

Response to Referee

We are grateful for the reviewer's valuable time and thoughtful comments. We have diligently addressed all the raised concerns, and our detailed responses are presented as follows.

General comments

1. In both the revised manuscript and the author response, the authors attribute a noticeable change in tropopause heights around the year 2016 to the introduction of a new IFS model cycle (Cy41r2) in ERA5. However, this explanation is not correct. The ERA5 reanalysis was produced using a single, consistent model cycle (IFS Cy41r2) and associated data assimilation system throughout the entire 1940–present record. It is a fundamental feature of reanalysis datasets that they are generated with one frozen model and assimilation system to ensure temporal consistency. Therefore, the observed changes around 2016 cannot be attributed to a switch in model cycle within ERA5. This point should be corrected in both the text and interpretation, and alternative explanations for the 2016 shift should be discussed.

Response: We sincerely thank the reviewer for catching this critical error in our interpretation. We apologize for the misunderstanding. We have now completely removed the incorrect attribution of the 2016 shift to a change in the ERA5 model cycle from both the manuscript and the discussion. Instead, we have revised the text to discuss other potential drivers, such as the strong 2015–2016 El Niño event and its potential impact on global tropopause characteristics, while acknowledging that the exact cause of this specific shift remains an open question and warrants further investigation.

Line 235: “A secondary shift occurred around 2016, marked by an increase in the mean difference and a decrease in the mean absolute difference. While this signal could be associated with climatic events like the 2015–2016 El Niño event, the mechanisms

behind it are not yet clear and warrant future study..."

2. The discussion of the apparent change in tropopause characteristics around 2006 should be refined: The ERA5 data assimilation system experienced a known issue during 2000–2006, leading to a cold bias in the lower stratosphere; this was corrected in the dedicated ERA5.1 reprocessing, which replaced ERA5 for that period. The Hoffmann and Spang (2022) ERA5 tropopause height dataset used in this study indeed employs ERA5.1 for 2000–2006 and ERA5 for the remaining years. Consequently, the discontinuity found around 2006 possibly reflects the transition between these two reanalysis versions. However, the launch of COSMIC GNSS-RO satellites in 2006 introduced a massive amount of high-precision temperature profile observations into the assimilation system, providing stronger observational constraints in the upper troposphere–lower stratosphere region. This increased data availability possibly contributed to the change in mean tropopause height and variability identified in the present study. I suggest discussing this aspect.

Response: We thank the reviewer for this excellent and insightful suggestion. We fully agree that clarifying the dual causes of the 2006 discontinuity is crucial. As suggested, we have now refined the discussion in the manuscript.

Line 224: "Instead, the record is marked by two transition points. The most pronounced shift occurs around 2006, characterized by a decrease in the mean difference and an increase in the mean absolute difference. This discontinuity is likely attributable to two concurrent events. First, ERA5 exhibits a significant cold bias in its stratospheric temperature analysis for the period 2000–2006. To address this, the ERA5.1 reanalysis was produced, which applied the background error covariance from the 1979–1999 period (Simmons et al, 2020). Thus, the discontinuity around 2006 likely stems from the transition from the ERA5.1 to the ERA5 reanalysis within the dataset. Secondly, the launch of COSMIC GNSS-RO satellites in 2006 markedly increased the availability of high-precision GNSS-RO data, strengthening the observational constraints on the upper

troposphere and lower stratosphere. Together, these developments may have caused the observed changes in mean tropopause height and variability...”

Specific comments

1. lines 85–95: The authors list “COSMIC and ROM-SAF” together as examples of GNSS-RO satellite datasets. This is not strictly correct: COSMIC (and COSMIC-2) are satellite missions providing GNSS-RO measurements, whereas ROM-SAF is the EUMETSAT processing facility that generates and distributes GNSS-RO products from several missions, notably the MetOp-A/B/C GRAS instruments. The authors should revise this sentence to clearly distinguish between satellite missions (e.g., COSMIC/FORMOSAT-3, COSMIC-2, MetOp-A/B/C GRAS) and processing centres (e.g., ROM-SAF) to ensure technical accuracy.

Response: We sincerely thank the reviewer for this precise and constructive comment. We agree that distinguishing between the satellite missions and the data processing centers is crucial for technical accuracy. We have revised the sentence in the manuscript accordingly to clearly reflect this distinction.

Line 76: “GNSS-based datasets from various satellite missions (e.g., COSMIC, COSMIC-2, MetOp-A, MetOp-B, and MetOp-C) provide high-density measurements with near-global coverage. Data products from these missions, including those from the MetOp-A (MetOp-B, MetOp-C) GRAS instruments as processed and distributed by the Radio Occultation Meteorology Satellite Application Facility (ROM SAF), are particularly suitable for analyzing global-scale tropopause characteristics (Son et al., 2011)...”

2. lines 135–137: The authors describe resampling the high-resolution radiosonde profiles to a uniform 10 m grid using cubic spline interpolation. This is acceptable given the similar native resolution (5–10 m) and should not introduce significant artifacts at this scale. While linear interpolation would likely be sufficient here, cubic spline interpolation does no harm and seems appropriate for producing

smoothly gridded profiles for subsequent analysis.

Response: We thank the reviewer for this positive feedback and for acknowledging the appropriateness of our methodology. We also consider linear interpolation sufficient, but we deliberately chose cubic spline interpolation to generate smoother vertical profiles, which aids the detailed structural analysis conducted in this study.

3. line 304: The manuscript still refers to “validation” of ERA5 against high-resolution radiosonde data. Since radiosonde observations are assimilated into ERA5, these datasets are not independent, and the comparison cannot be regarded as a true validation. Please replace “validation” with “evaluation” (or “intercomparison”) at this location, and ensure consistent use of terminology throughout the manuscript to avoid implying that ERA5 is being independently validated by the assimilated observations.

Response: We thank the reviewer for this important comment. As suggested, we have replaced “validation” with “intercomparison” throughout the manuscript to accurately reflect that the datasets are not independent and to maintain terminological consistency. We also made several additional modifications to avoid any implication of “validation”.

Line 288: “However, systematic intercomparisons of their biases and spatiotemporal patterns against high-resolution radiosonde data have yet to be fully conducted, particularly for newer reanalyses such as ERA5...”

Line 308: “A comprehensive intercomparison with high-resolution radiosonde observations demonstrates ERA5's exceptional performance in capturing TH characteristics, including absolute values, temporal variations, and spatial correlations...”

Line 327: “But many regions lack high-resolution, continuous radiosonde observations for intercomparison...”