

Response to Referee

We sincerely appreciate your thoughtful and exhaustive comments and suggestions, which significantly help us to improve the quality of the manuscript.

However, I made a big mistake in my understanding of the tropopause as defined by the World Meteorological Organization, which made my algorithm wrong and on which a large part of the whole article is based. So, we will adjust the whole article by removing the part of modification to the definition of WMO and adding the long term trend analysis of the tropopause.

We would like to express our sincere gratitude to the referee for your exceptionally informative, constructive, and detailed comments, and we would like to answer some of your questions if my algorithm is correct. We would appreciate it if you could give us your valuable comments for the revised manuscript.

General comment

This new study by Gou et al. explores the challenges in identifying the tropopause height using high-resolution radiosonde data based on the World Meteorological Organization (WMO) definition. The study highlights that the original WMO approach tends to underestimate the tropopause height in high-resolution datasets due to the presence of thin temperature inversions and gradient discontinuities. To address this issue, the authors propose two alternative methods, the Moving Average (MV) method and the Coarse-Fine (C-F) method, both of which provide more consistent tropopause height estimates compared to ERA5 reanalysis data. It is found that ERA5 systematically overestimates the tropopause height, particularly near the Hadley circulation edges, while the WMO-defined method underestimates it. The C-F method emerges as the most effective in preserving fine-scale structures while filtering out spurious lower-altitude tropopauses. The study concludes that modifications to the WMO method are necessary when using high-resolution radiosonde data for accurate long-term tropopause trend analysis.

The study addresses a crucial problem with practical implications for climate research and meteorology. The research employs robust statistical techniques, sensitivity analyses, and cross-validation with ERA5, ensuring a high level of scientific rigor. The methodology is well-described, but a deeper discussion on the physical basis of biases and additional validation with independent datasets could enhance transparency. The manuscript is mostly clear and concise. I would like to recommend that the paper be considered for publication, subject to the minor comments listed below.

Specific comments

1. 140: It is a little unclear to me what you mean by 'constant emitted temperature'?

Response: The term 'constant' here refers the radiative equilibrium temperature corresponding to infrared emission escaping to space through the atmospheric. This equilibrium temperature must equilibrate with Earth's absorbed solar radiation; deviation from this balance would drive persistent planetary warming or cooling.

Increased atmospheric water vapor enhances optical depth, elevating the effective emission height to colder atmospheric levels. To preserve radiative equilibrium (i.e., maintain constant emission temperature), the system likely modulates tropopause altitude, positioning the emission layer at an elevation that simultaneously compensates for altered optical depth and preserves radiative equilibrium temperature.

2. 1106: The dynamical tropopause in the tropics is usually defined by a potential temperature threshold, not by potential vorticity. Please clarify.

Response: This correction provided has been noted and is much appreciated. We want to show here that the dynamical tropopause is not available in that region.

This part has been rewritten:

Line 51: "...the dynamic tropopause (WMO, 1985; Hoinka, 1998), typically defined by potential vorticity (PV) thresholds of 1.5–4 potential vorticity unit (PVU) (Turhal et al., 2024), is less reliable in regions of low absolute potential vorticity, such as the tropics, and sometimes in mid-latitudes where strong anticyclonic flow prevails (Hoerling et al.,

1991)”

3. 1116: Lapse rate is defined as temperature difference over height difference. However, for radiosondes this is probably calculated via pressure differences? Could you please clarify and elaborate?

Response: The radiosonde data employed in this study, obtained from the University of Wyoming, ECMWF, NOAA, contain independent height variables.

4. 1136: Please clarify the shape of the moving window. Presumably a boxcar/tophat function was used?

Response: It's a boxcar function. It was a sliding average, which has now been removed from revised manuscript.

5. 1119: Reference paper for ERA5 should be cited: Hersbach H, Bell B, Berrisford P, et al. The ERA5 global reanalysis. Q J R Meteorol Soc. 2020; 146: 1999-2049. <https://doi.org/10.1002/qj.3803>

Response: Amended as suggested.

6. 1195: It would be good to know if the high resolution radiosonde data (or a downsampled version of the data) have been assimilated into ERA5? Presumably the data are not independent?

Response: 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).

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.

7. Fig. 6: These maps suggest that the statistical differences are due to a few individual outliers, probably related to inversions, rather than a bias across all profiles.

Response: The data used for this graph is wrong, but you can see the same type of comparison graph in our revised manuscript, where the prominence of individual cases is allowed in the statistics, and it will be diluted by a large amount of overall error, even though in the new manuscript their differences are small.

Technical corrections

Response: We sincerely appreciate your advice, amended as suggested.