case, where only one reanalysis is used, I recommend a substantial expansion of the discussion regarding the known differences between the most commonly used reanalyses.

We agree that the significance of our study increases when comparing results across different reanalyses, as we are aware of the general strengths and limitations of reanalysis data and the differences between reanalysis data sets. We have therefore repeated all analyses with three widely used reanalyses ERA-Interim, MERRA-2, JRA-55 as well as the recently published JRA3Q data set. For the introduction and illustration of our metrics to assess the structure of the LMS, we use ERA5 before we generalize our findings to all other data sets.

(ii) Pressure and height trends

Vertical shifts of tropopause and potential temperature levels in the paper are evaluated in the pressure coordinate system and the authors report trends in the pressures of these levels. However, the authors compare the detected trends for tropopause with previous studies from Xian and Homeyer (2019), Wilcox et al. (2012a) and Meng et al. (2021) in which trends are evaluated in a geometric vertical coordinate. Also in other parts of the paper (e.g., L215-216), the trends in pressure and altitude are interpreted as equivalent, which I do not consider to be correct as pressure levels in the troposphere shift upwards as a direct thermodynamic consequence of tropospheric warming (see, e.g., Fig. 1 from Eichinger and Šácha, 2020). So I'm concerned that comparing trends in geometric and pressure levels can be quite misleading. For example, if the pressure levels and the tropopause are moving upwards at the same rate, the tropopause pressure would not change, but the geometric height would. I would suggest redoing Figs. 2 and 4 in geometric coordinate, or the authors should show that trends in pressure level shifts are negligible compared to trends in tropopause/isentropic level height.

We acknowledge that a direct comparison between (tropopause or isentropic) pressure and height trends can be misleading due to the mentioned thermodynamic effects. We have therefore made a clearer distinction between pressure and height trends in the manuscript. As suggested, we have added tropopause (geopotential) height trends to Fig 2. Nevertheless, our analysis of pressure and potential temperature at the tropopause and lower stratospheric isentropes was conducted to lead towards an understanding of the LMS mass trends presented. Potential temperature trends are relevant in this context because we define the upper LMS boundary with respect to the potential temperature at the tropical tropopause. The pressure difference between upper and lower LMS boundary determines the LMS mass, whereas the (geometric or geopotential) altitude of theses surfaces plays no direct role in this context. Therefore, we do not discuss differences between pressure and height trends in further detail.

Minor comments

L37: UTLS is not defined

Thank you for the hint. The definition has been added.

L94-95: papers Oberländer-Hayn et al., 2016 or Šácha et al, 2024 showed that tropical upwelling does not intensify at the tropopause, but at pressure levels.

Thank you for the recommend reading. We understand that, according to Oberländer-Hayn et al. (2016) and Šácha et al. (2024), the expected acceleration of the BDC, including increasing tropical upwelling, is in fact to a large extent caused by a lifting of the circulation together with the tropopause (as well as other effects) instead of a pure acceleration of the vertical advection. This is an important information about BDC changes which has been added to the manuscript. Šácha et al. (2024) do report increased upwelling at the tropopause as well as the $100\,\mathrm{hPa}$ and $70\,\mathrm{hPa}$ in all reanalyses except ERA5, as evident from Fig. 3 in their study.

Line 96-98: According to Oberländer-Hayn et al. (2016) and Šácha et al. (2024), it is more precise to speak about a lifting of the circulation, which is connected to the tropopause expansion itself. Stratospheric temperature changes linked to stratospheric ozone additionally influence the BDC evolution.

L109: Do you use ERA5.1 for the 2000-2006?

Thank you for the reference to ERA5.1. The analysis had in fact been performed with ERA5 data so far. We have now updated all ERA5 analyses using ERA5.1 data for the time period 2000-2006. This update has only led to minor alterations of the trend results.

L116: I suggest to change subsection 2.2 name to "Tropopause detection"

„Tropopause detection“ seems indeed to be a more suitable tittle for subsection 2.2. Changed accordingly.

L117: cite ERA5 data

Thank you for the hint. The sentence is now formulated more generally, due to the inclusion of additional reanalyses.

L155: To determine the lapse rate tropopause according to WMO (1957) from all five reanalysis data sets, [...]

L117-215: The description of detection lapse rate tropopause is, in my opinion, insufficient. Although the authors refer to the meteorology (methodology) used in Birner, 2010, the clarity of the paper would benefit from a more detailed description of the meteorology (methodology) and data used.

We agree that the (lapse rate) tropopause detection is a critical factor for UTLS studies. Despite its clear definition, identifying the lapse rate tropopause offers scope for different approaches in practice. We have now provided a more detailed description of our lapse rate detection algorithm, which we hope makes our methodology clearer.

L155-165: To determine the lapse rate tropopause according to WMO (1957) from all five reanalysis data sets, we apply an algorithm closely following that of Birner (2010a), based on the work of Reichler et al. (2003). The algorithm computes the temperature lapse rate on $p^\kappa = p^{R_d/c_p}$ half levels (where p is the pressure, R_d is the specific gas constant and c_p is the heat capacity at constant pressure for dry air), according to equations 1)–4) in Reichler et al. (2003). Subsequently, the algorithm identifies the lowest half level at which the lapse rate becomes smaller than $2\,\mathrm{K/km}$. Following Birner (2010a), a preliminary tropopause is then identified by linear interpolation of p^{κ} and all other variables to a lapse rate of exactly $2\,\mathrm{K/km}$ (stratification threshold). Second, the algorithm checks whether the lapse rate remains on average below $2\,\mathrm{K/km}$ for all higher levels within $2\,\mathrm{km}$ (thickness criterion). Therefore, a temporal level at $2\,\mathrm{km}$ distance to the preliminary tropopause is added and the algorithm successively checks the average lapse rate for all levels between the preliminary tropopause and $2\,\mathrm{km}$ distance. If the preliminary tropopause does not fulfill the thickness criterion, the next higher half level fulfilling the stratification threshold is tested. [...]

L125-126: I find this sentence very vague: The detected lapse rate tropopause agrees well with the ERAInterim lapse rate tropopause in Gettelman et al. (2010) and Wilcox et al. (2012b). If a quantitative comparison was made, why is there no metric given? Moreover, the comparison refers to ERAInterim, a reanalysis from the same family as ERA5. Has validation been performed using other reanalyses? Maybe authors could cite previous studies on this matter: e.g., Hoffmann and Spang, 2022 or Zou et al., 2023.

The cited sentence was not ment to compare ERA5 and ERA-Interim lapse rate tropopauses in this case. The idea was to state that our lapse rate tropopause detection in ERA5 yields meaningful results. We admit that the sentence can be misunderstood and that this qualitative comparison is very vague. We have thus removed the sentence.

L164: Has there been a strong dependence on the mentioned regressors? There is no discussion of the significance of the regressors throughout the paper.

We conducted sensitivity test, running the DLM with and without regressors in order to check the robustness of the trends and their dependence on the regressors. However, our study does not aim to disentangle the contribution of different phenomena, represented by regression proxies, to the presented trends. Rather, we use the regressors for the sake of completeness and consistency with other studies on trends of UTLS characteristics (e.g. Seidel and Randel, 2006; Tegtmeier et al., 2020; Meng et al., 2021; Zou et al., 2023). We agree that a statement about the effect of the regressors was missing in the original manuscript. A short statement has now been added.

L215-220: [...] The same regressors have been used in different studies investigating changes of UTLS characteristics (e.g., Seidel and Randel, 2006; Meng et al., 2021; Tegtmeier et al., 2020; Zou et al., 2023). [...] In order to assess the robustness of the trends and the effect of the regressors, we conducted sensitivity tests in which the DLM was run with and without regressors. Overall, the results showed no strong dependency on the used regressors. However, cold point pressure trends, for example, become more significant when regressors are used (not shown).

L166: SOAD is not defined

Thank you for the hint. There has been a typo, it should be "SAOD". SAOD definition has been added.

L184: potential temperature trend should be 0.7 K/decade

Thank you! Changed.

L325: long-term

Done.

L337-338: the effect of density may be more complex due to changes in both temperature and pressure in the LMS region.

We recognize that we are not able to make a precise statement about density effects at this point. The point we want to make is: The contribution of atmospheric levels to the mass of an atmospheric layer decreases with altitude, assuming an exponential density decrease with altitude. We have changed the sentence accordingly.

L424-425: The "dynamical" upper boundary surfaces based on the tropical tropopause are indeed able to largely compensate for the tropopause rise in ERA5.

References

Meng, L., Liu, J., Tarasick, D. W., Randel, W. J., Steiner, A. K., Wilhelmsen, H., Wang, L., and Haimberger, L.: Continuous rise of the tropopause in the Northern Hemisphere over 1980–2020, Science Advances, 7, eabi8065, https://doi.org/10.1126/sciadv.abi8065, 2021.

Oberländer-Hayn, S., Gerber, E. P., Abalichin, J., Akiyoshi, H., Kerschbaumer, A., Kubin, A., Kunze, M., Langematz, U., Meul, S., Michou, M., Morgenstern, O., and Oman, L. D.: Is the Brewer-Dobson circulation increasing or moving upward?, Geophysical Research Letters, 43, 1772–1779, https://doi.org/10.1002/2015GL067545, 2016.

Seidel, D. J. and Randel, W. J.: Variability and trends in the global tropopause estimated from radiosonde data, Journal of Geophysical Research, 111, D21 101, https://doi.org/10.1029/2006JD007363, 2006.

Tegtmeier, S., Anstey, J., Davis, S., Dragani, R., Harada, Y., Ivanciu, I., Pilch Kedzierski, R., Krüger, K., Legras, B., Long, C., Wang, J. S., Wargan, K., and Wright, J. S.: Temperature and tropopause characteristics from reanalyses data in the tropical tropopause layer, Atmospheric Chemistry and Physics, 20, 753–770, https://doi.org/10.5194/acp-20-753-2020, 2020.

Zou, L., Hoffmann, L., Müller, R., and Spang, R.: Variability and trends of the tropical tropopause derived from a 1980–2021 multi-reanalysis assessment, Frontiers in Earth Science, 11, 1177 502, https://doi.org/10.3389/feart.2023.1177502, 2023.

Šácha, P., Zajíček, R., Kuchař, A., Eichinger, R., Pišoft, P., and Rieder, H. E.: Disentangling the Advective Brewer-Dobson Circulation Change, Geophysical Research Letters, 51, e2023GL105 919, https://doi.org/10.1029/2023GL105919, 2024.