We thank the reviewer for his comments. Below are our responses in blue.

The main changes were:

- We added the main conclusion to the abstract
- We modified the symbol size for the ozonesondeds and lidar in figure 1
- We added a rationale for the period and datasets used.
- We added a rationale for the usage of “zonal” means.
- We added a brief discussion of the WMO multiple tropopauses and tropopause breaks
- We added several references as suggested by the reviewers.

Review by Juan Antonio Añel

Review of (egusphere-2024-144) by Millan et al.

In this paper, the authors present the analysis of several relative coordinate systems to define the transition layer between the troposphere and the stratosphere, dealing with the regimes in the upper troposphere, the tropopause itself, and the lowermost stratosphere. They use ozone concentrations as a key fingerprint.

First of all, I have co-authored some works with some of the authors of this manuscript. However, we have not collaborated over the last few years; therefore, I have not perceived a conflict of interest, and I think I can provide an objective review of this paper.

Also, I recommend citing several of my works here. I am not trying to impose their citations on the authors. I suggest them because I think they cover gaps in this case and will help create a more balanced, complete, and informative manuscript for the readers. I let the authors and the editor judge on it.

We have added some of the suggested references even though we believe that the review process is not meant to be a tool for the reviewers to bolster their citations.

My main comment is that the manuscript would benefit from explaining why this topic is relevant and explaining the potential applications of these coordinates.

We believe the introduction provides already details to explain why this topic is important; in particular, just the first sentence succinctly summarizes why it is critical to quantify UTLS composition variations. Extensive discussion of the applications of these coordinates is beyond the scope of this paper but many of the references included in the introduction (e.g., Hoor et al., 2004; Pan et al, 2004; Hegglin et al., 2009; Manney et al, 2011; Schwartz et al., 2015; Olsen et al., 2019) describe such applications and will help any interested reader learn more.

As a general comment, given that equivalent latitude is the best-performing coordinate system here and the resolution is 5 degrees (table 3), I think it is important to note that finer resolutions could improve the result using it. Also, the authors use the known piecewise-constant method to compute the equivalent latitude, which results could be improved up to an additional 5% using a Region of Interest technique (Añel et al. 2013). This could be noted in the Discussion or Summary.

Añel JA, Allen DR, Sáenz G, Gimeno L, de la Torre L (2013) Equivalent Latitude Computation Using Regions of Interest (ROI). PLoS ONE 8(9): e72970. https://doi.org/10.1371/journal.pone.0072970

(https://doi.org/10.1371/journal.pone.0072970).

For measurements with sparse coverage, a finer resolution in the Equivalent latitude bin size may not necessarily improve the results since there will be less measurement to average for a given bin. That said, we have added in the Discussion section the following sentence: Given the importance of equivalent latitude, other methods to calculate it (e.g., Añel et al 2013) could be explored in the future.

Line 24: multiple tropopause and tropopause fold conditions play an essential role, introducing uncertainty on whether the region is under tropospheric or stratospheric conditions. The authors should mention some relevant literature here: Randel et al. 2007, Añel et al. 2008, Wang and Polvani, Añel et al. 2013.

Randel, W. J., D. J. Seidel, and L. L. Pan (2007), Observational characteristics of double tropopauses, J. Geophys. Res., 112, D07309, doi:10.1029/2006JD007904.

Añel et al. (2008) Climatological features of global multiple tropopause events, J. Geophys. Res., 113, D00B08, doi:10.1029/2007JD009697.

Wang, S., and L. M. Polvani (2011), Double tropopause formation in idealized baroclinic life cycles: The key role of an initial tropopause inversion layer, J. Geophys. Res., 116, D05108, doi:10.1029/2010JD015118.

Añel, J.A., de la Torre, L. and Gimeno, L., 2012. On the origin of the air between multiple tropopauses at midlatitudes. T*he Scientific World Journal*, 2012. https://doi.org/10.1100/2012/191028

Dr Añel is correct in pointing out that multiple tropopauses and tropopause fold conditions play a role in driving the ozone variability, we have added some of the suggested references.

Line 25: it would be valuable to add information on the uncertainty by satellite measurements because of vertical resolution. The impact of the vertical resolution is implied in the sentence "Measurements available in this region are spatially and temporally limited ..." The actual vertical resolution information is given in section 2.1.1, which describes the satellite measurements. We feel that it is not necessary to duplicate measurement details in the introduction.

Line 29: I would mention the cold-point tropopause (Gettelman et al., Pan et al. 2018) as it is relevant for water vapour and has been used many times instead of other definitions. Also, I suggest making it more explicit that Potential Vorticity (PV) can be used to distinguish between stratospheric air masses and tropospheric ones and track them, as they have very different values. They could add some examples, e.g. Chen et al. (2013).

A GETTELMAN, P.M. de F FORSTER, A Climatology of the Tropical Tropopause Layer, Journal of the Meteorological Society of Japan. Ser. II, 2002, Volume 80, Issue 4B, Pages 911-924

Pan, L. L., Honomichl, S. B., Bui, T. V., Thornberry, T., Rollins, A., Hintsa, E., & Jensen, E. J. (2018). Lapse rate or cold point: The tropical tropopause identified by in situ trace gas measurements. *Geophysical Research Letters*, 45, 10,756–10,763. https://doi.org/10.1029/2018GL079573

Chen X, Añel JA, Su Z, de la Torre L, Kelder H, van Peet J, et al. (2013) The Deep Atmospheric Boundary Layer and Its Significance to the Stratosphere and Troposphere Exchange over the Tibetan Plateau. PLoS ONE 8(2): e56909. https://doi.org/10.1371/journal.pone.0056909

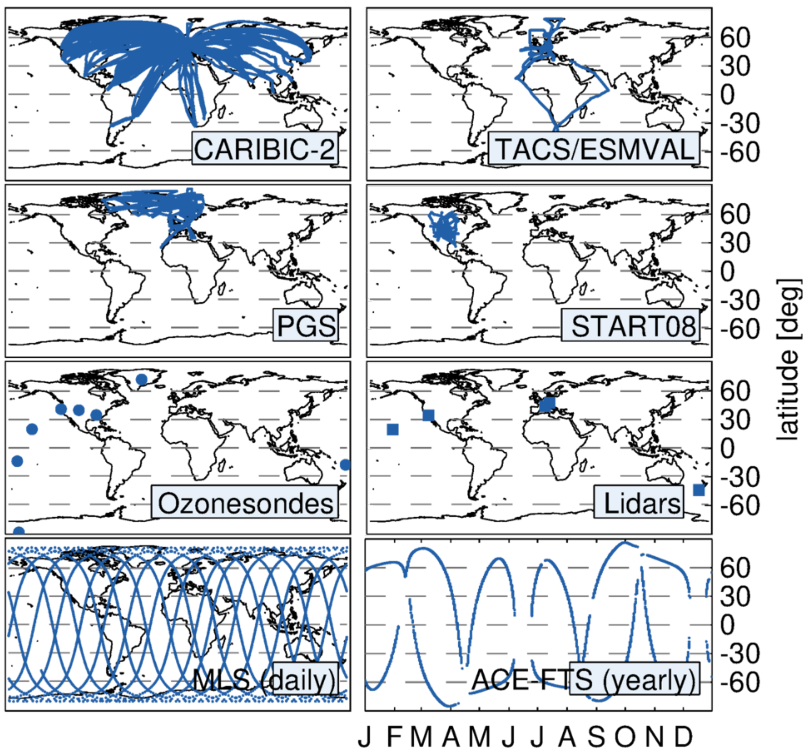
Eventually, OCTAV-UTLS is planning to look at the impact of coordinate systems for water vapor; in those studies, we will certainly include the cold point tropopause. For ozone, the cold-point tropopause is not commonly used.

Line 60: I think it is important to mention that the coordinates and definitions are also relevant and depend on the different phenomena to study (this probably can be addressed in the paragraph I mentioned before on the applications of this work); in many cases, even more critical than "regional" features.

The sentence now reads: Thus, the coordinates that are most helpful to study geophysical and transport properties of the data may be different for different regions and/or phenomena that are of interest.

Fig. 1: The dots and squares to locate the ozonesonde and lidar sites are too big to be informative. It would be good to have them in a smaller size.

Table 1 includes the lat lon info of the ozonesondes and lidars. Nevertheless, the figure was updated as suggested.



Line 90: I understand the reasons for it, but it would be good to add a line with the reasons to use Aura-MLS and ACE-FTS: lengthening the time series, measuring principle, etc.

The sentence was modified: In this study, we focus on two satellite limb sounders, Aura MLS and ACE-FTS, to exploit their long time-series and maximize the overlap with other datasets.

For example, I would move the current lines 108-110 here. We decided not to move that sentence, we want to introduce MLS before discussing how ACE-FTS compare to it.

Also, I have found it quite surprising that Toohey et al. (2013) and Hegglin and Tegtmeier (2017) are not cited in this subsection, as they directly discuss the bias in ozone measurements by the instruments used here, and several of the authors of this manuscript (and myself) are co-authors of both these works.

Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative, J. Geophys. Res. Atmos., 118, 11,847–11,862, doi:10.1002/jgrd.50874.

Hegglin and Tegtmeier (2017) https://doi.org/10.3929/ethz-a-010863911.

We added those citations after the sentence: "*In comparison with MLS, ACE-FTS has much lower sampling density and thus shows a seasonally varying sampling bias*"

Now it reads: In comparison with MLS, ACE-FTS has much lower sampling density and thus shows a seasonally varying sampling bias (Toohey et al., 2013; Millán et al., 2016; Hegglin and Tegtmeier, 2017).

Where Millán et al 2016 is: Millán, L. F., Livesey, N. J., Santee, M. L., Neu, J. L., Manney, G. L., and Fuller, R. A.: Case studies of the impact of orbital sampling on stratospheric trend detection and derivation of tropical vertical velocities: solar occultation vs. limb emission sounding, Atmos. Chem. Phys., 16, 11521–11534, https://doi.org/10.5194/acp-16-11521-2016, 2016.

Line 158: the year for "Smit and Thompson" is missing. Fixed

Line 161: It could seem evident that the 50 hPa region is outside the UTLS regime. Therefore, the results should be fine with the mentioned problem with the ozonesondes. However, it would be good to be clear with numbers about the reasons, mentioning that it is because the cases when the tropopause extends up to 50 hPa and above (double and triple tropopause cases reflecting a transition layer yet) are below the 5% for most of the planet and below 20% in only a few regions (which however coincide for example with the ozonesonde for Boulder). This information can be found in Añel et al. (2008)

We meant that top limit of the plots shown in the manuscript was close to 50 hPa so the drop-offs occur at pressure levels not studied in this manuscript. The sentence was changed to: These drop-offs were typically limited to pressures above ~50 hPa, which is approximately the upper limit of the vertical range used in this study. Therefore, the results shown here should generally be unaffected.

Line 231: VMR has not been defined before. VMR was changed to volume mixing ratio

Line 266. Rather than the studies cited, I would cite the primary studies dealing with the exposition of the multiple tropopause phenomenon (Randel et al. 2007 and Añel et al. 2008).

We have added these citations

Also, I understand that mentioning intrusions here is a generalization, which is not entirely accurate. MTs in this region are not necessarily associated only with intrusions understood in the sense of vertical movement but also with the latitudinal mix and overlapping of the tropical tropopause over the extratropical one and undergo latitudinal advection. This is mentioned later in line 274, but it should be added here and clarified to avoid misinterpretations.

Since this is mentioned in the same paragraph, we don't think we need to clarify it. To add it in line 266 will break the flow and be repetitive.

Line 274: regarding the horizontal mixing, again cite Wang and Polvani (2011) and Añel et et al. (2012). We added the Wang and Polvani reference. Anel et al 2012 is published in, what many consider, a predatory journal (https://en.wikipedia.org/wiki/The_Scientific_World_Journal) and hence we decided not to include such reference.

From Fig. 6, it is clear that the 4.5 PVU value catches better than the 2.0 PVU value in the stratospheric character and does much better in extratropical regions. This is not new at all. Later in the text (in the Discussion), the authors mention that it matches previous findings by Kunz et al. (2011a); however, already a prior work by Hoinka (1998) made clear that values above 3.5 PVU are a better representation of the extratropical tropopause. The result again makes a point against the extended use of the 2 PVU value to define the tropopause, which is clearly an overestimation. This point could be included in the Discussion.

Hoinka, K. P., 1998: Statistics of the Global Tropopause Pressure. Mon. Wea. Rev., 126, 3303–3325, https://doi.org/10.1175/1520-0493(1998)126<3303:SOTGTP>2.0.CO;2.

We added that reference before the Kunz et al. (2011a) citation.

Lines 282-283: please do not use parenthesis this way. https://eos.org/opinions/parentheses-are-are-not-for-references-and-clarification-saving-space

The sentence in question was changed to: Compared to other datasets, MLS displays larger RSTD values in the Northern extratropics and smaller values in the Southern extratropics in the tropopause-based coordinates. Despite its coarse vertical resolution potentially failing to properly resolve the tropopause, this RSTD values might be related to its better coverage of the region, i.e., MLS might sample more variability.

Both the datasets and code for the analysis should be better deposited in long-term repositories with DOI (e.g., PANGAEA, Zenodo). I know it is not a journal requirement, but it is good practice for the assets that the authors can do with reasonable effort.

The ozone datasets are available elsewhere as specified in the data availability section. The dynamical diagnostics are only available upon demand. That said, we are currently working on translating the IDL JETPAC software into python so that it is truly helpful for the community, i.e., without the expensive license for IDL, and that software will be made public.