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 #1

General:
This a very important and well-written paper. To understand long-term changes in the stratosphere, the tropical tropopause layer (TTL) is the most crucial region. Meteorological reanalyses are best estimates of the true state of the whole atmosphere in the past. As such, they are widely used to examine the atmospheric processes and to detect changes in the climate system. This paper gives important insights into the representation of the TTL in all relevant reanalysis products. Thus, I would like to recommend this paper for publishing in ACP with only few minor comments and some remarks.

General:
I think, this is a very important statement that all reanalyses temperatures at the cold point tropopause (or at the lapse rate tropopause) show warm bias if compared to the observations because of the vertical resolution problem. Interestingly, you also show that the height of the cold point tropopause in all reanalyses is always below that derived from the observations (up to 0.4 km, Fig 6). This is an important point in the current discussion if the (tropical) deep convection is able to cross the tropopause. In many studies, water vapor and ice observations are compared with the position of the cold point tropopause derived from the reanalyses data. Because of a systematic bias of the tropopause position in the reanalyses, the observed enhanced ice/water vapor values can be erroneously attributed to transport across the tropopause. Maybe you would like to discuss this point in your paper.

> *Thanks for pointing this out. This is indeed an interesting implication of the tropopause altitude comparison. We have added a statement to the summary.*

In your discussion of the inter-annual variability you quantify the contribution of the QBO, volcanic eruption and linear trends. However, you do not quantify the contribution of ENSO which is also a "major player" in such variabilities. Is it because you use a zonally averaged picture and to quantify ENSO, the zonally-resolved picture would be more appropriate? If this is the case, I would recommend to state this point more clearly.

> *Yes, including ENSO in the zonally averaged multilinear regression study does not allow for conclusive results as the zonally varying ENSO signals cancel each other out in the zonal mean analyses. We have also conducted multilinear regression of the zonally resolved temperature fields that will be discussed in a follow up publication, currently in preparation. The manuscript contains a statement explaining this '…* The influence of ENSO on TTL temperatures shows large longitudinal variations with positive anomalies over the Maritime Continent and West Pacific and negative anomalies over the East Pacific. While the zonally resolved response patterns agree well between observations and reanalyses (not shown here), the zonal mean responses are not significant (not shown here) …'.

P4 L22: Maybe you would like to mention also more recent papers for "off-line chemistry model applications", like Tao et al., 2019, ACP "Multitimescale variations..."

We have added the reference to the manuscript.

P5, L17-23: I wonder, why SHADOZ data are not mentioned here which are for me still a very important tropical data set

We have not used SHADOZ data, as this record (starting in 1999) is not long enough for the comparison of interannual variability and long-term changes evaluated here for the S-RIP core time period (1980–2010). For the zonal mean climatological analyses of the time period after 2000 we decided to use the GNSS-RO data as their uniform horizontal coverage allows to include tropical and zonal mean comparisons.

P6, L15: You explain "full-input" first in the line 41. Maybe you would like to reformulate

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). This information is given in line 15. We slightly reformulated the sentence to make this clearer.

P8, L12-15: "monthly-mean field have a warm bias of 0.5 K compared to 6-hourly data" this is not surprising. I would remove this type of motivation.

We have removed this sentence from the manuscript.

P9, L22-23: "the averaged maxima and minima values" - so you count all minima and maxima and divide it by its number? How do you define a local maximum or minimum? Maybe reformulate. In any case, this procedure is important to understand Fig. 11.

We have added the following information to the manuscript '… For each QBO cycle of this time series, the absolute temperature maximum and minimum are selected. In a second step, the means over all such temperature maxima and minima are calculated to give the averaged maximum and minimum values, respectively …'

P17, L4: I would count "volcanic" as a tropospheric variability

As the positive temperature anomalies in the upper TTL associated with volcanic eruptions are related to volcanic stratospheric aerosols, we have decided to list volcanic here under stratospheric variability.

P19, L5-6: "During the first 15 years" - or you mean during the last 15 years (higher altitude and lower pressure - I would expect the other way around)

Thanks for pointing this out. We refer here to the first 15 years and the wording was mixed up. The sentence has been corrected in the manuscript.

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.

Anonymous Referee #3

General comments:
This paper evaluates the temperature structure and tropopause characteristics in the tropical tropopause layer from various meteorological reanalysis data sets. The paper is generally well written and the results of the comparison are valuable for the community. Therefore, I recommend publication after the following specific and technical comments have been addressed.

Specific comments:
1. As accurately stated in p4 L44-45, this paper investigates "key characteristics of the temperature and tropopauses in the TTL". The title, however, gives the impression that other TTL properties are also being investigated (i.e., too broad). I suggest revising the title to indicate that the study focuses on the temperature structure and tropopause characteristics in the TTL.

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. The reasoning for choosing a certain data for certain analyses and is not always clear. Without sufficient explanation, it appears that the authors are cherry picking their results. For example:
a. Why doesn't the vertical profile for CFSR (green) in the right panel of Fig. 4 extend down to 140 hPa? Fig. 1 shows that CFSR has a model level at or just above the 140 hPa level. One of the key results, as presented in the text (e.g., Summary), is that tropical mean temperatures between 140 and 70 hPa in CFSR agrees best with those of GNSS-RO observations. I would like to see the CFSR data point near 140 hPa.

Indeed, the CFSR model level at around 138 hPa should be included in this evaluation. We have added CFSR at this level to Figure 4. Results remain unchanged as CFSR temperature at the model level around 138 hPa also agrees very well with the GNSS-RO data.

b. I would also like to see a panel using ERA5 data in Fig. 7. In all previous analyses and plots, ERA5 data are shown, but not here. Since ERA5 dataset is the newest of these reanalyses, readers will be most interested in seeing this result.

We have added the latitude–longitude comparison of cold point temperature for ERA5 to Figure 7. It shows a structure very similar to the other reanalyses when compared to the observations.

c. In Fig. 10, the temperature anomaly time series at 70 hPa (top panel) includes a time series using the RAOB radiosonde data. The second panel showing the temperature anomalies at the cold-point tropopause includes a time series using the IGRA radiosonde data. Why are the radiosonde data sources different in these two panels? Is there a reason for showing one data at 70 hPa and another at the cold point?

For consistency reasons we decided to rely on the radiosonde data sets used in Wang et al., 2013, where the authors provide detailed evaluations of the temporal variability and trends of radiosonde temperature in the TTL. Wang et al. (2013) use the unadjusted quality-controlled radiosonde data set IGRA for the cold point and several independently adjusted radiosonde temperature data sets RATPAC, HadAT, RAOBCORE, and RICH for temperatures

at 70 and 100 hPa. The motivation for evaluating interannual variability of cold point temperature, height and pressure only from the unadjusted temperature profiles is that temperature adjustments can change the location of the cold point tropopause in a profile. Therefore, we show RAOBCORE in the top panel at 70 hPa and IGRA in the lower panels for the cold point. The interannual anomalies at 70 hPa are shown only for RAOBCORE for a better clarity of the figure, while the other data sets are mentioned in the text. We have added a detailed explanation to chapter 2.1 (Observational data sets) to make clear which data sets are used at which levels for which reasons.

d. Why doesn't the right panel of Fig. 11 include data points from RATPAC, RICH and RAOBCORE (as in the left panel)?

Same reason as above.

e. The choice of radiosonde dataset in Fig. 12 is HadAT and RAOBCORE. Again, it is unclear why these two radiosonde data sets were chosen for this particular analysis. Perhaps it is best to stick to the same set of radiosonde data throughout the entire analyses?

For the trends at 70 and 100 hPa, we show the smallest and largest trends derived from the four adjusted radiosonde data sets as reported in Wang et al. (2012) and consider their range (including the reported error bars) as the observational uncertainty range. We have added this information to the Methods section.

3. There is a lot of discussion about the vertical resolution for obvious reasons (e.g., large impact on tropopause temperature). There is no mentioning of the horizontal resolution of the reanalyses data used for these comparisons. While the horizontal resolution likely plays a limited role, it would be good to document what resolution was used.

We have added information on the horizontal resolutions of the reanalyses data sets.

Technical comments:

- p5, L20: RATPAC data are mentioned, but none of the results shown in the paper use this data.

RATPAC results are shown in Figure 11.

- The second paragraph of Section 2.1 describes the various GNSS-RO measurements assimilated by the reanalyses, which are shown in Table 1. Table 1 also shows MetOp and C/NOFS data, but these are not mentioned in the text.

We have added the information to the text.

- p6, L32: ATOVS suite has a higher number of channels *compared to TOVS*?

We have changed the sentence accordingly.

- p6, L42 and p7, L12: What do you mean by "high vertical resolution"? How much higher are they compared to those of the reanalyses discussed in detail here?

*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.

- p7, L7: Is RICH also a radiosonde data (like RAOBCORE)? It is the first time this data set has been mentioned.

We have added RICH to Section 2.

- p7, L5: ERA-40 reanalysis data are not analyzed in this paper. Best to leave it out?

We have now included ERA-40 in two supplementary figures covering earlier time periods and therefore retain this text.

- p8, L12: Section 3.1 does not exist. Do you mean Section 3? Or Section 2.1?

We have changed the text to section 2.3.

- While I see the Fig. 3 caption describing the overlapped symbols, I suggest using a different symbol so that all the data points are visible.

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.

- It may be worthwhile to mention again at the beginning of Section 4 that the interannual variability of ERA5 variables are not analyzed due to the short data record. The sentence "In particular, . . .interannual variability" on p22, L25-28 is slightly misleading since the interannual variability in ERA5 is not analyzed.

We have added this information to the beginning of Section 4. We change the sentence on page 22 to 'In particular, the more recent reanalyses ERA-Interim, ERA5, MERRA-2, CFSR and JRA-55 mostly show very good agreement after 2002 in terms of the vertical TTL temperature profile, meridional tropopause structure and interannual variability.'

- p17, L34: I am having difficulty seeing the positive temperature anomalies related to Mt. Pinatubo eruption in Fig. 10 (top two panels).

Thanks for pointing this out. It is true, that following Mount Pinatubo only weak temperature anomalies occur at 70 hPa and no anomalies occur at the cold point. This is consistent with Fujiwara et al. (2015) who show that the positive temperature anomalies following Mount Pinatubo do not propagate down as far into the TTL as the ones following El Chichón. We have changed the text accordingly.

- The color of the lines for GNSS-RO and JRA-25 in Fig. 10 are difficult to distinguish. I suggest using a different color (or line style?) for JRA-25.

We have decided to keep the colors used for the reanalyses consistent with the S-RIP colour scheme. We have changed the color used for GNSS-RO to a slightly darker grey to make it more distinguishable from JRA-25.

- Fig. 11 caption: It would be helpful to mention the 1980-2010 time period in the caption.
  We have added the time period to the caption.

- p21, L31: “small negative bias at model levels *and small bias shift*, has the most realistic.”
  We have added ‘and a small bias shift’ to the text.

Anonymous Referee #4

This paper evaluates data quality of multiple atmospheric reanalyses focusing on ther- mal characteristics of the tropical tropopause layer (TTL). The comparisons are made against long-term archives of radiosonde and GNSS RO data, which provide the most accurate temperature measurements in the TTL. Purpose of the paper is very clear, methods are reasonable, and results are well organized. It provides valuable information on reanalysis data sets and is recommended for a publication in ACP after considering several minor issues listed below.

Minor issues

1. The title is too broad. The analyses focus mainly on long-term mean features and inter-annual variability of the TTL, while the title gives an expectation that it will cover overall aspects of the TTL. Annual cycle and intra-seasonal variation are also features of the TTL, particularly for dehydration processes. A more detailed title is required if authors decide not to include these features. One suggestion is making this paper as "part 1" covering long-term structures and inter-annual variability and left annual cycle and intra-seasonal variability for a future study (as this paper already has enough material, I think. . .).

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'. We make it clear from the onset that we focus on climatological and long-term characteristics by adding this information at the beginning of the abstract.

2. This is also related to comment #1. The temperature bias peaking near the equator and its potential connection to Kelvin waves (Figs. 6-8) are interesting results. This part is worth to be further investigated (even in a different paper) as it provides noble information for researchers studying the dehydration process based on reanalyses. Particularly, this feature could be "seasonally" different because temperature and circulation structures in the TTL undergo strong seasonality. The same is true for the Kelvin source over central Africa.

We agree with the reviewer that the evaluation of the seasonal cycle would be an interesting addition and should be covered in a future follow-up study. We have added a remark to the summary section.

3. Please provide some details describing how the CPT/LRT and their properties are calculated in this study. Several methods have been used to estimate properties of the CPT, and the results could be sensitive to the selected method, particularly for data set with coarse vertical resolution. This information will be helpful for readers to better understand the results provided in this paper.

We have added the following text to the manuscript 'We derive the cold point and lapse rate tropopause characteristics for each reanalysis using model-level data between 500 and 10 hPa at each grid point at 6-hourly temporal resolution. Zonal and long-term averages are calculated by averaging over all grid points, and represent the final step of data processing. For our calculations, the cold point tropopause is defined as the coldest model level. The lapse rate tropopause is defined as the lowest level at which the lapse rate decreases to 2 K km^{-1} or less, provided that the average lapse rate between this level and all higher levels within 2 km does not exceed 2 K km^{-1} (World Meteorological Organization, 1957).'

4. Given the accuracy and vertical resolution of ERA5 described in section 3, CPT temperature trend from ERA5 would be most reliable. It will be very useful if this information could be added in Fig. 12. (just suggestion)

At the moment, we have only acquired the 6 hourly data sets from ERA5 (used to calculated the cold point tropopause) for the period 2000 to 2017. We can also access a data set on the 37 standard CFMIP pressure levels for the time period 1979 to 2017, but the lower vertical resolution will impact the cold point tropopause. Therefore, we have decided to include long-term changes of ERA5 in a follow-up study that focuses on a detailed comparison of ERA-Interim and ERA5 and will evaluate long-term changes over different time periods (i.e., extending to 2018).

5. Dynamical aspect (e.g., upwelling) in the TTL is not covered in this paper. Some discussion may be beneficial (but not necessary).

We agree that a discussion of the dynamical aspects of the TTL would very interesting. However, this would require a large amount of additional material and is beyond the scope of the manuscript. Such a discussion can be found in the TTL chapter of the upcoming S-RIP report (currently under review) and related paper publications.

Technical comments

P3L20: Pan and Munchak (2011), Pan et al. (2018) could be good references for this paragraph

We have added the references to the manuscript.

P3L37: 0.5 km is roughly 5 hPa at this level, 5 hPa maybe more consistent.

Thanks for pointing this out. We have changed the text to 7 hPa.

P10L29: “near 100 hPa (ERA-Interim; -0.82 K)”. This is correct in Fig. 4 at ∼96 hPa, but could look inconsistent with Fig. 3 (right panel) as it shows ∼ -0.4 K at 100 hPa. Better to mention that it is on a model level, not 100 hPa.

We have changed the text according to the suggestions.

Fig. 4: Average on pressure level could be a bit misleading as it shows a smooth CPT. Additional figure on tropopause relative coordinate (e.g., Birner et al. 2002) could be useful.

We have decided to keep the current version of Figure 4, which aims at identifying reanalysis biases on the respective model levels. Tropopause-based averages would have the disadvantage to be affected by cold point altitude biases. Therefore, biases in a tropopause-based figure would belong to different model levels mixed up in the same tropopause-relative level.

P12L1: “comes at the expense . . . tropopause”. This expression could be a bit misleading because there is no clear causality.

We have changed the sentence to phrase this more carefully.

P13L5: “with respect to the zonal mean” => in meridional direction?

Yes, this phrase fits better. We have changed the sentence.

Fig. 7: ERA5 could provide an important clue on this issue as it has a good vertical resolution, but it is missing in the figure.

We have added ERA5 to the figure. It shows very similar structures when compared to the observations.

Fig. 8: Is the left figure different from that in Fig. 7?

Only the colour bar is different in order to contrast the comparison of CFSR over the whole time period better with the comparison during times of high Kelvin wave activity.

P17L24: datasets => data sets Fig. 10: RAOB is used for the first figure, but IGRA is used for the second figure. It will be helpful if an explanation is provided why authors made this choice. Periods ('.') are missing in several section titles and figure captions.

For consistency reasons we decided to rely on the radiosonde data sets used in Wang et al., 2013, where the authors provide detailed evaluations of the temporal variability and trends of radiosonde temperature in the TTL. Wang et al. (2013) use the unadjusted quality-controlled radiosonde data set IGRA for the cold point and several independently adjusted radiosonde temperature data sets RATPAC, HadAT, RAOBCORE, and RICH for temperatures at 70 and 100 hPa. The motivation for evaluating interannual variability of cold point temperature, height and pressure only from the unadjusted temperature profiles is that temperature adjustments can change the location of the cold point tropopause in a profile. Therefore, we show RAOBCORE in the top panel at 70 hPa and IGRA in the lower panels for the cold point. The interannual anomalies at 70 hPa are shown only for RAOBCORE for a better clarity of the figure, while the other data sets are mentioned in the text. We have added a detailed explanation to chapter 2.1 (Observational data sets) to make clear which data sets are used at which levels for which reasons.

References

Birner, T., A. Dornbrack, and U. Schumann, 2002: How sharp is the tropopause at midlatitudes? Geophys. Res. Lett., 29, 1700.

Pan, L. L., and L. a. Munchak. (2011). Relationship of cloud top to the tropopause and jet structure from CALIPSO data. J. Geophys. Res., 116, D12201.

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(19), 10-756.

~~The tropical tropopause layer in reanalysis data sets~~ Temperature and tropopause characteristics from reanalyses data in the tropical tropopause layer

Susann Tegtmeier[1], James Anstey[2], Sean Davis[3], Rossana Dragani[4], Yayoi Harada[5], Ioana Ivanciu[1], Robin Pilch Kedzierski[1], Kirstin Krüger[6], Bernard Legras[7], Craig Long[8], James S. Wang[9], Krzysztof Wargan[10,11], and Jonathon S. Wright[12]

[1]GEOMAR Helmholtz Centre for Ocean Research Kiel, 24105 Kiel, Germany

[2]Canadian Centre for Climate Modelling and Analysis, ECCC, Victoria, Canada

[3]Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, USA

[4]European Centre for Medium-Range Weather Forecasts, Reading, RG2 9AX, UK

[5]Japan Meteorological Agency, Tokyo, 100-8122, Japan

[6] Section for Meteorology and Oceanography, Department of Geosciences, University of Oslo, 0315 Oslo, Norway

[7]Laboratoire de Météorologie Dynamique, CNRS/PSL-ENS, Sorbonne University Ecole Polytechnique, France

[8]Climate Prediction Center, National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration, College Park, MD 20740, USA

[9]Institute for Advanced Sustainability Studies, Potsdam, Germany

[10]Science Systems and Applications, Inc., Lanham, MD 20706, USA

[11]Global Modeling and Assimilation Office, Code 610.1, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[12]Department of Earth System Science, Tsinghua University, Beijing, 100084, China

Abstract

The tropical tropopause layer (TTL) is the transition region between the well mixed, convective troposphere and the radiatively controlled stratosphere with air masses showing chemical and dynamical properties of both regions. The representation of the TTL in meteorological reanalysis data sets is important for studying the complex interactions of circulation, convection, trace gases, clouds and radiation. In this paper, we present the evaluation of **climatological and long-term** TTL **temperature and tropopause** characteristics in reanalysis data sets that has been performed as part of the SPARC (Stratosphere– troposphere Processes and their Role in Climate) Reanalysis Intercomparison Project (S-RIP).

The most recent atmospheric reanalysis data sets all provide realistic representations of the major characteristics of the temperature structure within the TTL. There is good agreement between reanalysis estimates of tropical mean temperatures and radio occultation data, with relatively small cold biases for most data sets. Temperatures at the cold point and lapse rate tropopause levels, on the other hand, show warm biases in reanalyses when compared to observations. This tropopause-level warm bias is related to the vertical resolution of the reanalysis data, with the smallest bias found for data sets with the highest vertical resolution around the tropopause. Differences of the cold point temperature maximise over equatorial Africa, related to Kelvin wave activity and associated disturbances in TTL temperatures. ~~Model simulations of air-mass transport into the stratosphere driven by reanalyses with a warm cold point bias can be expected to have too little dehydration.~~

Interannual variability in reanalysis temperatures is best constrained in the upper TTL, with larger differences at levels below the cold point. The reanalyses reproduce the temperature responses to major dynamical and radiative signals such as volcanic eruptions and the QBO. Long-term reanalysis trends in temperature in the upper TTL show good agreement with trends derived from adjusted radiosonde data sets indicating significant stratospheric cooling of around –0.5 to –1 K/decade. At 100 hPa and the cold point, most of the reanalyses suggest small but significant cooling trends of –0.3 to –0.6 K/decade that are statistically consistent with trends based on the adjusted radiosonde data sets.

Advances of the reanalysis and observational systems over the last decades have led to a clear improvement of the TTL reanalyses products over time. Biases of the temperature profiles and differences in interannual variability clearly decreased in 2006, when densely sampled radio occultation data started being assimilated by the reanalyses. While there is an overall good agreement, different reanalyses offer different advantages in the TTL such as realistic profile and cold point temperature, continuous time series or a realistic representation of signals of interannual variability. Their use in model simulations and in comparisons with climate model output should be tailored to their specific strengths and weaknesses.

1. Introduction

The tropical tropopause layer (TTL) is the transition region between the well-mixed, convective troposphere and the radiatively-controlled stratosphere. The vertical range of the TTL extends from the region of strong convective outflow near 12-14 km to the highest altitudes reached by convective overshooting events, around 18 km (Highwood and Hoskins, 1998; Folkins et al 1999; Fueglistaler et al., 2009; Randel and Jensen, 2013). Air masses in the TTL show dynamical and chemical properties of both the troposphere and the stratosphere, and are controlled by numerous processes on a wide range of length and time scales. Complex interactions among circulation, convection, trace gases, clouds and radiation in the TTL make this region a key player in radiative forcing and chemistry-climate coupling. As the TTL is the main gateway for air entering the stratosphere, stratospheric chemistry and composition, and especially the abundances of ozone, water vapour and aerosols, are strongly impacted by the properties of air near the tropical tropopause **(e.g., Mote et al., 1996; Holton and Gettelman, 2001; Fueglistaler et al., 2011)**.

The tropopause is the most important physical boundary within the TTL, serving to separate the turbulent, moist troposphere from the stable, dry stratosphere. The position of the tropopause is determined by the thermal properties of the TTL, as a negative, tropospheric vertical temperature gradient changes into a positive stratospheric temperature gradient. The role of the tropopause as a physical boundary is evident not only from the vertical temperature structure, but also from the distributions of atmospheric trace gases and clouds **(Pan and Munchak, 2011; Pan et al., 2018)**.

In the tropics, two definitions of the tropopause are widely used: one based on the cold point and one based on the characteristics of the lapse rate. The cold point tropopause is defined as the level at which the vertical temperature profile reaches its minimum (Highwood and Hoskins, 1998) and air parcels en route from the troposphere to the stratosphere encounter the lowest temperatures. Final dehydration typically occurs at these lowest temperatures, so that the cold point tropopause effectively controls the overall water vapour content of the lower stratosphere (Randel et al., 2004a) and explains its variability (Fueglistaler et al., 2009). While the cold point tropopause is an important boundary in the tropics where upwelling predominates, this definition of the tropopause is irrelevant for water vapor transport into the stratosphere at higher latitudes. The lapse rate tropopause, on the other hand, offers a globally-applicable definition of the tropopause, marking a vertical discontinuity in the static stability. The lapse rate tropopause is defined as the lowest level at which the lapse rate decreases to 2 K km^{-1} or less, provided that the average lapse rate between this level and all higher levels within 2 km does not exceed 2 K km^{-1} (World Meteorological Organization, 1957). The tropical lapse rate tropopause is typically ~0.5 km (~~10~~7 hPa) lower and ~1 K warmer than the cold point tropopause (Seidel et al., 2001).

Over recent decades, the thermal characteristics of the TTL and tropopause have been obtained from tropical radiosonde and Global Navigation Satellite System - Radio Occultation (GNSS-RO) upper air measurements. Radiosonde profiles offer temperature, wind and air pressure data at a high vertical resolution. However, climate records based on radiosonde data often suffer

from spatial inhomogeneities or time-varying biases due to changes in instruments and measurement practices (Seidel and Randel, 2006; Wang et al., 2012). Climate records from radio occultation data offer much better spatial coverage and density, but are only available starting from 2002. As a result, studies of long-term variability and trends in TTL and tropopause properties have also used reanalysis data (e.g., Santer et al., 2003; Gettelman et al., 2010, Xie et al., 2014).

Meteorological reanalysis data sets are widely used in scientific studies of atmospheric processes and variability, either as initial conditions for historical model runs or in comparisons with climate model output. Often, they are utilized as "stand-ins" for observations, when the available measurements lack the spatial or temporal coverage needed. Each atmospheric reanalysis system consists of a fixed global forecast model and assimilation scheme. The system combines short-range forecasts of the atmospheric state with available observations to produce best-guess, consistent estimates of atmospheric variables such as temperatures and winds. Spurious changes in the reanalysis fields can arise from changes in the quality and quantity of the observations used as input data, which complicates the analysis of variability and trends. Further discontinuities in reanalysis-based time series can originate from the joining together of distinct execution streams (Fujiwara et al., 2017).

Among the various TTL characteristics such as composition, radiation budgets and cloud properties, the vertical temperature structure and the position and temperature of the cold point are of particular importance for transport and composition studies. Many off-line chemistry-transport models or Lagrangian particle dispersion models are driven by reanalysis data sets (e.g., Chipperfield, 1999; Krüger et al., 2009; Schoeberl et al, 2012; **Tao et al., 2019**). Their representation of the cold point determines how realistically such models simulate dehydration and stratospheric entrainment processes. Process studies of TTL dynamics such as equatorial wave variability are also often based on the TTL temperature structure in reanalysis data sets (e.g., Fujiwara et al., 2012). Finally, reanalysis cold point temperature and height have been used in the past for comparison to model results and in investigations of long-term changes (e.g., Gettelman et al., 2010). Information on the quality and biases of TTL temperature and tropopause data are important for all above listed studies of transport, composition, dynamics and long-term changes of the TTL.

A comparison of the reanalysis products available at the end of the 1990s (including ERA-15, ERA-40 and NCEP-NCAR R1) with other climatological datasets showed notable differences in temperatures near the tropical tropopause (Randel et al., 2004b). While the ECWMF reanalyses agreed relatively well with radiosonde observations at 100 hPa, NCEP-NCAR R1 showed a warm bias of up to 3 K, probably resulting from low vertical resolution and the use of poorly-resolved satellite temperature retrievals (Fujiwara et al., 2017). Comparisons of winter temperatures at 100 hPa between more recent reanalyses, such as MERRA, NCEP CFSR and ERA-Interim, and Singapore radiosonde observations show better agreement, with reanalyses generally 1-2 K too cold at this level (Schoeberl et al, 2012). While many studies have highlighted the characteristics of individual reanalysis data sets, a comprehensive intercomparison of the TTL among all major atmospheric reanalyses is currently missing.

Here, we investigate whether reanalysis data sets reproduce key characteristics of the temperature and tropopauses in the TTL. This work has been conducted as part of the SPARC (Stratosphere–troposphere Processes and their Role in Climate) Reanalysis Intercomparison Project (S-RIP) (Fujiwara et al., 2017) and presents some of the key findings from the S-RIP report Chapter 8 on the TTL. Climatologies of the tropical cold point and lapse rate tropopauses as derived from modern reanalysis data sets are compared to high-resolution radio occultation data (Section 3). We also investigate temporal variability and long-term changes in tropopause levels and temperature within the TTL (Section 4). The observational and reanalysis data sets used in the evaluation are introduced in Section 2, and a discussion and summary of the results are provided in Section 5.

2 Data and methods

2.1 Observational data sets

~~High-resolution~~ Observations of the TTL **temperatures** are available from tropical radiosonde stations. However, climate records of radiosonde temperature, height and pressure data often suffer from inhomogeneities or time-varying biases due to changes in instruments or measurement practices (Seidel and Randel, 2006). Adjusted radiosonde temperature data sets at 100 hPa, 70 hPa and corresponding trends at the cold point have been created by removing such inhomogeneities (Wang et al., 2012, and references therein). In this chapter, we use ~~several~~ **the four** independently adjusted radiosonde data sets~~, including~~ RATPAC (Free et al., 2005), RAOBCORE (Haimberger, 2007), **RICH (Haimberger et al., 2012)** and HadAT (Thorne et al., 2005) **for evaluations at 70 and 100 hPa** ~~as well as the unadjusted, quality-controlled radiosonde data set IGRA (Durre et al., 2006) for over the S-RIP core time period (1981–2010).~~ **The interannual anomalies at 70 hPa are shown only for RAOBCORE to improve the clarity of the figure, but all data sets are discussed in the text. For trends at 70 and 100 hPa, we show the smallest and largest trends derived from the four adjusted radiosonde data sets as reported by Wang et al. (2012) and consider their range (including error bars) as the observational uncertainty range.**

Evaluations of the interannual anomalies of cold point temperature, height and pressure are based on the unadjusted, quality-controlled radiosonde data set IGRA (Durre et al., 2006) as temperature adjustments can change the location of the cold point in a profile. The trend of cold point temperature cannot be derived from the unadjusted IGRA data set due to inhomogeneities or time-varying biases caused by changes in instruments and measurement practices (see Wang et al., 2012 for a detailed discussion). Instead we use adjusted cold points derived from the adjusted radiosonde data sets discussed above.

Since 2002, high-resolution temperature and pressure data in the TTL are also available from satellite retrievals based on the GNSS-RO technique. Recent studies have demonstrated good agreement between GNSS-RO and radiosonde temperature profiles (e.g. Anthes et al., 2008; Ho et al., 2017). We use a **monthly mean** zonal mean data set constructed from measurements collected by the Challenging Minisat Payload (CHAMP, Wickert et al., 2001), Gravity Recovery and Climate Experiment (GRACE, Beyerle et al., 2005), Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC, Anthes et al., 2008), Metop-A

(von Engeln et al., 2011), Metop-B, Satélite de Aplicaciones Científicas-C/Scientific Application Satellite-C (SAC-C, Hajj et al., 2004), and TerraSAR-X (Beyerle et al., 2011) missions. All data are re-processed or post-processed occultation profiles with moisture information ('wetPrf' product) as provided by the COSMIC Data Analysis and Archive Center (CDAAC, https://cdaac-www.cosmic.ucar.edu/cdaac/products.html). **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.** The observational temperature records at reanalysis model levels in the TTL region have been determined by interpolating **each** GNSS-RO temperature profile **to the reanalysis model levels** with the barometric formula, taking into account the lapse rate between levels. For each profile, the cold point and lapse rate tropopause characteristics were identified based on the cold point and WMO criteria, respectively. **Zonal and long-term averages of the tropopause metrics and temperatures at model levels are calculated by averaging over all grid points and represent the final step of data processing.**

We also use a daily data set of cold point temperatures obtained from all GNSS-RO missions, gridded on a 5°x5° grid between 30°N and 30°S. For each 5° wide latitude band, we apply a two-dimensional fast Fourier transform to detect Kelvin wave anomalies for planetary wavenumbers 1–15, periods of 4–30 days and equivalent depths of 6–600 following the theoretical dispersion curves for Kelvin waves as in Wheeler and Kiladis (1999). We allow a wider range of equivalent depths, since it has been shown that Kelvin waves tend to propagate faster around the tropical tropopause than they do in the troposphere (Kim and Son, 2012). The filtered anomalies represent cold point temperature variability that propagates in the same wavenumber-frequency domain as Kelvin waves, i.e. when the temperature is modulated by Kelvin waves present around the tropopause. The spatial variance of the filtered signals is used to calculate a monthly index as a measure of the amount of Kelvin wave activity in the TTL. The index is calculated as the 1σ standard deviation over the filtered anomalies at all spatial grid points. Time periods of enhanced Kelvin wave activity are defined as the months when the index is larger than the long-term mean plus the 1σ standard deviation of the whole time series. Based on this definition, we determined 20% of all months to be characterized by enhanced Kelvin wave activity.

2.2 Reanalysis data sets

We evaluate eight "full-input" reanalyses, **where a full-input reanalysis is** defined ~~here~~ as **a** systems that assimilate surface and upper-air conventional and satellite data. In this paper, we focus on ERA-Interim (Dee et al., 2011), ERA5 (Hersbach et al., 2018), JRA-25 (Onogi et al., 2007), JRA-55 (Kobayashi et al., 2015), MERRA (Rienecker et al., 2011), MERRA-2 (Gelaro et al., 2017), NCEP-NCAR Reanalysis 1 (Kistler et al., 2001; referred to hereafter as R1), and CFSR (Saha et al., 2010). We limit our analyses to the S-RIP core intercomparison period 1980–2010. Due to availability at the time of the evaluations, ERA5 is only evaluated over

2002–2010. Details of each reanalysis, including model characteristics, physical parameterizations, assimilated observations, execution streams, and assimilation strategies have been summarized by Fujiwara et al. (2017). ~~MERRA-2 data access was through the Global Modeling and Assimilation Office (GMAO, 2015).~~

Global temperature fields in the reanalysis data sets are ~~produced~~ **constrained** by assimilating conventional (surface and balloon), aircraft, and satellite observations. The most important sources of assimilated data for stratospheric temperatures are the microwave and infrared satellite sounders of the TOVS suite (1979–2006) and the ATOVS suite (1998–present). All of the above reanalysis systems assimilate microwave and infrared radiances from these instruments, except for NCEP-NCAR R1 which assimilates temperature retrievals instead. Measurements from the ATOVS suite, which has a higher number of channels **compared to TOVS**, have been assimilated from about 1998, although the exact start dates differ among the reanalyses. The introduction of ATOVS considerably improved the vertical resolution of the assimilated data. Some of the reanalyses (ERA-Interim, ERA5, MERRA, MERRA-2, and CFSR) also assimilate radiances from the hyperspectral infrared sounders AIRS (2002–present), IASI (2008–present), and/or CrIS (2012–present), although the latter was not available for assimilation during the intercomparison period considered here. ~~Radiance biases associated with instrument changes, inaccurate calibration offsets, orbital drifts or long-term CO_2 changes can propagate into the reanalysis fields (e.g. Rienecker et al., 2011).~~ **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.**

All full-input reanalyses assimilate upper-air temperature observations from radiosondes which are available at a very high vertical resolution. Systematic errors in radiosonde profiles caused by effects of solar radiative heating on the temperature sensor (Nash et al., 2011) have typically been corrected either onsite or at the reanalysis centre before assimilation (Fujiwara et al., 2017). 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). Earlier reanalyses (ERA-40 and JRA-25) used simplified homogenization approaches that mostly corrected for daily and seasonal variations. Although the detailed quality control procedures for radiosonde and other conventional data imported from the global distribution network can vary among the individual reanalyses, the conventional data archives are often shared among the centres (see also Fujiwara et al., 2017).

Recent reanalysis systems have also included information from GNSS-RO instruments by assimilating observations of the bending angle up to 30 km (Cucurull et al., 2013). Assimilating these high vertical resolution data affects reanalysis temperature and provides an additional ‘anchor’ for adaptive bias correction of satellite radiances. JRA-55 assimilates refractivity profiles up to 30 km, which are functions of temperature, humidity, and pressure. For all recent reanalyses, the advent of the COSMIC mission in 2006 significantly increased the number of GNSS-RO profiles available for assimilation. Details of the various GNSS-RO data assimilated

by ERA5, ERA-Interim, JRA-55, MERRA-2 and CFSR up to the end of 2010 are listed in **Table 1**. **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.**

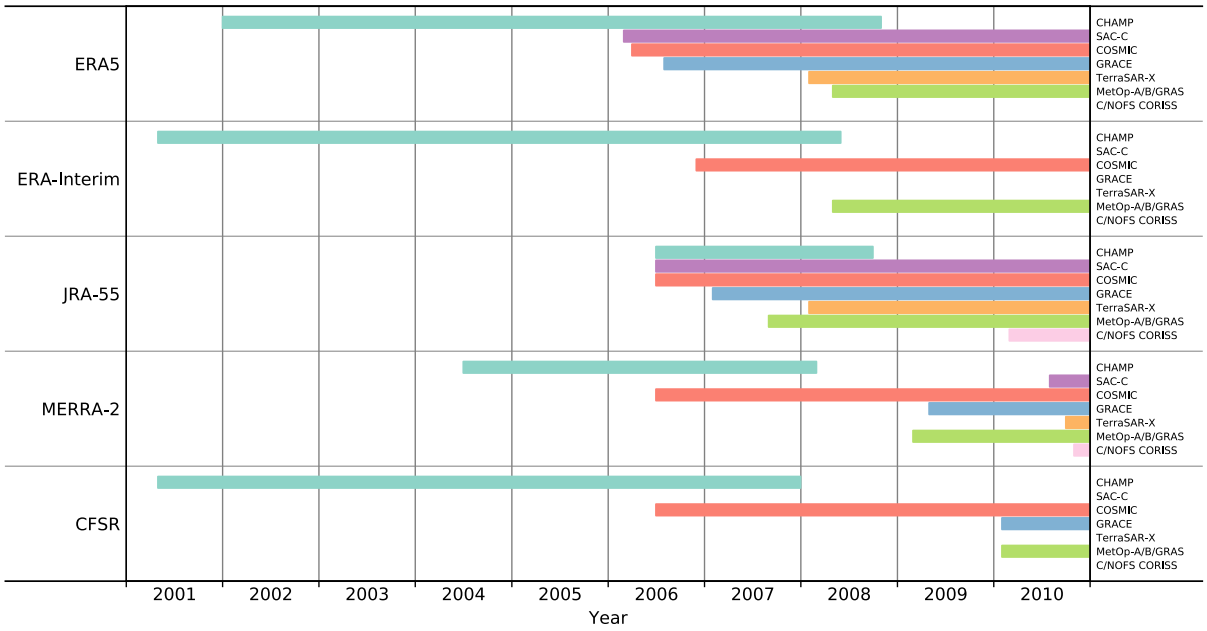

Table 1. List of GNSS-RO data assimilated by the reanalysis systems with starting dates prior to the end of 2010.

Among the observational data sets, radiosonde and GNSS-RO data are our best source of information about the TTL. While the reanalyses assimilate **different** versions of these data, it is not automatic that they reproduce the data **within their uncertainty** ~~exactly~~. For instance, discrepancies exist between reanalysis stratospheric temperatures and those derived from their radiance input data (Long et al., 2017). In fact, it is a subject of ongoing research how well reanalyses fit the data they assimilate (Simmons et al., 2014, Wright and Hindley 2018). ~~In general, the~~ **The** data assimilation ~~methodology relies on an interplay among a vast number of diverse observations, model simulations and bias correction schemes.~~ **systems combine information from a model, a set of observations and a priori information weighted by their respective uncertainties.** The degraded vertical resolution of the reanalyses, compared to radiosonde and GNSS-RO data also leads to differences, especially for derived quantities such as the tropopause location and temperature, which will be investigated in the following evaluations.

The reanalysis models resolve the TTL with different vertical resolutions, as illustrated in **Fig. 1**. The number of model levels between 200 and 70 hPa varies among the reanalyses from a low of 4 (NCEP-NCAR R1) to a high of 21 (ERA5), corresponding to vertical resolutions between ~1.5 km and ~0.2 km. In addition to the native model levels, all reanalyses provide post-processed data on standard pressure levels with at least four levels situated between 200 and 70 hPa (**Fig. 1**). **The horizontal resolutions of the reanalysis products are approximately 0.25°×0.25° (ERA5), 0.7°×0.7° (ERA-Interim), 0.63°×0.5° (MERRA-2), 0.66°×0.5° (MERRA), 0.56°×0.56° (JRA-55), 1.13°×1.13° (JRA-25), 0.5°×0.5° (CFSR), and**

1.9°×1.9° (R1). ~~We show that reanalysis-based estimates of tropopause temperature, pressure and height compare much better to observations when they are derived from model-level data than when they are derived from pressure-level data (Section 3.1). Another sensitivity study demonstrates that tropopause temperatures directly calculated from monthly-mean fields have a warm bias of 0.5 K compared to tropopause temperatures based on 6-hourly data (not shown here). Therefore, we derive the cold point and lapse rate tropopause characteristics for each reanalysis using model-level data at each grid point at 6-hourly temporal resolution. Zonal and long-term averages are then calculated by averaging over all grid points, and represent the final step of data processing.~~

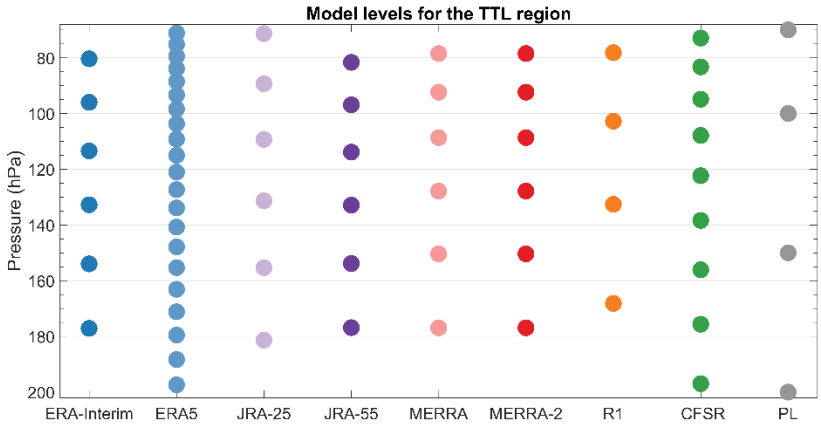

Figure 1. Model-level pressure values for different reanalysis data sets in the TTL using a fixed surface pressure of 1013.25 hPa. Standard pressure levels (PL) in the TTL region are also shown.

2.3 Methods

Given the strong gradients of temperature and static stability in the TTL, the vertical resolution of the reanalysis data sets is an important factor in cold point and lapse rate tropopause calculations. For each reanalysis, tropopause heights and temperatures can be derived either from model- or pressure-level data (**Fig 1**). A comparison of the CFSR cold point tropopause based on model- and pressure-level temperature data is shown here to demonstrate the clear advantage of the finer model-level resolution (**Fig. 2**). The cold point tropopause from CFSR model-level data for the time period 2002–2010 agrees well with radio occultation results, with differences of less than 1.5 K and 0.2 km at all latitudes. The tropopause derived from CFSR pressure-level data, on the other hand, shows larger differences. This estimate is up to 0.4 km too low and up to 3 K too warm, illustrating the need to use data with high vertical resolution to identify and describe the tropopause. **We derive the cold point and lapse rate tropopause characteristics for each reanalysis using model-level data between 500 and 10 hPa at each grid point at 6-hourly temporal resolution. Zonal and long-term averages are calculated by averaging over all grid points, and represent the final step of data processing. For our calculations, the cold point tropopause is defined as the coldest model level. The lapse rate tropopause is defined as the lowest level at which the lapse rate decreases to 2 K km^{-1} or less, provided that the average lapse rate between this level and all higher levels within 2 km does not exceed 2 K km^{-1} (World Meteorological Organization, 1957).** ~~The following climatological tropopause comparisons are all based on model-level data.~~

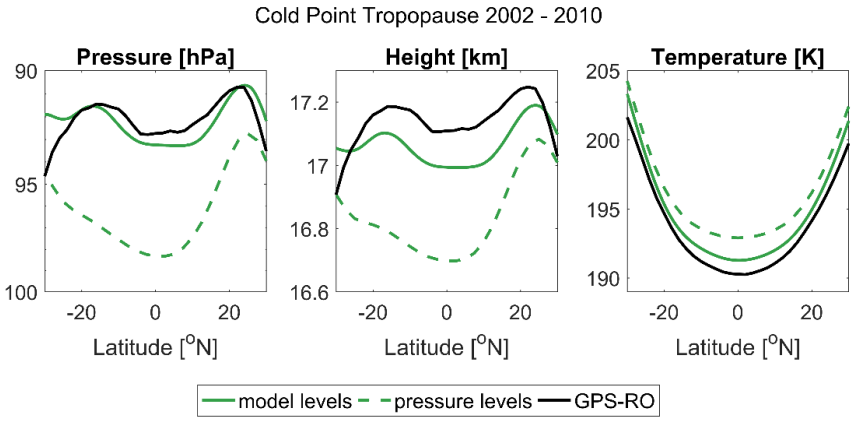

Figure 2. Latitudinal distributions of zonal-mean cold point tropopause pressure (left), altitude (centre) and temperature (right) based on radio occultation data (black) and CFSR model-level (green solid) and pressure-level (green dashed) data during 2002–2010.

The evaluation of the interannual variability (Section 4) is based on time series of deseasonalized monthly temperature, pressure and altitude anomalies calculated relative to the mean annual cycle during 2002–2010. To study variability driven by tropospheric and stratospheric forcing, we identify and isolate the variations based on a standard multivariate regression analysis:

$$T(t) = A_1 \cdot \mathrm{QBO1}(t) + A_2 \cdot \mathrm{QBO2}(t) + B \cdot \mathrm{ENSO}(t) + D \cdot \mathrm{VOL}(t). \quad (1)$$

Here $\mathrm{QBO1}(t)$ and $\mathrm{QBO2}(t)$ are orthogonal time series representing QBO variations constructed as the first two EOFs of the Freie Universität Berlin (FUB) radiosonde stratospheric winds (Naujokat, 1986). $\mathrm{ENSO}(t)$ is the multivariate ENSO index (https://www.esrl.noaa.gov/psd/enso/mei/) and $\mathrm{VOL}(t)$ is the stratospheric aerosol optical depth from the Global Space-based Stratospheric Aerosol Climatology (Thomason et al., 2018). The standard error of the regression coefficients was derived based on the bootstrap method **(Efron and Tibshirani, 1993)**. The QBO temperature amplitude is calculated as the difference between the averaged maxima and averaged minima values of the time series of the QBO temperature variations $A_1 \cdot \mathrm{QBO1}(t) + A_2 \cdot \mathrm{QBO2}(t)$. **For each QBO cycle of this time series, the absolute temperature maximum and minimum are selected. In a second step, the means over all such temperature maxima and minima are calculated to give the averaged maximum and minimum values, respectively.**

The long-term trends of the reanalyses temperature time series have been derived as the regression coefficient of a linear function that provides the best fit in a least-squares sense. The **uncertainty in each long-term** trend ~~s error bars are~~ **is calculated** as the standard error of the slope with ~~an~~ **the** effective sample size **adjusted to account for the corresponding lag-1 autocorrelation coefficient**. Significance is tested based on two-tailed test with a 95% confidence interval.

3 Temperature and tropopause characteristics

Tropical mean temperatures from reanalyses at two standard pressure levels (100 hPa and 70 hPa) and at the two tropopause levels are compared to radio occultation data for the time period 2002–2010 (**Fig. 3**). At 100 hPa, reanalysis temperatures agree well with radio occultation data with differences between –0.35 K (too cold; ERA-Interim and ERA5) and 0.43 K (too warm; CFSR). At 70 hPa, the agreement is even better, with differences ranging from –0.29 K (JRA-55) to 0.12 K (JRA-25). However, nearly all reanalyses show warm biases at both tropopause levels, with differences of up to 1.2 K compared to the observations. Most likely, the excess warmth of tropopause estimates based on reanalysis products stems from the limited vertical resolution of the reanalysis models in the TTL region. The best agreement is found for the reanalysis with the highest vertical resolution (ERA5; 0.05 K too warm at the cold point tropopause). The data set with the lowest vertical resolution (NCEP-NCAR R1) is 2.2 K too warm, outside the range displayed in Figure 3.

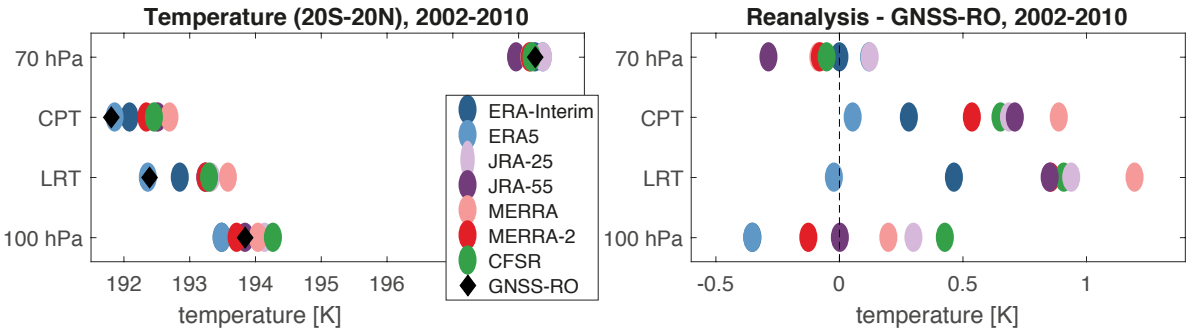

Figure 3. Tropical mean (20°S-20°N) temperatures at 100 hPa, the lapse rate tropopause (LRT), the cold point tropopause (CPT) and 70 hPa from reanalyses and GNSS-RO data during 2002–2010 (left panel). Differences between the GNSS-RO and reanalysis temperatures are shown in the right panel. At 100 hPa, ERA-Interim is hidden by ERA-5, at the LRT, MERRA-2 is hidden by JRA-55, and at 70 hPa, ERA5 is hidden by JRA-25 and MERRA is hidden by MERRA-2.

Temperature profile comparisons between 140 and 70 hPa at the native model levels ~~resolution~~ have been conducted for the five most recent reanalyses (ERA5, ERA-Interim, JRA-55, MERRA-2, CFSR). All reanalyses tend to be colder than the observations in the tropical mean (**Fig. 4**), but differences are relatively small and the agreement is good overall. CFSR and ERA5 agree best with the radio occultation data with mean biases of around –0.06 K and –0.28 K, respectively, averaged over the whole vertical range. ERA-Interim and MERRA-2 agree very well at upper levels but show large deviations **on model levels** near 100 hPa (ERA-Interim; –0.82 K) and below 110 hPa (MERRA-2; –0.67 K), respectively. The evaluation demonstrates that temperature comparisons at standard pressure levels (**Fig. 3**) can be biased by up to 0.5 K, with CFSR showing a positive bias (0.45 K) at the 100 hPa standard pressure level but very good agreement (-0.05 K) at nearby native model levels. Such biases can result from vertical interpolation of temperature data in regions with large lapse rate changes.

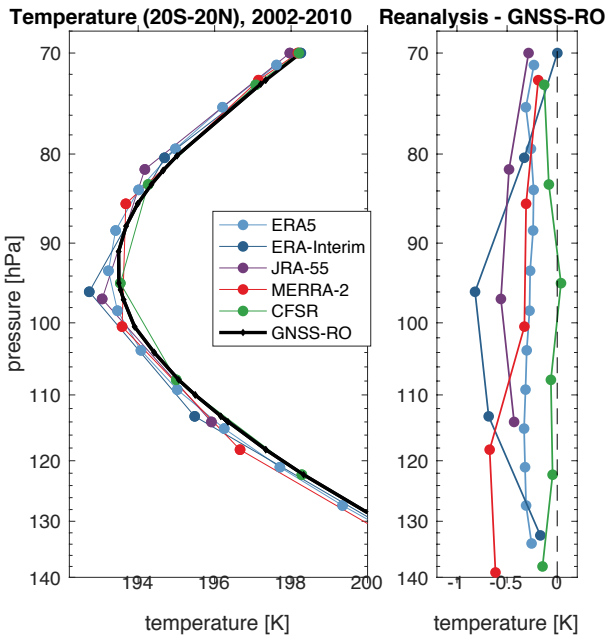

Figure 4. Tropical mean (20°S–20°N) temperature profiles at reanalysis model levels between 140 and 70 hPa (left panel) during 2002–2010 and differences between reanalysis and GNSS-RO temperatures (right panel).

Comparing the temperature profiles to the tropopause values (**Fig. 3 and 4**) reveals that despite the five reanalyses having negative biases at model levels, they mostly have positive biases at the cold point and lapse rate tropopause levels. As the discrete values corresponding to reanalysis model levels are unable to reproduce the true minimum temperature as recorded in a near-continuous profile, this difference is expected for the cold point tropopause. Similarly, the lapse rate tropopause criteria might typically be fulfilled at lower levels for data at coarser resolution, thus resulting in a warm bias at the lapse rate tropopause on average. Overall, our results indicate that the negative temperature bias at model levels is more than cancelled out by the positive bias introduced when calculating the cold point and lapse rate tropopauses. Linking the temperature profile and tropopause comparisons, this 'bias shift' is about 0.3 K for ERA5, 0.6 K for CFSR and 1 K or larger for ERA-Interim, MERRA-2 and JRA-55. In consequence, ERA5, with both a small negative bias at the model levels and a small bias shift provides the most realistic tropopause temperatures. CFSR also has a relatively small bias shift, but the relatively unbiased temperature profile does not permit any error cancelation via this shift, so that cold point and lapse rate tropopause levels based on CFSR are systematically too warm.

Agreement between the reanalysis temperature profiles and GNSS-RO data clearly improves when the comparison is restricted to the 2007–2010 time period, when the more densely-sampled COSMIC data were assimilated (**Table 1**). This point is illustrated by comparison of temperature time series from reanalyses and observations at two model and both tropopause levels (**Fig. 5**). For ERA5, ERA-Interim and MERRA-2, the cold bias with respect to GNSS-RO at model levels decreases after 2007, most likely because of the high number of daily COSMIC profiles available for assimilation from this time onwards. Cold biases at model levels are accompanied by warm biases in the tropopause temperatures, which, for ERA-Interim and ERA5, increase after 2007. **As the increase at all levels is very similar, this indicates** that the

advantage of a reduced temperature bias at model levels comes at the expense of an increased temperature bias at the tropopause. CFSR and MERRA-2 show no such systematic change of their tropopause temperatures over time when compared to GNSS-RO data. JRA-55 is the only reanalysis product for which cold point and lapse rate tropopause temperatures agree slightly better with GNSS-RO estimates after 2007.

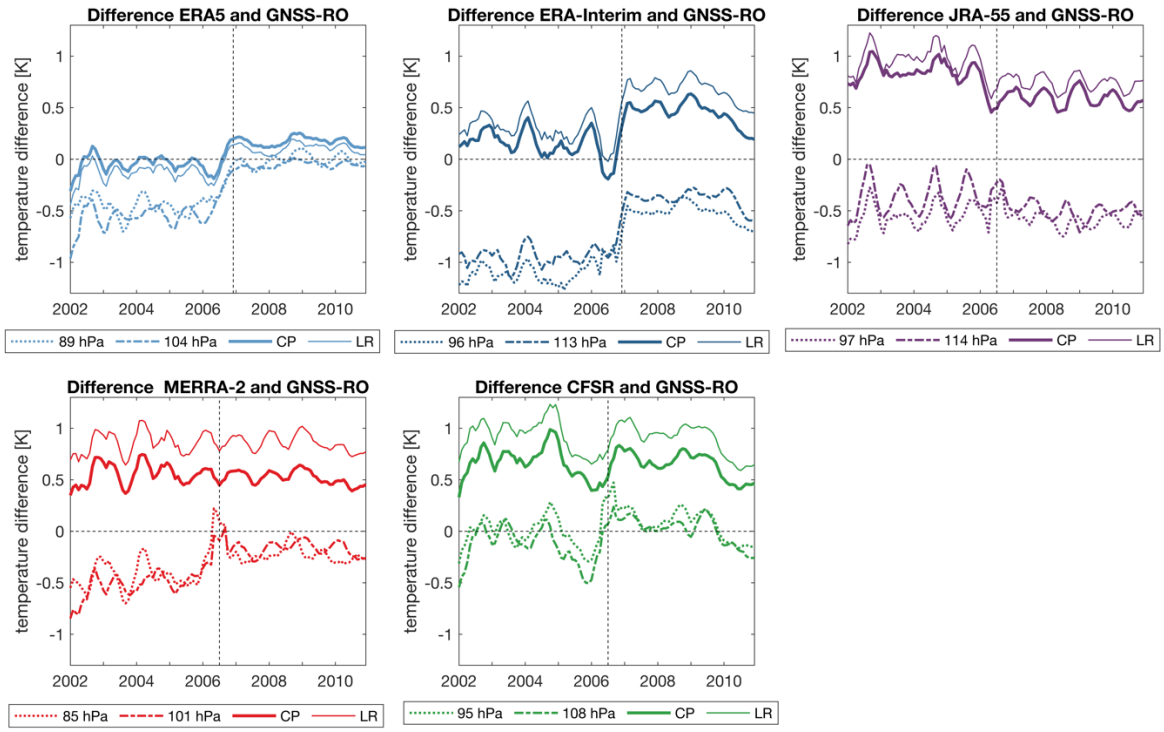

Figure 5. Tropical mean (20°S-20°N) time series of temperature differences between reanalysis and radio occultation at the cold point (CP) and lapse rate (LR) tropopause levels, as well as selected reanalysis model levels. Vertical lines indicate when the assimilation of COSMIC radio occultation data started.

Evaluations of the latitudinal structure of the cold point tropopause for 2002–2010 are based on comparisons to radio occultation data (**Fig. 6**). All reanalysis data sets produce tropopause levels that are too low and too warm, with the latter related to vertical resolution as explained above. The observations show that average cold point temperatures are lowest right around the equator. The reanalyses fail to reproduce this latitudinal gradient, indicating more constant cold point temperatures across the inner tropics between 10°S and 10°N with a less pronounced minimum at the equator. As a consequence, the largest differences in cold point tropopause temperatures relative to GNSS-RO data are at the equator and the best agreement is around 20°S/20°N for all reanalysis data sets.

The cold point altitude and pressure exhibit little north–south variability, ranging from 16.9 km (94 hPa) to 17.2 km (91.8 hPa). With respect to the seasonal cycle, it is well known that the temperature and altitude of the cold point tropopause are linked, with the coldest temperatures and highest altitudes observed during boreal winter (e.g., Seidel et al., 2001; Kim and Son, 2012). This relationship does not hold **in the meridional direction** ~~with respect to the zonal mean~~: the highest cold point altitudes are located around 20°S/20°N, while the lowest cold point temperatures are located near the equator. The higher altitude/lower pressure of the cold point tropopause around 20°S/20°N results from zonally-variable features linked to tropospheric pressure regimes, such as particularly low tropopause pressures over the Tibetan plateau during boreal summer (Kim and Son, 2012). The reanalysis data sets capture most of this latitudinal structure, showing roughly constant differences between about 0.1 and 0.2 km (0–2 hPa). The largest differences are found for NCEP-NCAR R1 in the Southern Hemisphere, where the cold point tropopause based on R1 is both higher and warmer than observed. The best agreement with respect to cold point temperatures is found for ERA5 and ERA-Interim, which are around 0.2 K and 0.4 K warmer than the radio occultation data, respectively. All other reanalysis data sets are in close agreement with each other, with differences from the observations of between 0.5 K and 1 K. The altitude and pressure of the cold point tropopause are captured best by ERA5, CFSR, MERRA, MERRA-2 and JRA-55, which all produce cold point tropopauses that are slightly too low (~0.1 km). ERA-Interim, despite very good agreement in cold point temperature, shows slightly larger biases in cold point altitude (~0.2 km) relative to the GNSS-RO benchmark. **Zonal mean cold point tropopause temperatures, altitudes and pressures during 1981-1990 and 1991-2002 are shown for all reanalyses in the supplementary Figure S1.**

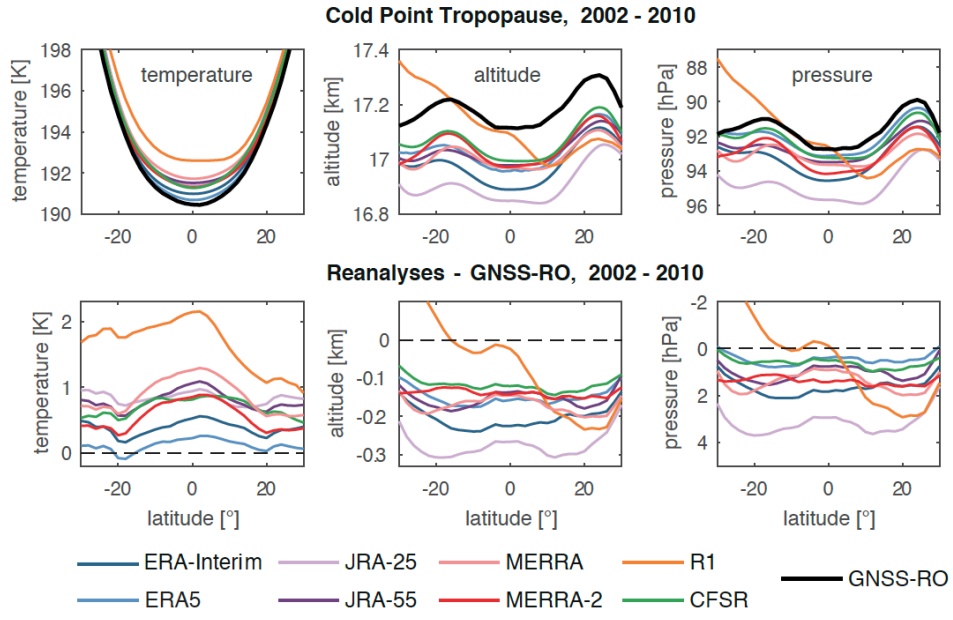

Figure 6. Latitudinal distributions of zonal-mean cold point tropopause temperature (left), altitude (centre) and pressure (right) based on radio occultation data and reanalysis products during 2002–2010 (upper row). Differences between radio occultation and reanalysis estimates are shown in the lower row.

We investigate the temperature biases and their maxima near the equator by analysing latitude–longitude variations in the cold point tropopause relative to GNSS-RO estimates for four of the reanalyses (**Fig. 7**). To show differences at relatively high spatial resolution, we focus on the period 2007–2010. A wealth of observational studies has shown that the coldest tropopause temperatures are located over the ‘‘maritime continent’’ and the West Pacific (Highwood and Hoskins, 1998), with secondary minima over equatorial South America and Africa coinciding with other centres of deep convective activity (Gettelman et al., 2002). The colocation of tropospheric convective activity with zonal asymmetries in cold point temperature can be explained by the radiative cooling effects of cirrus clouds overlying deep convection (Hartmann et al., 2001) or diabatic cooling associated with convective detrainment (Sherwood et al., 2003). Furthermore, it has been suggested that the response of equatorial waves to convective heating influences the structure of the cold point tropopause (Kim and Son, 2012; Nishimoto and Shiotani, 2012; Nishimoto and Shiotani, 2013). The dominant wave modes responsible for cold point temperature variability are linked to equatorial Kelvin waves and the Madden-Julian oscillation.

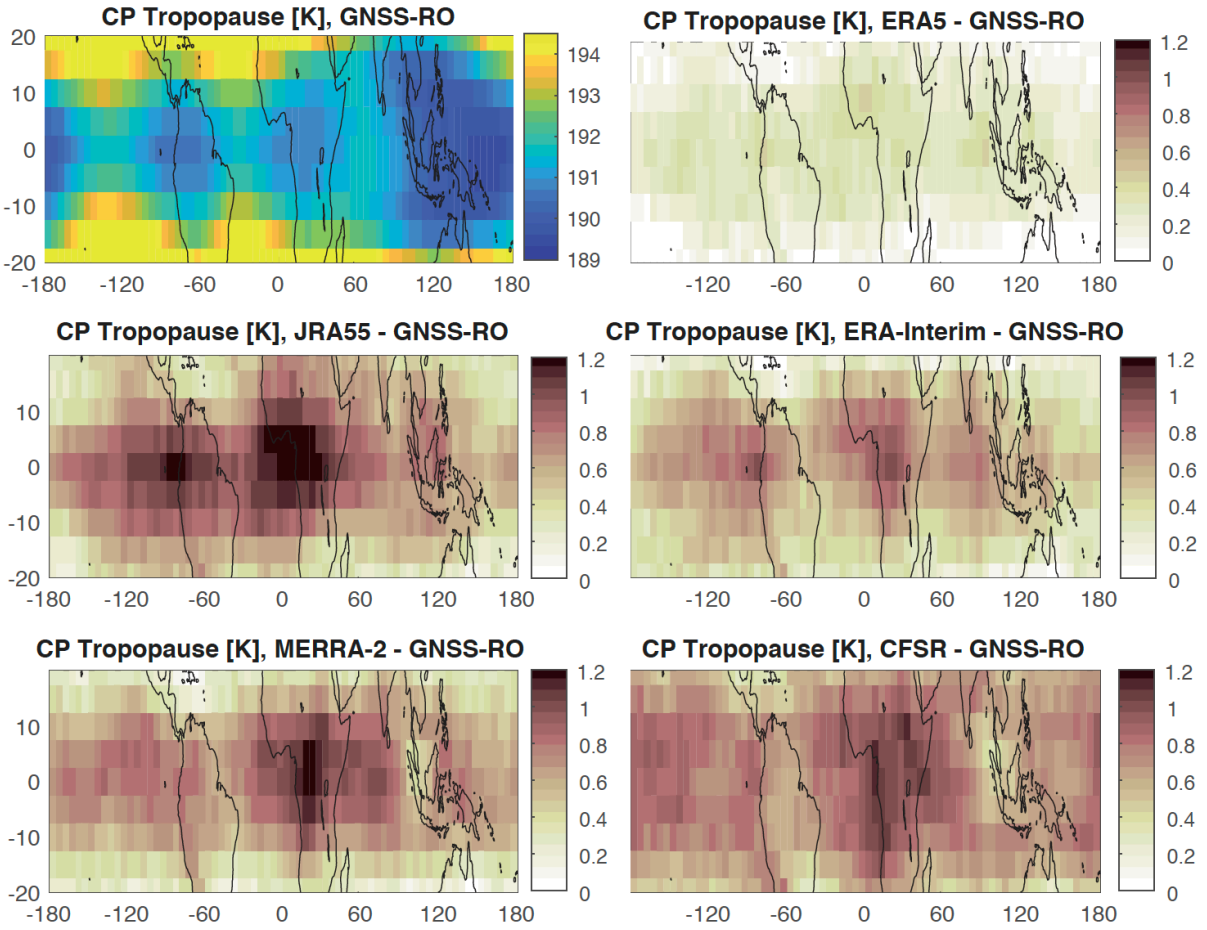

Figure 7. Latitude–longitude distributions of annual mean GNSS-RO cold point temperatures (upper left) and differences between cold point temperatures from individual reanalyses and those from GNSS-RO during 2007–2010 (lower panels).

For the analysed reanalyses (**ERA5**, ERA-Interim, MERRA-2, JRA55, and CFSR), differences with respect to the observations are largest in the inner tropics over central Africa, reaching values 50% to 100% greater than the zonal mean differences. This region is characterized by a local cold point minimum that results from deep convection and its interaction with equatorial waves. There is also evidence of a secondary maximum in the differences over equatorial South America or the East Pacific, although the magnitude and location of this maximum differ among the reanalyses.

The convective centre over the Western Pacific warm pool, where the cold point tropopause is coldest, does not show enhanced biases relative to the observations. One possible explanation for the bias distribution might link the enhanced temperature differences to Kelvin wave activity that maximizes over Central Africa but is weaker over the West Pacific (Kim et al., 2019). As the Kelvin waves disturb the temperature profile at small vertical scales, the reanalyses may be particularly unsuited to estimate cold point temperatures in regions of strong Kelvin wave activity. We average cold point temperatures from reanalyses and observations over time periods of enhanced Kelvin wave activity. For CFSR, composite differences for periods with enhanced wave activity are compared in **Fig. 8** to mean differences averaged over the whole 2007–2010 period. While mean biases over Central Africa are less than 1 K, average differences during periods of enhanced Kelvin wave activity are as large as 1.4 K. The same is true for other reanalyses (not shown here), with the exception of ERA-Interim, suggesting that in most cases Kelvin waves contribute to the spatial structure of biases in cold point tropopause estimates based on reanalysis products.

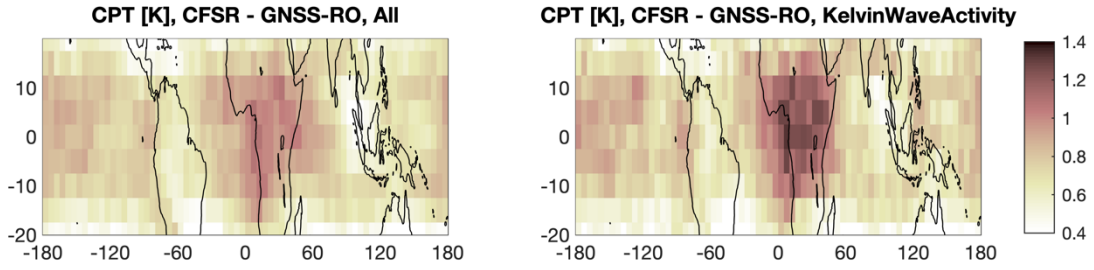

Figure 8. Latitudinal-longitude sections of the differences between GNSS-RO and CFSR cold point temperatures for 2007-2010 (left panel) and for time periods of high wave activity (right panel).

The zonal mean lapse rate tropopause (**Fig. 9**) at the equator is found at similar temperatures and heights as the cold point tropopause, being only slightly warmer and lower. Poleward of 10°S/10°N, however, the lapse rate tropopause height decreases considerably faster than the cold point height, since the cold point is more often located at the top of the inversion layer while the lapse rate tropopause is located at the bottom of the inversion layer (Seidel et al., 2001). Lapse rate tropopause temperatures based on reanalysis data are on average about 0.2 K to 1.5 K too warm when compared to radio occultation data (see **Fig. 3** and associated

discussion) with best agreement for ERA5 and ERA-Interim. Consistent with this temperature bias, lapse rate tropopause levels based on reanalysis data are about 0.2 km to 0.4 km lower than those based on radio occultation data. The latitudinal structure of lapse rate tropopause temperatures reveals slightly larger biases at the equator and better agreement between 10°–20° in each hemisphere, and is generally very similar to the latitudinal distribution of biases in cold point temperatures (**Fig. 6**). The altitude of the lapse rate tropopause shows considerable ~~zonal~~ **meridional** variability, ranging from 14.5 km to 16.7 km. All reanalyses capture the plateau in lapse rate tropopause altitudes between 20°S and 20°N and the steep gradients in these altitudes on the poleward edges of the tropics. **Zonal mean lapse rate tropopause temperatures, altitudes and pressures during 1981-1990 and 1991-2002 are shown for all reanalyses in the supplementary Figure S2.**

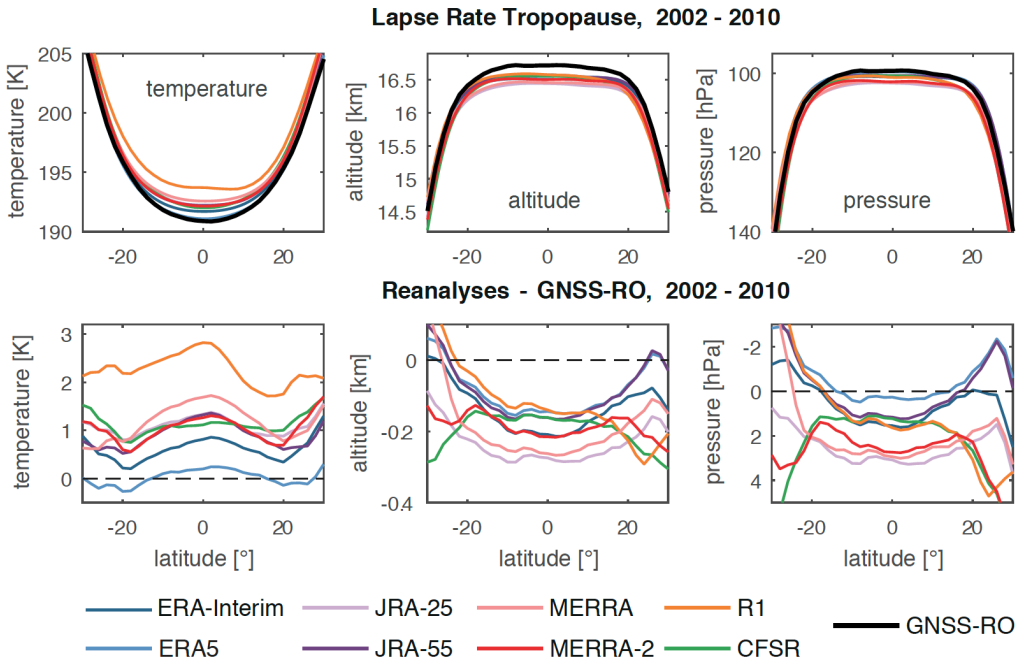

Figure 9. Latitudinal distributions of zonal-mean lapse rate tropopause temperature (left), altitude (centre) and pressure (right) based on radio occultation data and reanalysis products during 2002–2010 (upper row). Differences between radio occultation and reanalysis estimates are shown in the lower row.

4 Interannual variability and long-term changes

It has long been recognized that inter-annual variations in TTL temperatures are strongly affected by both tropospheric (e.g. ENSO) and stratospheric (e.g. QBO, solar, volcanic) variability (Randel et al., 2000; Zhou et al., 2001, Krüger et al 2008). Time series of deseasonalized monthly 70 hPa temperature anomalies and cold point temperature, pressure and altitude anomalies are shown in **Fig. 10**. Anomalies are calculated relative to the mean annual cycle during 2002–2010 for each dataset. **The interannual variability of ERA5 is not analysed due to the short data record available at the time of the analysis.** The performance of the reanalyses with respect to both the spread among reanalyses and their agreement with observations is much better at the 70 hPa level than at the cold point level. The older reanalyses NCEP-NCAR R1 and JRA-25 generally show larger deviations from the RAOBCORE time series. The level of agreement among the reanalyses and between reanalyses and observations improves over time, with a step-like improvement around 1998–1999 that is likely associated with the TOVS-to-ATOVS transition. The higher vertical resolution of measurements from the ATOVS suite (see, e.g. Figure 7 in Fujiwara et al., 2017) is known to reduce differences among the reanalysis with respect to stratospheric temperature (Long et al., 2017) and polar diagnostics (Lawrence et al., 2018). Within the TTL, temperature biases **decrease** ~~improve~~ from values of 1–2 K to around 0.5 K following the TOVS-to-ATOVS transition. This agreement improves further after 2002, when many of the more recent reanalyses started assimilating AIRS and GNSS-RO data (**Table 1**; see also Figure 8 in Fujiwara et al., 2017).

Interannual variability at 70 hPa is dominated by the stratospheric QBO signal, which is reproduced by all reanalysis data sets. The amplitudes of the QBO temperature variations in all datasets based on a multilinear regression analyses over 1981-2010 are shown in **Fig. 11**. At 70 hPa, the observational radiosonde data sets give QBO variations of 2.1–2.2 K. Reanalyses agree well with the observations and show QBO variations of 2–2.4 K. The only exception is NCEP-NCAR R1, which clearly underestimates the signal compared to radiosondes and other reanalyses, with an amplitude of 1.7 K. Best agreement with the radiosonde data sets is found for MERRA-2, MERRA and CFSR. The influence of ENSO on TTL temperatures **(not shown here)** shows large longitudinal variations with positive anomalies over the maritime continent and West Pacific and negative anomalies over the East Pacific. While the zonally resolved response patterns agree well between observations and reanalyses ~~(not shown here)~~, the zonal mean responses are not significant ~~(not shown here)~~. Positive temperature anomalies following the eruptions of El Chichón in 1982 ~~and Mount Pinatubo in 1991~~ can be detected in **Fig. 10** for all reanalyses, consistent with the results of Fujiwara et al. (2015). **Following the Mount Pinatubo eruption in 1991, small positive temperature anomalies are evident at the 70 hPa level around the beginning of 1992. However, no positive temperature anomalies are found at the cold point during this time (see Fujiwara et al., 2015 for a more detailed analysis).**

At the cold point, NCEP-NCAR R1 is a clear outlier, with much warmer temperature anomalies than any other data set during the period prior to 2005 **(Fig. 10)**. However, differences among the more recent reanalyses are also relatively large, with ERA-Interim (on the lower side) and CFSR (on the upper side) showing differences as large as 2 K in the early years of the comparison. Given that existing homogenized radiosonde data sets also show deviations of up

to 1.5 K at this level (Figure 2 in Wang et al., 2012), we cannot deduce which reanalysis data set is most realistic. Note that the radiosonde time series from IGRA shown here should not be used for evaluating long-term changes (see Wang et al., 2012 for details), but only for assessing the representation of interannual variability. Periods of particularly pronounced interannual variability alternate with relatively quiescent ones. The amplitude of **interannual variability (Fig. 10) and** the QBO temperature signal **(Fig. 11)** are weaker at the cold point than at 70 hPa, but are still well captured by all of the reanalysis data sets except for NCEP-NCAR R1. ~~The amplitude of interannual variability is also smaller at the cold point than at 70 hPa, but with larger month-to-month variations (not shown here).~~

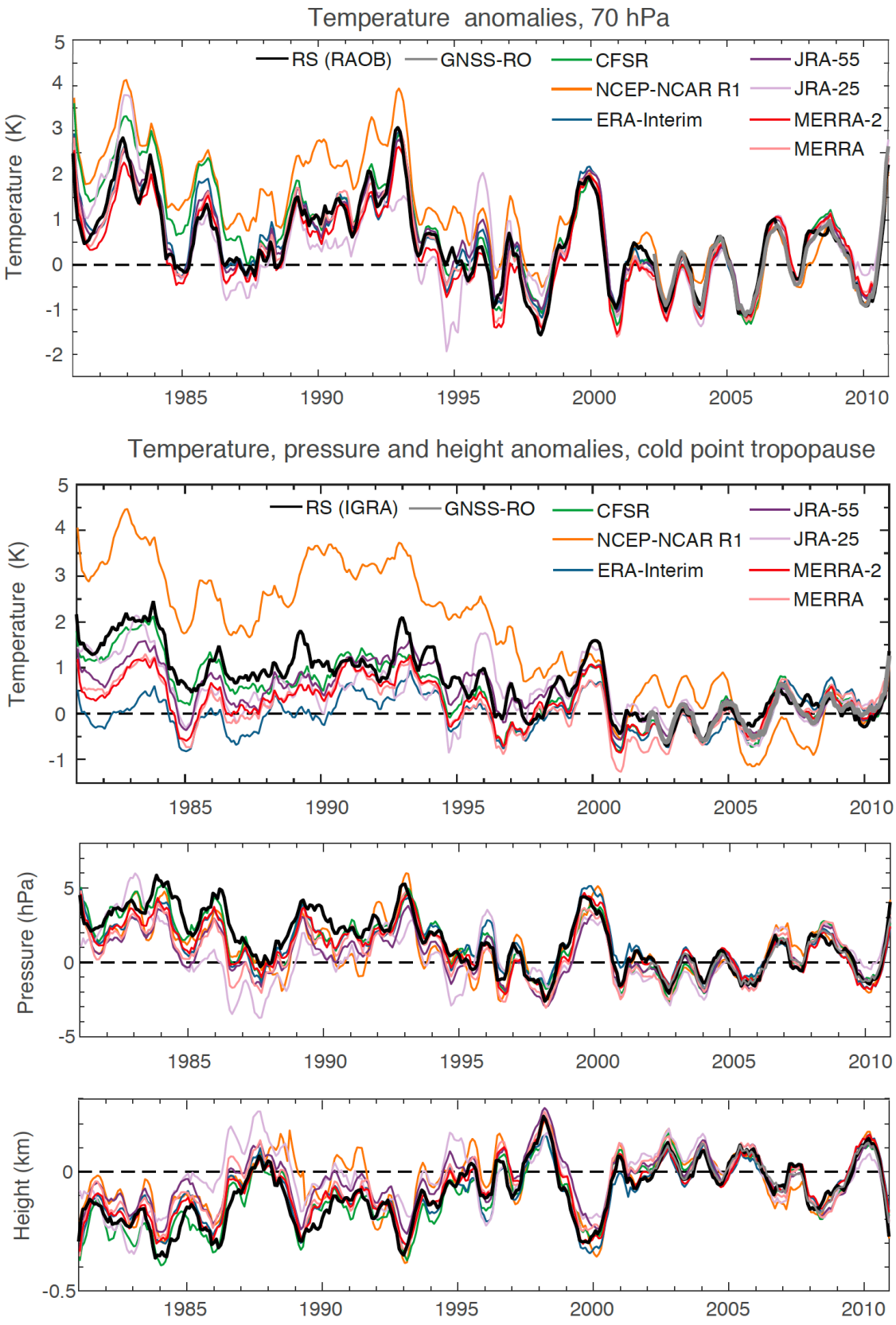

Figure 10. Time series of deseasonalized anomalies in 70 hPa temperature (upper), cold point temperature (upper middle), cold point pressure (lower middle) and cold point altitude (lower) averaged over the tropics (20°S–20°N) and evaluated relative to the reference period 2002–2010. Time series are shown for reanalysis products, radiosonde data (RAOBCORE and IGRA) and radio occultation data (GNSS-RO). Time series are smoothed with a 7-months running mean.

Interannual variability in cold point pressure and altitude **(Fig. 10)** shows better agreement among the data sets than that in cold point or 70 hPa temperature. During the first 15 years of the record, the reanalysis cold point tropopause levels are mostly shifted toward ~~higher~~ **lower** altitudes and ~~lower~~ **higher** pressures, consistent with ~~lower~~ **higher** temperatures during this period. Anomalies in cold point temperature are in most cases matched by anomalies in cold point pressure and altitude, with a warmer cold-point temperature (e.g. around 1999–2000) corresponding to lower tropopause (negative altitude anomaly and positive pressure anomaly) and vice versa. The older reanalyses NCEP-NCAR R1 and JRA-25 again show the largest overall differences. The agreement improves over time, with the most consistent results found for the period after 2002.

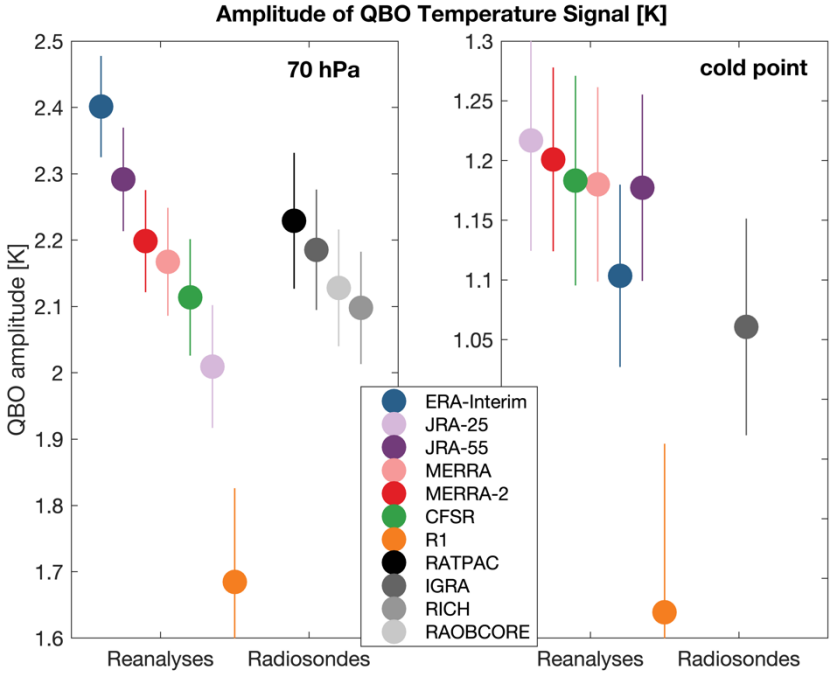

Figure 11. Amplitude of QBO temperature signal for 10°S-10°N at 70 hPa and the cold point derived from a multilinear regression analyses for radiosonde and reanalysis data sets **for the period 1981-2010**.

Long-term temperature changes are evaluated over the 1979-2005 time period due to the availability of adjusted tropopause trends from radiosonde data sets (see Wang et al., 2012 for details). Both radiosonde records suggest significant cooling at the 70 hPa level (**Fig. 12**). Trends derived from reanalysis data can be problematic due to changes in the assimilated observations. Given this potential limitation, it is of interest to examine whether the reanalysis trends are consistent with the hypothetically more reliable trends derived from homogenized observational records. At 70 hPa, temperature trends based on the reanalysis data sets span almost exactly the same range (–0.5 to –1.1 K/decade) as those based on the radiosonde data sets (–0.5 to –1 K/decade). All reanalysis- and observationally-based trends are significant at this level, confirming the stratospheric cooling reported by many previous studies (e.g., Randel

et al., 2009). Satellite data from the Microwave Sounding Unit channel 4 (~13–22 km) suggests smaller trends of around –0.25 K/decade over 1979–2005 (Maycock et al., 2018) or –0.4 K/decade over 1979–2009 (Emanuel et al., 2013). However, the much broader altitude range of this MSU channel includes both stratospheric and tropospheric levels, which impedes a direct comparison with trends at 70 hPa.

At the 100 hPa and cold point levels, the situation is completely different. The available adjusted radiosonde data sets show in some cases uncertainties larger than the respective temperature trends at these levels. Only a few of the available data sets indicate a statistically significant cooling based on a methodology that adjusts the cold point trend to account for nearby fixed pressure-level data and day–night differences (Wang et al., 2012). Based on the trends shown in Wang et al. (2012) for five adjusted radiosonde data sets, we show here the smallest and largest reported trends and consider their range (including the reported error bars) as the observational uncertainty range. Similar to the observations, the reanalysis data sets suggest a large range in cold point temperature trends, from no trend at all (0 K/decade for ERA-Interim) to a strong cooling of –1.3 K/decade (NCEP-NCAR R1). The latter is outside of the observational uncertainty range and can thus be considered unrealistic. All other reanalyses suggest small but significant cooling trends of –0.3 K/decade to –0.6 K/decade. JRA-25, JRA-55, MERRA, and MERRA-2 agree particularly well and produce trends in the middle of the observational uncertainty range. Overall, due to the large uncertainties in radiosonde-derived cold point temperature trends, all reanalyses except for R1 are statistically consistent with at least one of the observational data sets.

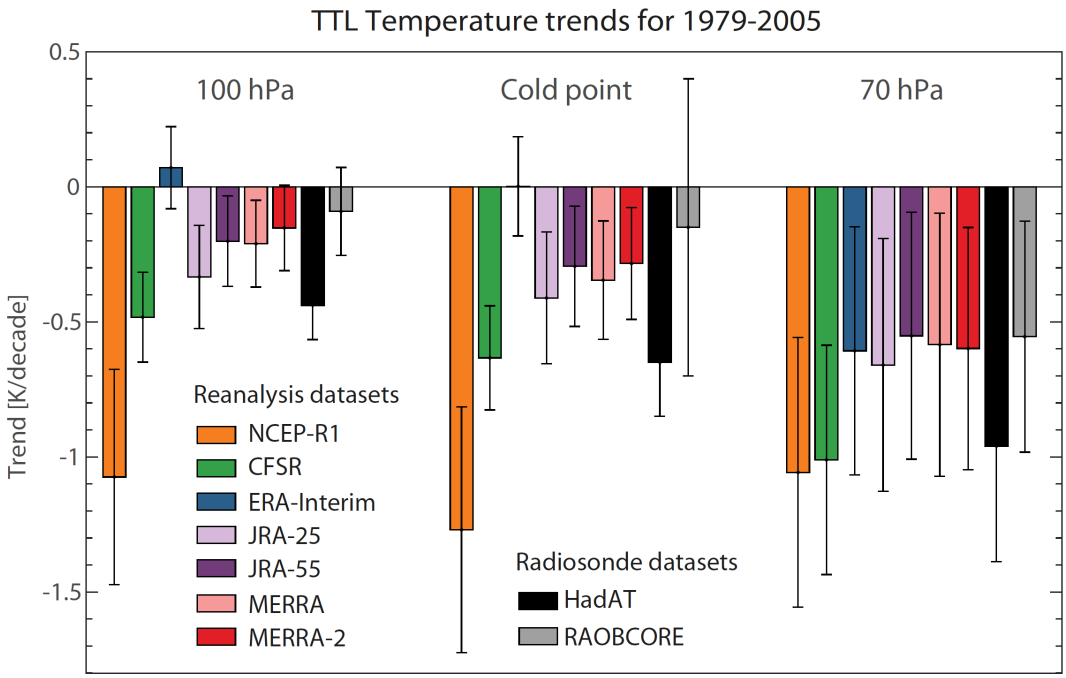

Figure 12. Linear trends in tropical mean (20°S–20°N) temperature (K/decade) at 100 hPa, the cold point and 70 hPa for the time period 1979-2005. Error bars indicate ±2σ uncertainty in the trend and account for serial auto-correlation.

Temperature trends at 100 hPa are very similar to trends at the cold point level, and again suggest consistency among most of the reanalysis and radiosonde data sets, with the notable exception of R1. Nearly all data sets suggest slightly smaller cooling trends (–0.15 K/decade to -0.5 K/decade) relative to the cold point consistent with the fact that the cold point is at slightly higher altitudes than 100 hPa. Among the data sets, only ERA-Interim produces a warming trend (0.07 K/decade), although this result is not statistically significant.

5 Summary

Meteorological reanalyses are widely used in scientific studies of TTL processes being utilizied as ''stand in observations'' or for driving transport models. The most recent atmospheric reanalysis data sets (ERA5, ERA-Interim, MERRA-2, JRA-55, and CFSR) all provide realistic representations of the major characteristics of temperature structure within the TTL **for 2002-2010**. There is good agreement between reanalysis estimates of tropical mean temperatures between 140 and 70 hPa and GNSS-RO retrievals, with relatively small cold biases for most data sets. CFSR shows the best agreement with GNSS-RO in this layer with a mean bias of –0.06 K. Agreement between the temperature profiles and the GNSS-RO data clearly improves when the comparison is restricted to the period after 2007, when the densely-sampled COSMIC data were assimilated by all reanalyses.

Temperatures at the cold point and lapse rate tropopause levels show warm biases in reanalyses when compared to observations. This tropopause-level warm bias is opposite to the cold bias found at all model levels and is most likely related to difficulties in determining the true cold point and lapse rate tropopause levels from discrete temperature profiles with coarse vertical resolution. Our analysis confirms that the magnitude of the bias shift is consistent with the vertical resolution of the reanalysis data, with the smallest bias shifts found for data sets with the highest vertical resolution around the tropopause (ERA5 and CFSR). The negative temperature bias at model levels is often cancelled out by the positive bias introduced when identifying the lapse rate and cold point tropopause locations. As a result, ERA5, which has a small negative bias at model levels **and a small bias shift**, has the most realistic tropopause temperatures, while CFSR, which produces the most realistic model-level temperature profile, has a warm bias of 0.6–0.9 K at the cold point and lapse rate tropopause levels. Older reanalyses like MERRA, JRA-25 and especially NCEP-NCAR R1 show the largest temperature biases at the tropopause levels.

The zonal structure of tropopause temperature reveals that the biases in reanalysis relative to observations maximise at or near the equator. All of the recent reanalyses produce a realistic horizontal structure of cold point temperature with minima corresponding to the centres of tropical deep convection. Differences between reanalyses and observations are greatest over equatorial Africa. These enhanced differences are possibly related to Kelvin wave activity and associated disturbances in TTL temperatures that also maximize in this region. **Further investigation of seasonal variability in the cold point tropopause, including detailed analysis of this feature, will be conducted in a follow-up study.**

Interannual variability in reanalysis temperatures is best constrained in the upper TTL (70 hPa), with larger differences at lower levels such as the cold point and 100 hPa. The reanalyses reproduce the temperature responses to major dynamical and radiative signals such as volcanic

eruptions and the QBO. Agreement among the reanalyses and between the reanalyses and observations generally improves over time, with a step-like improvement around the TOVS-to-ATOVS transition in 1998–1999 and in 2006 with the beginning of assimilation of COSMIC GNSS-RO data. Interannual variability is lower at the cold point and 100 hPa relative to 70 hPa, but with larger month-to-month fluctuations causing larger discrepancies among the reanalyses. As at 70 hPa, NCEP-NCAR R1 is a clear outlier. Interannual variability in cold point pressure and altitude shows better agreement than that in TTL temperature. Anomalies in cold point temperatures are in most cases matched by corresponding anomalies in cold point pressure and altitude.

Long-term reanalysis trends in temperature at 70 hPa show good agreement with trends derived from adjusted radiosonde data sets. All reanalyses and observational data sets indicate significant stratospheric cooling at this level of around –0.5 K/decade to –1 K/decade. At the 100 hPa and cold point levels, both adjusted radiosonde data sets and reanalyses indicate large uncertainties in temperature trends. Reanalysis-based estimates at the cold point range from no trend at all (0 K/decade for ERA-Interim) to strong cooling of –1.3 K/decade (NCEP-NCAR R1). While the latter is outside of the observational uncertainty range and can be considered unrealistic, all other reanalyses data sets agree with at least one of the observational data sets within uncertainties. The bulk of the reanalyses are in good agreement at these levels, suggesting small but significant cooling trends of –0.3 K/decade to –0.6 K/decade that are statistically consistent with trends based on the adjusted radiosonde data sets.

Advances of the reanalysis and observational systems over the last decades have led to a clear improvement of the TTL reanalyses products over time. In particular, the ~~modern~~ **more recent** reanalyses ERA-Interim, ERA5, MERRA-2, CFSR and JRA-55 **mostly** show ~~a~~ very good agreement after 2002 in terms of the vertical TTL temperature profile**,** meridional tropopause structure and interannual variability. Temperatures at the cold point and lapse rate, on the other hand, are too high for most ~~old and modern~~ reanalyses**, regardless of production date**. As these differences maximise over Central Africa, a centre of deep convective activity, chemistry-transport models driven by reanalyses and simulating air mass transport into the stratosphere, can be expected to have too little dehydrations and too high water vapor. **Furthermore, all reanalyses place the cold point tropopause too low in altitude relative to observations. This displacement can have important implications for studies that compare water vapor and ice observations with the position of the cold point tropopause derived from reanalyses data, as enhanced ice and water vapour contents could be erroneously attributed to deep convection crossing the tropopause.**

Depending on the particular application, different reanalyses offer different advantages such as a realistic cold point temperature (e.g., ERA5), small bias in the TTL temperature profile (e.g., CFSR), realistic spatial distribution of the cold point temperature (e.g., ERA-Interim), continuous TTL temperature time series through 2006 (e.g., JRA55), or a realistic representation of signals of interannual variability (e.g., MERRA-2). Their use in model simulations and in comparisons with climate model output should be tailored to their specific strengths and weaknesses.

Author contributions. ST developed the idea for this paper and carried out the evaluations with contributions from all co-authors. SD and BL provided the reanalyses tropopause and profile data. RPK provided the GNSS-RO tropopause, wave activity and temperature profile data. J.S. Wang provided the radiosonde tropopause data. ST wrote the manuscript with contributions from all co-authors.

Acknowledgements:

We thank the reanalysis centres for providing their support and data products. We thank C. Bloecker from the Global Modeling and Assimilation Office, NASA Goddard Space Flight Center for providing information on the GNSS-RO data assimilated in MERRA-2. ERA5 data were generated using Copernicus Climate Change Service Information. **MERRA-2 data access was through the Global Modeling and Assimilation Office (GMAO, 2015).** The work of S. Tegtmeier was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – TE 1134/1. Contributions from J.S. Wright were supported by the National Natural Science Foundation of China (20171352419) via a joint DFG–NSFC funding initiative.

References

Anthes, R. A., Bernhardt, P. A., Chen, Y., Cucurull, L., Dymond, K. F., Ector, D., Healy, S. B., Ho, S.-P., Hunt, D. C., Kuo, Y.-H., Liu, H., Manning, K., McCormick, C., Meehan, T. K., Randel, W. J., Rocken, C., Schreiner, W. S., Sokolovskiy, S. V., Syndergaard, S., Thompson, D. C., Trenberth, K. E., Wee, T.-K., Yen, N. L., and Zeng, Z.: The COSMIC/FORMOSAT-3 Mission: Early Results, Bulletin of the American Meteorological Society, 89, 313, https://doi.org/10.1175/BAMS-89-3-313, 2008.

Beyerle, G., Schmidt, T., Michalak, G., Heise, S., Wickert, J., and Reigber, C.: GPS radio occultation with GRACE: Atmospheric profiling utilizing the zero difference technique, Geophysical Research Letters, 32, L13806, https://doi.org/10.1029/2005GL023109, 2005.

Beyerle, G., Grunwaldt, L., Heise, S., Köhler, W., König, R., Michalak, G., Rothacher, M., Schmidt, T., Wickert, J., Tapley, B. D., and Giesinger, B.: First results from the GPS atmosphere sounding experiment TOR aboard the TerraSAR-X satellite, Atmospheric Chemistry & Physics, 11, 6687–6699, https://doi.org/10.5194/acp-11-6687-2011, 2011.

Chipperfield, M. P., Multiannual simulations with a three-dimensional chemical transport model, J. Geophys. Res., 104(D1), 1781– 1805, doi:10.1029/98JD02597, 1999.

Cucurull, L., J. C. Derber, and R. J. Purser: A bending angle forward operator for global positioning system radio occultation measurements. Journal of Geophysical Research: Atmospheres, 118, 14-28. doi: 10.1029/2012JD017782, 2013.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteorol. Soc., 137: 553–597. doi: 10.1002/qj.828, 2011.

Durre, I., R. S. Vose, and D. B. Wuertz, Overview of the Integrated Global Radiosonde Archive, J. Clim., 19(1), 53–68, doi:10.1175/JCLI3594.1, 2006.

Efron, B., and R. J. Tibshirani, An Introduction to the Bootstrap, 436 pp., Chapman and Hall, New York, 1993.

Emanuel, K., S. Solomon, D. Folini, S. Davis, and C. Cagnazzo, Influence of Tropical Tropopause Layer Cooling on Atlantic Hurricane Activity. J. Climate, 26, 2288–2301, https://doi.org/10.1175/JCLI-D-12-00242.1, 2013.

Folkins, I., M. Lowewenstein, J. Podolske, S. Oltmans, and M. Proffitt, A barrier to vertical mixing at 14 km in the tropics: Evidence from ozonesondes and aircraft measurements, J. Geophys. Res.,104, 22095-22102, 1999.

Free, M., Seidel, D. J., Angell, J. K., Lanzante, J., Durre, I., and Peterson, T. C., Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new data set of large-area anomaly time series, J. Geophys. Res., 110, D22101, doi:10.1029/2005JD006169, 2005

Fueglistaler S, Dessler A E, Dunkerton T J, et al., Tropical tropopause layer, Rev. Geophys., 47, RG1004, doi:10.1029/2008RG000267, 2009.

Fueglistaler, S., Haynes, P. H., and Forster, P. M.: The annual cycle in lower stratospheric temperatures revisited, Atmos. Chem. Phys., 11, 3701-3711, https://doi.org/10.5194/acp-11-3701-2011, 2011.

Fujiwara, M., Suzuki, J., Gettelman, A., Hegglin, M. I., Akiyoshi, H., and Shibata, K., Wave activity in the tropical tropopause layer in seven reanalysis and four chemistry climate model data sets, J. Geophys. Res., 117, D12105, doi:10.1029/2011JD016808, 2012.

Fujiwara, M., Hibino, T., Mehta, S. K., Gray, L., Mitchell, D., and Anstey, J.: Global temperature response to the major volcanic eruptions in multiple reanalysis data sets, Atmos. Chem. Phys., 15, 13507-13518, https://doi.org/10.5194/acp-15-13507-2015, 2015.

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, Atmos. Chem. Phys., 17, 1417-1452, https://doi.org/10.5194/acp-17-1417-2017, 2017.

Gelaro, R., W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, and B. Zhao: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Gettelman, A., M. L. Salby, and F. Sassi: Distribution and influence of convection in the tropical tropopause region. J. Geophys. Res., 107, 4080, doi:10.1029/2001JD001048, 2002.

Gettelman, A., et al., Multimodel assessment of the upper troposphere and lower stratosphere: Tropics and global trends, J. Geophys. Res., 115, D00M08, doi:10.1029/2009JD013638, 2010.

Global Modeling and Assimilation Office (GMAO), MERRA-2 inst6_3d_ana_Nv: 3d,6-Hourly, Instantaneous, Model-Level, Analysis, Analyzed Meteorological Fields V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [1.1.2015], 10.5067/IUUF4WB9FT4W, 2015.

Gorbunov, M. E., Benzon, H.-H., Jensen, A. S., Lohmann, M. S., and Nielsen, A. S.: Comparative analysis of radio occultation processing approaches based on Fourier integral operators, Radio Science, 39, RS6004, https://doi.org/10.1029/2003RS002916, 2004.

Haimberger, L., Homogenization of Radiosonde Temperature Time Series Using Innovation Statistics. J. Climate, 20, 1377–1403, https://doi.org/10.1175/JCLI4050.1, 2007.

Haimberger, L., Tavolato, C., and Sperka, S.: Homogenization of the Global Radiosonde Temperature Dataset through Combined Comparison with Reanalysis Background Series and Neighboring Stations, J. Climate, 25, 8108–3131, https://doi.org/10.1175/JCLI-D-11-00668.1, 2012.

Hajj, G. A., Ao, C. O., Iijima, B. A., Kuang, D., Kursinski, E. R., Mannucci, A. J., Meehan, T. K., Romans, L. J., de la Torre Juarez, M., and Yunck, T. P.: CHAMP and SAC-C atmospheric occultation

results and intercomparisons, Journal of Geophysical Research (Atmospheres), 109, D06 109, https://doi.org/10.1029/2003JD003909, 2004.

Hartmann, D. L., J. R. Holton, and Q. Fu: The heat balance of the tropical tropopause, cirrus, and stratospheric dehydration. Geophys. Res. Lett., 28, 1969–1972, 2001

Hersbach, H., et al., Operational global reanalysis: progress, future directions and synergies with NWP, ERA Report Series, 27, 2018.

Highwood, E. J. and Hoskins, B. J., The tropical tropopause. Q.J.R. Meteorol. Soc., 124: 1579-1604. doi:10.1002/qj.49712454911, 1998.

Ho, S.-p., Peng, L., and Vömel, H.: Characterization of the longterm radiosonde temperature biases in the upper troposphere and lower stratosphere using COSMIC and Metop-A/GRAS data from 2006 to 2014, Atmospheric Chemistry & Physics, 17, 4493–4511, https://doi.org/10.5194/acp-17-4493-2017, 2017.

Holton, J. R. and Gettelman, A.: Horizontal transport and the dehydration of the stratosphere, Geophys. Res. Lett., 28, 2799–2802, 2001.

Kim, J. and S. Son: Tropical Cold-Point Tropopause: Climatology, Seasonal Cycle, and Intraseasonal Variability Derived from COSMIC GPS Radio Occultation Measurements. J. Climate, 25, 5343–5360, https://doi.org/10.1175/JCLI-D-11-00554.1, 2012.

Kim, Y.-H., Kiladis, G. N., Albers, J. R., Dias, J., Fujiwara, M., Anstey, J. A., Song, I.-S., Wright, C. J., Kawatani, Y., Lott, F., and Yoo, C.: Comparison of equatorial wave activity in the tropical tropopause layer and stratosphere represented in reanalyses, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-110, in review, 2019.

Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP–NCAR 50-year reanalysis: monthly means CD-ROM and documentation, B. Am. Meteorol. Soc., 82, 247–267, 2001.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 reanalysis: general specifications and basic characteristics, J. Meteorol. Soc. Jpn., 93, 5–48, doi:10.2151/jmsj.2015-001, 2015.

Krüger, K., Tegtmeier, S., and Rex, M.: Long-term climatology of air mass transport through the Tropical Tropopause Layer (TTL) during NH winter, Atmos. Chem. Phys., 8, 813-823, https://doi.org/10.5194/acp-8-813-2008, 2008.

Krüger, K., Tegtmeier, S., and Rex, M.: Variability of residence time in the Tropical Tropopause Layer during Northern Hemisphere winter, Atmos. Chem. Phys., 9, 6717-6725, https://doi.org/10.5194/acp-9-6717-2009, 2009.

Kursinski, E. R., Hajj, G. A., Schofield, J. T., Linfield, R. P., and Hardy, K. R.: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System, Journal of Geophysical Research: Atmospheres, 102, 23 429–23 465, https://doi.org/10.1029/97JD01569, 1997.

Lawrence, Z. D., Manney, G. L., and Wargan, K.: Reanalysis intercomparisons of stratospheric polar processing diagnostics, Atmos. Chem. Phys., 18, 13547-13579, https://doi.org/10.5194/acp-18-13547-2018, 2018.

Long, C. S., Fujiwara, M., Davis, S., Mitchell, D. M., and Wright, C. J.: Climatology and interannual variability of dynamic variables in multiple reanalyses evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP), Atmos. Chem. Phys., 17, 14593-14629, https://doi.org/10.5194/acp-17-14593-2017, 2017.

Maycock, A.C., et al., Revisiting the mystery of stratospheric temperature trends. Geophys. Res. Lett., doi:10.1029/2018GL078035, 2018.

Mote, P. W., Rosenlof, K. H., McIntyre, M. E., Carr, E. S., Gille, J. C., Holton, J. R., Kinnersley, J. S., Pumphrey, H. C., Russell, J. M., and Waters, J. W.: An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor, J. Geophys. Res., 101(D2), 3989– 4006, doi:10.1029/95JD03422, 1996.

Nash, J., Oakley, T., Vömel, H., and Li, W.: WMO intercomparison of high-quality radiosonde systems, Yangjiang, China, 12 July–3 August 2010, Instruments and Observing Methods Re- port No. 107, WMO/TD-No. 1580, WMO, Geneva, Switzerland, 238 pp., 2011.

Naujokat, B., An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics, J. Atmos. Sci., 43, 1873-1877, 1986.

Nishimoto, E., and M. Shiotani, Seasonal and interannual variability in the temperature structure around the tropical tropopause and its relationship with convective activities, J. Geophys. Res. Atmos., 117, D02104, doi:10.1029/2011JD016936, 2012.

Nishimoto, E., and M. Shiotani, Intraseasonal variations in the tropical tropopause temperature revealed by cluster analysis of convective activity, J. Geophys. Res. Atmos., 118, 3545– 3556, doi: 10.1002/jgrd.50281, 2013.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hat- sushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Taka- hashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Taira, R.: The JRA-25 reanalysis, J. Meteorol. Soc. Jpn., 85, 369–432, doi:10.2151/jmsj.85.369, 2007.

Pan, L. L., and Munchak, L. A.: Relationship of cloud top to the tropopause and jet structure from CALIPSO data, J. Geophys. Res., 116, D12201, doi:10.1029/2010JD015462, 2011.

Pan, L. L., Honomichl, S. B., Bui, T. V., Thornberry, T., Rollins, A., Hintsa, E., & Jensen, E. J.: Lapse Rate or Cold Point: The Tropical Tropopause Identified by In Situ Trace Gas Measurements. Geophysical Research Letters, 45(19), 10-756, 2018.

Randel W. J., and Jensen, E., Physical processes in the tropical tropopause layer and their roles in a changing climate. Nat Geosci 6:169–176, 2013.

Randel, W. J., F. Wu, and D. J. Gaffen, Low frequency variations of the tropical tropopause from NCEP reanalyses. J. Geophys. Res., 105, 15 509–15 523, 2000.

Randel, W. J., F. Wu, S.J. Oltmans, K. Rosenlof, and G.E. Nedoluha, Interannual Changes of Stratospheric Water Vapor and Correlations with Tropical Tropopause Temperatures. J. Atmos. Sci., 61, 2133–2148, https://doi.org/10.1175/1520-0469, 2004a.

Randel, W. J., P. Udelhofen, E. Fleming, M. Geller, M. Gelman, K. Hamilton, D. Karoly, D. Ortland, S. Pawson, R. Swinbank, F. Wu, M. Baldwin, M. Chanin, P. Keckhut, K. Labitzke, E. Remsberg, A.

Simmons, and D. Wu, The SPARC Intercomparison of Middle-Atmosphere Climatologies. J. Climate, 17, 986–1003, https://doi.org/10.1175/1520-0442(2004)017, 2004b.

Randel, W. J., Shine, K. P., Austin, J., Barnett, J., Claud, C., Gillett, N. P., et al., An update of observed stratospheric temperature trends. Journal of Geophysical Research, 114, D02107. https://doi.org/10.1029/2008JD010421, 2009.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeis- ter, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective anal- ysis for research and applications, J. Climate, 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.- Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, B. Am. Meteorol. Soc., 91, 1015–1057, doi:10.1175/2010BAMS3001.1, 2010.

Santer, B.D., R. Sausen, T.M.L. Wigley, J.S. Boyle, K. AchutaRao, C. Doutriaux, J.E. Hansen, G.A. Meehl, E. Roeckner, R. Ruedy, G. Schmidt, and K.E. Taylor, Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes. J. Geophys. Res., 108, no. D1, 4002, doi:10.1029/2002JD002258, 2003.

Schoeberl, M. R., A. E. Dessler, and T. Wang., Simulation of stratospheric water vapor and trends using three reanalyses." Atmospheric Chemistry and Physics, 12 (14): 6475-6487 doi:10.5194/acp-12-6475-2012, 2012.

Seidel, D. J., and Randel, W. J., Variability and trends in the global tropopause estimated from radiosonde data, J. Geophys. Res., 111, D21101, doi:10.1029/2006JD007363, 2006

Seidel, D. J., R. J. Ross, J. K. Angell, and G. C. Reid, Climatological characteristics of the tropical tropopause as revealed by radiosondes, J. Geophys. Res., 106(D8), 7857–7878, doi:10.1029/2000JD900837, 2001.

Sherwood, S. C., T. Horinouchi, and H. A. Zeleznik: Convective impact on temperatures observed near the tropical tropopause. J. Atmos. Sci., 60, 1847–1856, 2003.

Simmons, A. J., Poli, P., Dee, D. P., Berrisford, P., Hersbach, H., Kobayashi, S. and Peubey, C., Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. Q.J.R. Meteorol. Soc., 140: 329-353. doi:10.1002/qj.2317, 2014.

Tao, M., Konopka, P., Ploeger, F., Yan, X., Wright, J. S., Diallo, M., Fueglistaler, S., and Riese, M.: Multitimescale variations in modeled stratospheric water vapor derived from three modern reanalysis products, Atmos. Chem. Phys., 19, 6509–6534, https://doi.org/10.5194/acp-19-6509-2019, 2019.

Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, Revisiting radiosonde upper air temperatures from 1958 to 2002, J. Geophys. Res., 110, D18105, doi:10.1029/2004JD005753, 2005.

Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., and Peter, T.: A global space-based stratospheric aerosol climatology: 1979–2016, Earth Syst. Sci. Data, 10, 469-492, https://doi.org/10.5194/essd-10-469-2018, 2018.

von Engeln, A., Andres, Y., Marquardt, C., and Sancho, F.: GRAS radio occultation on-board of Metop, Advances in Space Research, 47, 336–347, https://doi.org/10.1016/j.asr.2010.07.028, 2011.

Wang, J. S., Seidel, D. J., and Free, M., How well do we know recent climate trends at the tropical tropopause?, J. Geophys. Res., 117, D09118, doi:10.1029/2012JD017444, 2012.

Wheeler, M. and G.N. Kiladis, Convectively Coupled Equatorial Waves: Analysis of Clouds and Temperature in the Wavenumber–Frequency Domain. J. Atmos. Sci., 56, 374–399, https://doi.org/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2, 1999.

Wickert, J., Reigber, C., Beyerle, G., König, R., Marquardt, C., Schmidt, T., Grunwaldt, L., Galas, R., Meehan, T. K., Melbourne, W. G., and Hocke, K.: Atmosphere sounding by GPS radio occultation: First results from CHAMP, Geophysical Research Letters, 28, 3263–3266, https://doi.org/10.1029/2001GL013117, 2001.

World Meteorological Organization: Definition of the tropopause, Bulletin of the World Meteorological Organization, 6, 136–137, 1957.

Wright, C. J. and Hindley, N. P.: How well do stratospheric reanalyses reproduce high-resolution satellite temperature measurements?, Atmos. Chem. Phys., 18, 13703-13731, https://doi.org/10.5194/acp-18-13703-2018, 2018.

Xie, F., Li, J., Tian, W., Li, Y., & Feng, J., Indo-Pacific warm pool area expansion, Modoki activity, and tropical cold-point tropopause temperature variations. *Scientific reports*, *4*, 4552. doi:10.1038/srep04552, 2014.

Zhou, X.L., M.A. Geller, and M.H. Zhang, Tropical Cold Point Tropopause Characteristics Derived from ECMWF Reanalyses and Soundings. J. Climate, 14, 1823–1838, https://doi.org/10.1175/1520-0442, 2001.