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.