

Response to Referee

We thank the reviewer for your time and effort in providing such insightful and constructive feedback.

We have carefully addressed all the concerns raised. The suggestions have been invaluable in improving the clarity and rigor of our manuscript. Below is our point-by-point response to the comments.

Specific comments

1. Though not stated anywhere in the manuscript, it is implied from Figure 2 that ERA5 output used are those with a coarsened vertical grid (37 pressure levels?) rather than the native high-resolution vertical grid. Using the coarsened data is unfair to the model for these evaluations and likely introduces greater uncertainty and bias to the results than that intrinsic to the reanalysis. The full-resolution output should be used.

Response: We sincerely thank the reviewer for raising this critical technical point. Figure 2 has been regenerated using the 137 hybrid sigma-pressure levels data. We wish to specifically clarify that the core ERA5 data used for the comparison with radiosondes in our study were always sourced from the Hoffmann and Spang (2022) ERA5-based product, which is itself built upon the full-resolution 137 model levels. The previous use of the 37-level data was limited to the schematic demonstration in Figure 2 and has now been corrected.

2. Are the radiosonde data used for comparison independent from the ERA5 assimilation? Given the source, I would expect them to all have been assimilated and thus find the profile differences in Figure 2 to be quite shocking! I'm not sure I've seen such disagreement before, which makes me question whether the colocation in space and/or time is correct. Far more attention and discussion should be given to these issues and their implications. While it is not uncommon to find large differences in tropopause height, large differences in the temperature profiles

(beyond isolated layers here and there) almost never occur.

Response: We thank the reviewer for this insightful and critical comment.

About the assimilation, the radiosonde data integrated into ERA5 are based on standard pressure levels with lower resolution, and ERA5 does utilize a downsampled version of the high resolution radiosonde observations (Ingleby, 2017).

Although high-vertical-resolution radiosonde data are part of the assimilation process in established reanalysis data products, it's still provide a good opportunity to quantify uncertainties in the lapse rate tropopause determination from reanalysis data (Hoffmann and Spang, 2022).

We have corrected the collocation errors, and the temperature profiles now show high consistency. Despite this, a statistically significant difference in tropopause height remains. This indicates that the discrepancy is not due to gross temperature biases but may stem from the challenges of reanalysis in capturing the tropopause. The role of the tropopause determination algorithm should also be considered.

Ingleby, B.: An assessment of different radiosonde types 2015/2016, Technical memorandum, <https://www.ecmwf.int/en/elibrary/80268-assessment-different-radiosonde-types-20152016>, 2017.

Hoffmann, L., and Spang, R.: An assessment of tropopause characteristics of the ERA5 and ERA–Interim meteorological reanalyses, *J. Atmos. Chem. Phys.*, 22, 4019–4046, <https://doi.org/10.5194/acp-22-4019-2022>, 2022.

3. Furthermore on the radiosonde record, I am surprised by the minimum spatial coverage of the data used, especially since many prior studies utilizing such have had far greater global coverage. And finally, it is not clearly communicated what the vertical resolution of the native radiosonde data is. For example, I don't know what lines 136-137 mean. Is the resolution 5-10 m? If so, what is the point of using cubic interpolation to 10 m as such would be unnecessary?

Response: Thank you for these insightful questions. Regarding the spatial coverage, we

acknowledge that our dataset is more limited than some prior studies. This is because our work specifically requires long-term records with high vertical resolution, which are not widely available across the global radiosonde network. As for the vertical resolution, the native data varies by station, typically between 5-10 meters. We applied cubic interpolation to a standardized 10-meter grid not to invent new data, but to create smooth, consistent profiles. This process is essential for our subsequent analysis of the fine-scale structure within the atmospheric boundary layer, as it reduces the potential for aliasing artifacts in our results.

4. Beyond the simple tasks of measuring absolute differences, biases, and RMSE between the radiosondes and reanalyses, several additional points in the paper which are implied to be novel have been demonstrated in prior studies. This is true, for example, for the instances of large tropopause height differences between the model and radiosondes discussed in lines 178-181.

Response: We thank the reviewer for this important comment, which helps us better position the contribution of our work. We agree that the existence of tropopause height discrepancies between reanalyses and observations has been documented in prior literature.

However, using the state-of-the-art ERA5 reanalysis and a high-resolution radiosonde dataset, we provide a systematic assessment and attribution analysis of the spatial patterns, seasonal cycle, and long-term trends of these differences on a global scale. As elaborated in our response to Comment #2, our analysis specifically emphasizes that resolving the tropopause and inversion layers within fine-scale thermal structures at high-vertical-resolution is crucial.

5. There is comparatively greater emphasis given to changes in some of the statistics over time, with an implied trend suggesting that further analysis is warranted (e.g., lines 191-194), but such variability seems easily explained by the large differences in sample sizes. Also, the seasonality of the tropopause is discussed in half a

paragraph, but nothing new is provided there.

Response: We thank the reviewer for this comment. We rewrite those sentences and add some new opinion to explain the change.

Line 221: “Table 1 details the statistical differences for each year from 2000 to 2023, revealing a gradual increase in observation data over time. Our intercomparison statistics, detailed in Table 1, reveal that the ERA5-radiosonde TH differences are not homogeneous over time. Instead, the record is marked by two transition points linked to major updates in the ERA5 system...”

6. The statistical significance testing and determination is not well explained. For example, it is not clear to the reader whether the shaded regions are significant or insignificant in Figure 6.

Response: We thank the reviewer for pointing out this lack of clarity. We have modified Figure 6 and its caption as follows to ensure the reader can immediately understand which regions exhibit statistically significant trends.

7. Lines 63-65 imply that COSMIC is the only source of RO data, but that is not true.

Response: We thank the reviewer for this insightful comment. Amended as suggested.

Line 76: “GNSS-based datasets (e.g., COSMIC and the Radio Occultation Meteorology Satellite Application Facility (ROM SAF)) provide high-density measurements with near-global coverage, making them particularly suitable for analyzing global-scale tropopause characteristics (Son et al., 2011)...”

8. 5 & 10 year markers on Figure 1 are indistinguishable

Response: We agree with the reviewer and thank you for pointing out the poor distinguishability of the markers in the original Figure 1. We have redesigned the legend/markers (using triangles for 10-year and circles for 5-year data points).

9. The description of Figure 7 doesn't make sense. What is meant by annual average

variation? These are just time series and diagnosed trend lines, correct?

Response: We thank the reviewer for this comment and agree that the previous description of Figure 7 was unclear and potentially misleading. We have removed the trend lines from the figure, as their presence did not align with the figure's primary intended purpose.

We have refocused the figure to clearly convey its two main objectives:

- i. To demonstrate the strong consistency in the interannual variability of the data between ERA5 and radiosonde observations.
- ii. To visually present the effective recording periods and data coverage of the radiosonde data across different latitude bands.