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.

Reply to Reviewer 2

We appreciate the reviewer's interest in our study. However, we would like to clarify some apparent misunderstandings: The LMS mass is calculated based on the formalism of Appenzeller et al. (1996). However, our study does not address mass transport between the tropics and subtropics, nor does it utilize Transformed Eulerian Mean equations. Our study presents the mass of the lowermost stratosphere (LMS) and associated trends. It is correct that we define and compare different boundary surfaces enclosing the LMS.

In the following, we address all comments in detail (Reviewer's comments in italic, 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

Overall, the manuscript reads well and sounds reasonable, and I find it a nice work. However, a concern arises from the beginning, whether the authors have used here ERA5 or ERA5.1. It is now well known that ERA5 has a cold bias in the lowermost stratosphere; therefore, its data are not reliable for some time in the region of the atmosphere that is the focus here. To address this issue, ERA5.1 was produced. The authors may have used the corrected ERA5.1 data; however, they should make it explicit. If they have used ERA5, there is a chance that the results here are not entirely valid (as they are based upon data known to be erroneous), and they should update the study using ERA5.1. On the other hand, if they have used ERA5.1, this should be made explicit in the text.

Thank you for pointing out the cold bias in ERA5, rectified in ERA5.1. We agree that it is important to use ERA5.1 for our study of the lowermost stratosphere. All ERA5 analyses have been revised using ERA5.1 for the affected time period, i.e. 2000-2006. This revision has only led to minor alterations of the trend results.

As some of the authors know (they have approached us in conferences to talk about it), some colleagues and I have proposed similar metrics to the ones here used for the UTLS region since years ago (joint PV and potential temperature changes), to determine the transition from tropics to the extratropical region (where the LMS mass is computed here), and regularly updated them. Although not published in a paper, they have been widely presented in SPARC workshops. Examples are:

- *Añel, J. A., Gettelman, A., Castanheira, J. M., de la Torre, L. (2018) Tropical widening from isentropic and potential vorticity fields, SPARC OCTAV-UTLS Workshop. 7 - 9 November 2018, Mainz (Germany)*
- *Añel, J. A., Gettelman, A., Castanheira, J. M. (2015) Tropical widening from isentropic and PV fields. SPARC Regional Workshop on the Role of the Stratosphere in Climate Variability and Prediction, 12-13 January 2015, Granada (Spain)*
- *Añel, J. A., Gettelman, A., Castanheira, J. M. (2009) Tropical broadening vs. tropopause rising. "The Extratropical UTLS" SPARC Workshop, 19-22 October 2009, Boulder (CO, USA)*

These metrics are based on the equivalent latitude of cross-points between isentropic lines and the PV field, so they are especially relevant for the discussion between lines 70 and 81 and the paragraphs after line 270 (which are the same as we have shown in the past) and Figure 9 (which is precisely the same kind of plot we have been presenting). Essentially, the results we have presented in the conferences have always matched those based on the tropopause break associated with the jet for all the reanalysis (ERA-Interim, NCEP2, JRA-55 and MERRA) and WACCM4 (the CCMI version with 66 vertical levels and in a configuration with 133 levels). It is comparable to the crossing between the 350 K isoline or 2 PVU and the thermal tropopause *used here. For the fairness and completeness of the discussion, it should be cited. The discussion in the Introduction in line 79, paragraphs 280-295, and the conclusions in line 367 are the right parts of the text to attribute the original idea and past results. Actually, I find it quite unfortunate that our works have not been cited in this submitted version. I have attached to this review one of our SPARC presentations to illustrate it.*

We did not mean to withhold any results and acknowledge your ideas and effort on the subject of tropical widening. Unfortunately, the mentioned talks are not easy to reference or to find for the interested reader. We therefore respectfully refrain from citing the mentioned work and instead refer to peer-reviewed literature, also referring to the ACP reference guidelines (`https://www.atmospheric-chemistry-and-physics.net/submission.html#references`).

Additionally, I have read the manuscript several times, and I find a gap (maybe biased by my own scientific interests) in all the exposition and discussion related to the changes in the structure of the UTLS: the changes in the structure of the tropopause itself. I agree with the authors that mass changes are the relevant variable and that they have a criterion to delimit the region where it is computed. Additionally, they mention the overlapping of the tropical tropopause over the extratropical as an issue. However, from the point of view of the lapse-rate definition, it is clear that the region studied here is changing, which is evident in the broader area of vertical stability, fingerprinted by an increase in the numbers of multiple tropopauses (e.g. Castanheira et al. 2009), correlated with increasing UTLS baroclinicity. I think it is relevant to mention this around lines 96 and 225 and, if possible, to add in the manuscript some discussion on how the metrics presented here could be related to the changes in the vertical stability and the widening of the tropopause region.

We agree that the increase of double tropopause frequency should be mentioned in our manuscript, as it can be related to changes in the thermodynamic structure of the UTLS. We have added according information, referring to the study by Castanheira et al. (2009).

L90-92: In addition to the general widening of the tropics, the frequency of double tropopause events, i.e. poleward excursions of the tropical above the extratropical tropopause, is found to have increased (Castanheira et al., 2009; Xian and Homeyer, 2019). This trend likely reflects an increase in baroclinicity in UTLS, driven by the GHG-induced climate change (Castanheira et al., 2009).

Finally, using different periods (since 2000 and 1980) sometimes makes the text confusing. I do not see the point of beginning in 2000 and then extending the analysis back to the 1980s. Does it provide some fundamental new insight here? I doubt it. The authors could think about simply removing the part pre-2000.

We take note of the experience that the discussion of different time periods can be confusing. Nevertheless, the main strength of the DLM (dynamic linear regression model) trend analysis is to infer potential trend reversal dates without prior specification. Even though our focus is on the time period after 1998, the non-linear DLM trends show interesting features also before this time. Especially the identification of trend reversal dates, as well as the specification of continuous trends, are helpful for interpreting the results after 1998. For example, considering both, the full time period 1979-2019 as well as changes after 1998, helps putting our results for the SH lapse rate tropopause pressure in context with trend studies focusing on different time periods.

Minor comments

Line 42: citing Hoerling et al. (1991) about the potential vorticity and the tropopause is right. However, the numbers given by Hoerling et al. limit the location of the tropopause to 1-3 PVU, which has been proven too restrictive. Hoinka et al. (1999) have a good discussion, showing that 3.5 PVU approaches extratropical tropopause better. I recommend adding a citation to Hoinka et al. so that those readers without a profound knowledge of the topic have a more comprehensive and updated view of the issue of using the PV criterion to "find" the tropopause.

In fact, Hoerling et al. (1991) compare tropopause definitions for PV threshold values between 1-5 PVU. Hoinka (1999) choose 3.5 PVU, referring to Hoerling et al. (1991).

Line 43: when discussing the chemical tracers, I think it is fair to add the e90 by Prather et al., e.g. (2011) https://doi.org/10.1029/2010JD014939

We acknowledge the existence of other tracers beside ozone, including e90, that allow for the definition of a chemical tropopause as shown by Prather et al. (2011). However, we aim to give a brief overview rather than a complete list of tropopause definitions in the mentioned text passage.

Lines 88-94: I find this paragraph explicative and well-referenced. The authors mention that the issue of the BDC trends is an ongoing discussion. They refer to models, satellite data, and reanalysis. First, I would clarify in the text that Tegtmeier et al. (2020) refer to reanalysis. Then, I recommend citing a more up-to-date study, recently published by Sacha et al. (2024), which shows consistent results from models and uncertainties from reanalysis (https://doi.org/10.1029/2023GL105919).

We appreciate the reference to the study by Šácha et al. (2024), which we have cited within the mentioned paragraph on BDC trends. Additionally, we have included the study by Zou et al. (2023), pointing out that the direction of tropical tropopause temperatures appears to have changed after 2005. Tegtmeier et al. (2020) report trends in the tropical tropopause layer for the period 1979-2005, comparing reanalysis data and observations.

L96-97: [...] 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.
L105-110: [...] This is consistent with tropical lower stratospheric temperature trends close to zero within this period, inferred by Zou et al. (2023) from reanalyses data. The temperature reduction in the tropical tropopause region and at the cold point reported for the time period 1979–2005 by Tegtmeier et al. (2020) is consistent with increased tropical upwelling. However, different reanalyses often show a significant spread when compared, whether in terms of, e.g., temperature trends in the TTL region (e.g., Tegtmeier et al., 2020) or dynamical tropical upwelling (e.g., Šácha et al., 2024).

I would remove the explanation on the Bayesian basis of the DLM, the paragraph beginning in line 98.

We suppose this comment is related to line 168 of the original version of the manuscript. Its Bayesian nature distinguishes the DLM from other methods for time series analysis, e.g. multiple linear regression. We consider the short explanation on the DLM principles useful, especially for the understanding of the presented trend uncertainties.

Also, the explanation of the accessibility to the DLM model code is already included in the "Code and data availability" section. Regarding this, a minor issue: GitHub is not a suitable repository to store assets from scientific research or papers; GitHub states it on its webpage and offers an integration with Zenodo to store code that needs long-term archival, as the one used in papers, providing a DOI for it. I strongly recommend copying the DLM code in Zenodo and citing it instead of the GitHub repository.

Here, the goal was to show the source of the dlmmc which is available on github and can be referenced by Alsing (2019).

Also, instead of making the LMS code available upon request (which outcome is never assured), I recommend depositing it a permanent repository. ACP does not enforce this, but it is the usual practice in many other journals, including some of the EGU.

We agree that publication of our code and data makes it easier for the scientific community to benefit from our code and data and to compare methods and results. We have therefore made the code for LMS mass calculation and trend estimation available on zenodo, together with the different LMS boundary fields (i.e. lapse rate tropopause, PPT10mean, PPTcp10mean as well as the cold point) for all five reanalyses. See `https://zenodo.org/records/13890232`.

I think the colour scale for the DLM trend state in Fig. 3 should be improved. The discussion focuses on values lower than 7.5 hPa, and this is mostly yellow with independence of the values. It would be good if *the authors could provide a colour palette that helps to perceive the differences between 0 and 7.5 hPa.*

This is a valid point. We have improved the color scale.

Lines 235-236: I would delete this mention of the polar vortex. Overall, the link between ozone recovery and its thermal effect and the material barrier that the polar vortex represents to latitudinal mixing is well-known and clear. However, I do not think it is actually relevant to the discussion here and only introduces some confusion.

We think this is an interesting observation and always good to point out consistencies, also if relationships are well known.

In the Conclusions, I would emphasise the "model" dependence of the results shown here for the LMS changes. There are some disagreements with previous works and probably with another reanalysis if it was checked.

We are aware of the general strengths and limitations of reanalysis data and the differences between reanalysis datasets, specifically relevant for trend analyses. We realize that the manuscript was lacking a clear discussion on the uncertainties arising from this. Therefore, we have extended our analysis, comparing the previously presented results for ERA5 to three widely used reanalyses ERA-Interim, MERRA-2 and JRA-55 as well as the recently published reanalysis JRA3Q. In order to better address uncertainties of our findings, we point out robust features and discuss discrepancies across the different reanalyses.

References

Alsing, J.: dlmmc: Dynamical linear model regression for atmospheric time-series analysis, Journal of Open Source Software, 4, 1157, https://doi.org/10.21105/joss.01157, 2019.

Appenzeller, C., Holton, J. R., and Rosenlof, K. H.: Seasonal variation of mass transport across the tropopause, Journal of Geophysical Research: Atmospheres, 101, 15 071–15 078, https://doi.org/10.1029/96JD00821, 1996.

Castanheira, J. M., Añel, J. A., Marques, C. A. F., Antuña, J. C., Liberato, M. L. R., De La Torre, L., and Gimeno, L.: Increase of upper troposphere/lower stratosphere wave baroclinicity during the second half of the 20th century, Atmospheric Chemistry and Physics, 9, 9143–9153, https://doi.org/10.5194/acp-9-9143-2009, 2009.

Hoerling, M., Schaack, T., and Lenzen, A.: Global objective tropopause analysis, pp. 1816–1831, 1991.

Hoinka, K. P.: Temperature, Humidity, and Wind at the Global Tropopause, Monthly Weather Review, 127, 2248–2265, https://doi.org/10.1175/1520-0493(1999)127<2248:THAWAT>2.0.CO;2, 1999.

Prather, M. J., Zhu, X., Tang, Q., Hsu, J., and Neu, J. L.: An atmospheric chemist in search of the tropopause, Journal of Geophysical Research, 116, D04 306, https://doi.org/10.1029/2010JD014939, 2011.

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.

Reply to Reviewer 3

We thank the Reviewer for the careful reading and constructive criticism of our manuscript. The comments and remarks have helped to improve the paper. We understand the main concerns raised by the Reviewer, regarding the discussion of uncertainties and justification of the choice of data and methodology. We have extended our analysis accordingly. In the following, we address all comments in further detail (reviewers comments in italic, 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

We realize that our manuscript can be improved by a more thorough discussion of uncertainties. In order to investigate the robustness of our findings, we have included a comparison of ERA5 with three widely used reanalyses, namely ERA-Interim, MERRA-2 and 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. Furthermore, we have included a brief statement regarding the impact of the regressor variables on the trends. Beyond the aforementioned dependencies on data and regressors, we are confident that our method ensures a good treatment of uncertainties. One advantage of the DLM over other methods for time series analysis is the rigorous treatment of uncertainties in the data and the regression coefficients by simultaneous estimation of all model components. Further details are provided in the replies to the specific comments.

Specific comments

The figures are clear and support the narrative. However, some figures (e.g., Figures 2 and 4) could benefit from putting individual lines into separate panels. Furthermore, the shading in these figures denotes the associated standard deviation, however, the statistical significance is tested on top of it. I would rather display 95% confidence or credible intervals (Šácha et al., 2024) straight away.

Thank you for the suggestions. We have changed the figures accordingly, now presenting tropopause trends (Fig. 2), isentropic pressure trends (Fig. 5a–c) and pressure trends of the upper LMS boundaries (Fig. 9a–c) for ERA5 in individual panels. In addition to the mean and standard deviation, additional lighter shading now denotes the 5-95 percentile. In order to illustrate the reanalysis comparison without overloading the paper, we present the respective results in single panels. For the sake of clarity, we omit the presentation of uncertainty in these comparison figures but refer to the ERA5 plots as uncertainties are of the same order of magnitude across the reanalyses (Figs. 3, 5d–f and 9d–f).

Explain the selection criteria for the reanalysis and its period since ERA5 goes beyond the year 1979

We are aware of the potential and limitations of reanalysis data in general and of the individual reanalyses in particular, especially regarding long-term trend studies. Therefore, we agree that our study can benefit from a comparison of different reanalyses. While illustrating our metrics for LMS characterization with ERA5, we have included a comparison of the results in ERA-Interim, MERRA-2, JRA-55 and JRA3Q. We limit our trend analysis to the satellite era, starting in 1979. The time period 1979-2019 makes the different reanalyses reasonably comparable (MERRA-2 starting in 1980, 2018 being the last full year of ERA-Interim).

L134-135: For this study, ERA5 monthly mean data for the time period 1979–2019 has been used (Hersbach et al., 2020), 1979 marking the beginning of the satellite era. [...]
L144-146: In addition, we use monthly mean data from ERA-Interim (Dee et al., 2011), MERRA-2 (Gelaro et al., 2017), JRA-55 (Kobayashi et al., 2015) and JRA3Q (Kosaka et al., 2024) for the same time period as ERA5. However, ERA-Interim is only available until 2018 and the MERRA-2 time series begins in 1980.

Why do you use only 2000 samples?

This is a valid question since the number of Markov Chain Monte Carlo samples drawn from the posterior distribution of the DLM analysis is a choice left to the DLM user. In order to chose an appropriate DLM setup, we have performed sensitivity tests, comparing different numbers of DLM samples (e.g., 200, 1000, 2000, 5000, 10000) and their effect on the mean and spread of the DLM background trend state. The number of DLM samples does indeed influence the mean and spread of the resulting trends. A larger number of DLM samples does not necessarily reduce the spread, i.e. the uncertainty of the trend results but the trend with respect to, for example, latitude becomes smoother and thus more reliable. However, we consider the improvement from a number of 2000 to 10000 samples as negligible and not in proportion to the considerably higher computing effort. We therefore chose 2000 samples (plus 1000 warm-up samples). The same choice has also been made by Minganti et al. (2022). Karagodin-Doyennel et al. (2022) present DLM results for a sample number of only 200, while other studies compute 10000 samples (e.g., Laine et al., 2014; Šácha et al., 2024) and even 100000 samples (e.g., Ball et al., 2019).

L227-229: In this study, the DLM runs provide 2000 possible model state estimates after an additional 1000 samples that are considered as warm-up and discarded. Sensitivity experiments conducted as part of this study have shown that increasing the number of DLM samples (to e.g. 10 000) does not significantly improve the results, but comes at a considerable computational cost.

Since the authors use other regressors, I would appreciate discussion of their impact and whether they contribute to reduce the uncertainty of the discussed trends.

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). However, 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).

Using vector figures instead of raster ones may help to improve the quality of your publication.

Thank you for the suggestion. We have changed all figures (except Figs. 3 and 5) to vector figures. (Figs. 3 and 5 become too large and are thus compressed.)

I think the whole community would appreciate an adoption of Open Science approaches to allow the reproducing the extensive analysis in this study (e.g. Laken, 2016), especially when authors use DLMMC which has been made open. In particular, I would recommend any kind of willingness of the authors to publish the code allowing to reproduce the figures in the paper. There are multiple ways how to proceed, either to allow access upon request or via portals allowing to assignment Digital Object Identifier (DOI) to the research outputs, e.g. ZENODO. I think it could enhance the quality and reliability of this publication.

We agree that publication of our code and data makes it easier for the scientific community to benefit from our code and data and to compare methods and results. We have therefore made the code for LMS mass calculation and trend estimation available on zenodo, together with the different LMS boundary fields (i.e. lapse rate tropopause, PPT10mean, PPTcp10mean as well as the cold point) for all five reanalyses.

See https://zenodo.org/records/13890232.

Technical comments

L11 0.5° latitude per decade?

It is 0.5° latitude between 1998-2019.

L166 replace SOAD with SAOD and define

Done.

References

Ball, W. T., Alsing, J., Staehelin, J., Davis, S. M., Froidevaux, L., and Peter, T.: Stratospheric ozone trends for 1985-2018: sensitivity to recent large variability, preprint, Gases/Remote Sensing/Stratosphere/Physics (physical properties and processes), https://doi.org/10.5194/acp-2019-243, 2019.

Karagodin-Doyennel, A., Rozanov, E., Sukhodolov, T., Egorova, T., Sedlacek, J., Ball, W., and Peter, T.: The historical ozone trends simulated with the SOCOLv4 and their comparison with observations and reanalyses, Atmospheric Chemistry and Physics, 22, 15 333–15 350, https://doi.org/10.5194/acp-22-15333-2022, 2022.

Laine, M., Latva-Pukkila, N., and Kyrölä, E.: Analysing time-varying trends in stratospheric ozone time series using the state space approach, Atmospheric Chemistry and Physics, 14, 9707–9725, https://doi.org/10.5194/acp-14-9707-2014, 2014.

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.

Minganti, D., Chabrillat, S., Errera, Q., Prignon, M., Kinnison, D. E., Garcia, R. R., Abalos, M., Alsing, J., Schneider, M., Smale, D., Jones, N., and Mahieu, E.: Evaluation of the N_2O Rate of Change to Understand the Stratospheric Brewer-Dobson Circulation in a Chemistry-Climate Model, Journal of Geophysical Research: Atmospheres, 127, e2021JD036 390, https://doi.org/10.1029/2021JD036390, 2022.

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.