Review of

The extratropical tropopause - Trace gas perspective on tropopause definition choice

by Bauchinger et al.

The manuscript demonstrates the use of various tropopause definitions for the interpretation of aircraft measurements. The study is of interest for researchers working with in-situ or other high-resolution data in the UTLS region. The authors provide a detailed analysis of how the choice of the tropopause definition can impact our understanding of tracer gradients and air mass origin. These findings are further explored and illustrated with a few selected case studies. The manuscript is well written and clearly structured. I recommend publication after the following comments have been addressed.

Major comments:

1. The chemical ozone tropopause definition is based on a climatology that links ozone mixing ratios to a vertical distance above the thermal tropopause. The ozone profiles used to construct this climatology are from one particular station. It is not clear and not sufficiently discussed if this ozone climatology is representative for the regions of the flight campaigns. The manuscript could provide some sensitivity analysis (potentially as a supplementary material) on how climatologies based on multiple station data or satellite data would impact the presented results.

2. The motivation of the use of a chemical tropopause (given in section 3.3) argues that that the same data set can be used for the definition of the tropopause and the trace gas of interest. But this is not the case in this manuscript, or? The motivation also argues that the chemical tropopause avoids mismatches in dynamical parameters. However, with a one-station ozone climatology applied to a large geographical region it seems that these advantages of the chemical tropopause do not apply. Furthermore, it seems that the chemical tropopause is used in a different way (based on a climatology) while the dynamical tropopauses are based on the respective dynamical situation. All of this needs to be clarified and discussed more clearly.

3. To illustrate the ozone based and N2O based tropopause definitions, it might be helpful to show a latitude-altitude cross section of each gas. Or some typical profiles.

4. The use of the cross-tropopause gradient in Figure 7 seems not sufficient to assess the quality of separation into tropospheric and stratospheric air as the authors explain starting in line 374. Instead of using a supplement and whose interpretation needs to be combined with the interpretation of Figure 7, it would seem easier to evaluate the second derivative instead of the first.

5. Some of the data is not freely available as stated in the manuscript. Following the link to the HALO database, MISSION: PHILEAS webpage, I found 'Begin free data access: 2028-09-29'.

Minor comments

Section 2 Data: What is the time period considered? How many flights are included in the analysis?

Line 83: in-fight -> in flight

Line 21 (significant uncertainties in our understanding) and line 25 (knowledge gaps and uncertainties): I do not disagree, but these general statements need to be explained and backed up more. What are the uncertainties and remaining questions? Describe the general knowledge gaps and give some examples illustrating why these are relevant.

Line 98: I don't really understand the argument: '*the routes taken were not tailored to explore special features in the atmosphere*'.

Line 110: Was 'MS data set' defines?

Line 250: This seems not really true north of 40N.

Line 258-263: Can downwelling impact tracer distributions and PV in different ways and therefore impact the differences in seasonality?

Line 269-270: The statement 'most notable … spring' is not clear to me.

Line 286: Does 'with' need to be removed. Something seems wrong in this sentence.

Line 286-291: Wouldn't this impact ozone and thus the chemical ozone tropopause in similar ways?

Line 295 and other places: The term seasonal trends seems to refer to seasonal variations and might be confusing.

Line 308-309: Does this show in the mean or in the standard deviation?

Figure 6: For comparability, it would be better to show the full y-range for the upper panels as well.