Response to interactive comments on "The tropical tropopause layer in reanalysis data sets" by Tegtmeier et al.

We thank the reviewer for his/her comments which have helped us to improve the paper in revision. Comments are reproduced below, followed by our responses in *italics*.

Anonymous Referee #2

This paper evaluates the vertical structure of the temperature fields from a number of meteorological reanalyses in the tropical tropopause layer (TTL). While the evaluation of reanalyses in this region is important for the user community and fits the focus of ACP, I found several limitations that should be addressed before publishing in ACP.

General comments:
1. From the title, I would expect that the paper also discusses wind or humidity fields in the TTL from the reanalyses, which is not the case. The title should thus be changed and I suggest "Vertical structure of temperature fields from atmospheric reanalyses in the tropical tropopause layer". Or maybe you may have a better suggestion.

> *We agree that the title was too broad and have changed it to 'Temperature and tropopause characteristics from atmospheric reanalyses in the tropical tropopause layer'.*

2. I understood (Sect. 2.2) that reanalysis temperature fields in the TTL are constrained by satellite radiance observations (from 1978 onward), radiosonde profiles (from 1978 onward) and GNSS-RO (between 2002-2006 onward depending on the reanalysis). On the other hand, reanalysis temperature fields are also evaluated by radiosondes and GNSS-RO data. A proper evaluation should be done with independent datasets (i.e. not assimilated) which seems not to be the case. Please clarify and/or comment.

> *We agree that ideally an evaluation would be based on independent data sets. Unfortunately, there is no independent temperature data set with the required spatial coverage, uniform sampling and vertical resolution available in the TTL region.*

3. I found that the intercomparison method lack of details and/or clarity. GNSS-RO data used for the validation of the temperature are provided as zonal mean (P5L27). Is it on a daily or a monthly basis? It is also said that GNSS-RO are interpolated at the reanalysis levels (P5L35-37). A proper comparison of the reanalysis with the observations should be done by (1) mapping the reanalyses at the observation geolocation (by using additional information like averaging kernels or weighting function if necessary) to avoid sampling errors and then (2) calculating the cold point and lapse rate tropopause from the reanalyses in the space of the observations to which they are compared. If done differently, it should be justified.

> *The intercomparison of GNSS-RO data to reanalyses model level temperature (e.g., Figure 4 and 5) is based on the following method. For each individual profile the temperature is interpolated from the two adjacent levels to the reanalyses model level based on the barometric formula. In a second step, the monthly mean tropical mean values are calculated.*
> *The intercomparison of GNSS-RO data to reanalyses cold point and lapse rate tropopause (e.g., Figure 3, 5 and 6) is based on the following method. For each profile, the cold point*

and lapse rate tropopause characteristics were identified based on the cold point and WMO criteria, respectively. In a second step, the monthly mean zonal mean and monthly mean tropical mean values are calculated.

Zonal averages of GNSS-RO data do not suffer from uneven sampling patterns as they are evenly distributed over longitude on a monthly basis (see Fig. 3 of Yu, K., Rizos, C., Burrage, D. et al., An overview of GNSS remote sensing, EURASIP J. Adv. Signal Process. (2014) 134. https://doi.org/10.1186/1687-6180-2014-134).

We have added information to section 2.1 of the manuscript to explain the methodology in more detail.

Please, comment and/or clarify.

4. It is said that GNSS-RO and radiosonde data are provided at high vertical resolution but their values are not given in the manuscript. Please, provide the vertical resolution of these two datasets.

We have added the following information to the manuscript 'The GNSS-RO 'wetPrf' temperature profiles from CDAAC are provided on a 100-m vertical grid from the surface to 40 km altitude. The effective physical resolution is variable, ranging from ~1 km in regions of constant stratification down to 100-200m where the biggest stratification gradients occur e.g. at the top of the boundary layer or at a very sharp tropopause (Kursinski et al., 1997; Gorbunov et al., 2004), most often being somewhere in between.'

Regarding the vertical resolution of radiosondes, in addition to mandatory levels (which near the tropical tropopause are 150, 100, 70, and 50 hPa), individual radiosonde soundings include data at "significant levels," where the observations between mandatory reporting levels depart from a linear interpolation, such as would occur at the tropopause. As the number of significant levels can vary over time and with station, a conclusive statement on the vertical resolution is not possible. We have therefore removed 'high-resolution' from the sentence.

5. Section 3 discusses the reanalyses between 2002 and 2010. Except that GNSS-RO data are not available before that time, is there other reasons to not show the results at earlier time? If not, I recommend providing similar figures (without GNSS-RO data) than Fig. 6 and 9 for, e.g. 1980-1990 and 1990-2000, in a supplement. This would be very instructive for the users of the reanalyses.

We have added the zonal mean evaluations of the lapse rate and cold point tropopause for the time periods 1981-1990 and 1991-2000 in a supplement.

6. There is a long discussion about the use of model- or pressure-levels which is confusing because it seems obvious that using a low-resolution standard pressure grid (only four levels in the TTL) would introduce biases. Fig. 3 is also confusing. I understand that values at 70 and 100 hPa are from the standard pressure but that the CP and LP values are calculated from the model levels. I guess that showing the temperature bias at 70 and 100 hPa from the difference profiles of Fig. 4 would provide (after interpolation) much accurate values. I would suggest to move all the discussion related to the standard pressure levels in a supplement or an appendix and to show in the main body of the paper only results obtained on the model levels.

We agree with the reviewer, that it is not surprising that the low-resolution standard levels introduce biases when used for tropopause calculations. Our sensitivity test is used to illustrate how large such a bias can be for the tropopause temperature, altitude and

pressure calculations. We have simplified and shortened the discussion of this issue, to make this point clearer.

Figure 3 uses the 70 and 100 hPa levels to present the comparison for all reanalyses at the same level. This temperature comparison on pressure levels offers additional information to the comparison on model levels presented in Figure 4. This additional information is valuable for studies that have or will use pressure levels instead of model levels in the TTL regions.

7. Both notations MERRA2 and MERRA-2 are used throughout the paper. Please, choose one of them.

We have changed the notations to MERRA-2.

Technical corrections:

P2 L20-22: "Model simulations. . ." This is not shown in the paper so it should be removed from the abstract.

We have removed the sentence from the abstract.

P3 L11-14: "As the TTL. . ." Please add references at the end of the sentence.

We have added three references to the sentence.

P5 L27: "We use zonal mean. . ." On which time basis? Daily? Monthly? Other?

We have added the information 'monthly mean' to the sentence.

P6 L15: What do you mean by "full input"?

We use the term "full-input" reanalyses here for systems that assimilate surface and upper-air conventional and satellite data (compared to systems that only assimilate surface observations). We have reformulated the sentence to make this clearer.

P6 L24-25: "MERRA-2 . . ." The meaning of this sentence is not clear. Please, clarify.

We have moved the sentence to the acknowledgements.

P6 L26: I would replace "produced" by "constrain" which is more accurate.

We have changed the wording as suggested.

P6 L38-40: "Radiance biases. . ." I don't understand what message the authors want to give with this sentence. Please, clarify.

We have replace the sentence with '… Because radiance biases associated with instrument changes, inaccurate calibration offsets, orbital drifts or long-term CO_2 changes can cause unwanted biases in the resulting reanalysis temperature fields (e.g. Rienecker et al., 2011), a variational bias correction scheme is used during the data assimilation procedure to remove or minimize any radiance biases. This ensures that any temperature changes introduced by the circumstances outlined above are kept small, which is important when looking for long term changes. …'.

P6 L41: ". . .from radiosondes which. . ." Are these radiosonde data the same than those used for the evaluation? See also the general comment related to this issue.

To a large degree the assimilated radiosonde data profiles are the same as the ones used for the homogenized radiosonde data sets. As each radiosonde data set uses different

criteria on which stations and profiles to include, there exist small differences between assimilated and homogenized radiosonde data sets.

Also not the sentence later in the paragraph 'In order to avoid discontinuities or inconsistencies in temperature time series from radiosondes, several reanalysis systems use homogenized temperature data sets such as RAOBCORE (ERA-Interim, JRA-55, MERRA, MERRA-2) and RICH (ERA5).'.

P7 L10: ". . .from GNSS-RO instruments. . ." Same comment as above.

The GNSS-RO instruments assimilated by the reanalyses are the same used for the evaluation of the data sets. We have added the sentence ' ... In addition to the GNSS-RO data sets discussed in section 2.1, C/NOFS-CORISS (Communications/Navigation Outage Forecasting System Occultation Receiver for Ionospheric Sensing and Specification) is assimilated by some of the reanalyses.' *to make this clearer.*

P7 L26: "While the reanalyses assimilate versions of these data..." Do you mean "different versions of these data. . ."?

We have added 'different' to the sentence.

P7 L27: Replace "exactly" by "within their uncertainty" which is more accurate.

We have changed the sentence accordingly.

P7 L30-P8 L1: "In general, the. . ." This sentence does not describe data assimilation methodology. Instead, I suggest "Data assimilation systems combines the information from a model, a set of observations and a priori information weighted by their uncertainties."

We have changed the sentence accordingly.

P7 L12: I don't see the "Section 3.1" in the paper.

We have changed the text to 'Section 2.3'.

P9 L21: Please, add a reference to the "bootstrap method".

We have added the reference: Efron, B., and R. J. Tibshirani (1993), An Introduction to the Bootstrap, 436 pp., Chapman and Hall, New York.

P9 L27-28: "The trend error..." I don't understand the meaning of this sentence. Please, clarify.

We have changed the sentence to '... The uncertainty in each long-term trend is calculated as the standard error of the slope with the effective sample size adjusted to account for the corresponding lag-1 autocorrelation coefficient. ...'

P10 L19-21: "At 100 hPa, ERA-Interim is..." I suggest redoing the figure by using different symbols (star, cross, *) allowing to see the values of all reanalyses.

We have produced different versions of this figure (including different symbols or symbols slightly shifted vertically), but found that the visibility does not improve sufficiently. Therefore, we prefer to keep the figure in its current version and to mention the overlaps in the caption.

P10 L22: Remove "resolution" in ". . .native model level resolution. . ."

We have changed the text accordingly.

P14 L5: I would replace "...over the Maritime continent..." by "...over the sea..." because a continent is one of the several large landmasses that make up the Earth.

As the expression 'maritime continent' has been used in many TTL publications to refer to the overall region including landmasses and sea, we prefer to keep the expression. We will follow the example from Fueglistaler et al. (2009) and use maritime continent in quotes at the first occurrence.

P14 Figure 7: I would be very interesting to also show the results of ERA5. Is there any reason to not show it?

We have added the latitude–longitude comparison of cold point temperature for ERA5 to Figure 7.

P15 L13: Replace "to estimating" by "to estimate".

We have changed the text accordingly.

P16 L4-5: What do you mean by "variability" in ". . .considerable zonal variability. . ."?

We have changed the sentence to ' … The altitude of the lapse rate tropopause shows considerable meridional variability, ranging from 14.5 km to 16.7 km. …'.

P16 Figure 9: Add "pressure" in the upper right panel of the figure, as in Figure 6.

We have added the label 'pressure' to the panel.

P17 L17: "decrease" would be more appropriate than "improve".

We have changed the sentence accordingly.

P17 L29-31: "The influence of ENSO. . ." I do not see any figure showing the influence of ENSO on the TTL temperature. Please, clarify.

As we focus here on the zonal mean interannual variability, we do not show the longitudinal temperature variations associated with the ENSO signal. We have moved the phrase 'not shown here' from the next sentence to this sentence, to make this clear from the onset.

P17 L30: As explained above, change "Maritime Continent" by "sea" or "ocean".

As the expression 'maritime continent' has been used in many TTL publications to refer to the overall region including landmasses and sea, we prefer to keep the expression. We will follow the example from Fueglistaler et al. (2009) and use maritime continent in quotes at the first occurrence.

P17 L37-P19 L5: This part is not very clear because it is never clear to which figure (10 or 11) the text refers. Please, clarify.

The text refers to Figure 10, except for the last sentence. We have added this information.

P17 Figure 10: Why not starting the time series in 1978 or 1980.

For consistency with the S-RIP report and other publications, we use here the S-RIP climatological core time period January 1981 to December 2010.

P21 L14-15: "...all provide realistic..." It should specify that the period of validity of this result is 2002-2010.

We have changed the sentence accordingly.