Reply to anonymous Referee #1 (acpd-14-C11680–C11684)

This study compares trajectories in the TTL calculated with three different temperature datasets. The focus of the evaluation is on the statistical evaluation of the minimum temperature along the trajectories as they cross the tropical tropopause, and the corresponding water vapour entering the stratosphere. It is shown that the overall humidity values and in particular the seasonal cycle and interannual variability are only very weakly sensitive to the choice of the temperature dataset. The objective of the study is well justified and the results are in principle relevant; however, important aspects of the paper are not well explained and/or conceptually fuzzy - as outlined in my comments below. Therefore major revisions are required to turn this study into a fully consistent and convincing paper.

Reply:

Thanks for those helpful comments. We have made substantial changes to the manuscript to include answers to all aspects. The detailed answers for each question can be found below, with line numbers from the updated manuscript.

Major comments:

A) *Is this paper really about the impact of "temperature resolution"? First, the study only considers the vertical resolution aspect, and not the horizontal nor the temporal one. In particular temporal resolution might also matter, but this is not mentioned in the paper. Then, "resolution" to me sounds very technical (e.g., like running a model with two different resolutions). But this is not exactly the problem, nor what you do. My point is that if the MERRA assimilation cycle was running with a model with higher vertical resolution, then the resulting field would not necessarily capture more of the, e.g., gravity wave signals because capturing them is not only an issue of resolution but also of the representation of the wave triggering mechanisms. I would like the authors to discuss more critically and explicitly what they actually do. I think it is good and relevant, but it is not well described by "resolution". Maybe then the authors might also consider rephrasing the title of their study.*

Reply:

This is an excellent point. We intend to investigate the impacts of vertical variability of tropopause temperatures on trajectory modeling of water vapor. It is known that local tropopause temperature could experience much variability in the vertical, so the real question is that if the temperature datasets we used already captured enough variability in tropopause temperature. If not, how big impact could it be? In the updated manuscript, we have changed the title to "… temperature vertical variability…" to make the objective of this paper clearer.

B) *Three datasets are used and I am perfectly fine with the first two of them: MERRA is used because it is "standard", available for a long time period etc. GPS is used* *because it is based upon independent observations with fine vertical resolution. My expectation is that this data set should be as close to reality as possible. But then why use the synthetically created MERRA-Twave dataset? I understand that this dataset would be valuable if we did not have GPS. But since we have GPS what can this dataset tell us in addition? I suggest to better motivate the use of this third dataset, or to focus on the analysis with MERRA and GPS only.*

Reply:

We included both GPS and MER-Twave datasets because they have their own advantages and limitations. GPS provides sparse sampling in the tropics (only ~800-1100 profiles per day), which means the variability in GPS is smaller than reality, although its mean is more accurate given the precise profiling. In contrast, MER-Twave has better variability but not accurate mean, since it is designed to have similar temperature variability to radiosondes but with mean reserved to original MERRA data (Kim and Alexdander, 2013). In summary, the mean temperature is closer to reality in GPS than in MER-T and MER-Twave, but the temperature variability is closer to reality in MER-Twave than in MER-T and GPS. We have added this discussion in lines 240-249.

C) *The discussion of the impact of atmospheric waves is insufficient. The general statement at the beginning of section 2.2.2 "Waves are underrepresented in reanalyses" does not make sense. Clearly planetary and synoptic-scale waves are/should be perfectly captured by reanalyses. It remains unclear, which part of the wave spectrum is considered here. Kelvin waves, gravity waves? When discussing the role of, e.g., gravity waves on the temperature field in the TTL, then maybe also the temporal resolution should be discussed. Six-hourly fields, from MERRA or GPS, cannot capture the temporal propagation / evolution of waves. This could potentially also affect the minimum temperature along the trajectories.*

Reply:

The description of underrepresented waves can be found in Kim and Alexander, 2013. According to their results (their Fig. 1b-d), at reanalysis model levels temperature variability at time scales shorter than ~10 days are weaker than observations. Thus, underrepresented waves include a part of Kelvin waves, mixed Rossby-gravity waves, and gravity waves. However, the problem in using reanalysis data for trajectory simulations is associated not only with these waves (< 10 days), but also with slow-scale waves (>10 days), since it involves interpolation between reanalysis vertical levels. As shown in Kim and Alexander, 2013, conventional interpolation (either linear or higher order) in-between model vertical levels degrades temperature variability even at longer time scales (> 10 days). This is because observed temperature profiles have strong curvatures in-between coarse model levels due to the existence of fine vertical-scale waves.

We only considered a vertical resolution issue, since horizontal or temporal resolution of current reanalyses is good enough to resolve most of TTL waves (Note that we do not mean that horizontal and temporal resolutions are good enough to resolve wave generation mechanisms.). A large portion of TTL waves has horizontal and temporal scales much larger and longer than reanalysis resolution, therefore, temperature behaves almost linearly in-between model horizontal and temporal resolution. However, temperature does not behave linearly in vertical space due to the fact that a significant portion of TTL waves have vertical wavelengths shorter than ~4 km (see Figure S4 in supporting information of Kim and Alexander, 2015), which could make waves less represented by the ~1.2 km vertical resolution in reanalyses.

The above discussion has been included in section 2.2.2.

D) *The paper has not been very carefully written. Several sentences/formulations are unclear:*

- *p. 29210 line 11: what is meant by "finite resolution"? Every resolution is finite, do you mean "fine"? (This problem occurs in several places.)*

Reply:

Yes, we mean fine resolution. All "finite" have been corrected to "fine".

- *p. 29213 line 15: "the carrying methane" sounds odd to me. Not clear how the methane values are initialized in the trajectories.*

Reply:

Corrected. We have modified in lines 122-126 to include the whole story of methane carried in our model.

- *p. 29213 line 19: what is meant by "limited in the tropical 110-50 hPa"?*

Reply:

Tropical 110-50 hPa is where the most dehydration happens. Refer Fig. 5a-c.

- *p. 29213 line 28: "total diabatic heating rates from all sky": please explain this better.*

Reply:

It means total heating rates due to long-wave and short-wave radiation, moist physics, friction, etc. It has been modified in lines 95-97.

- *p. 29214 line 8: "not represented well in current coarse model levels": you probably mean "... in models with coarse vertical resolution"; but I think this is not really the point (see comment A): even with more levels MERRA would not correctly capture all gravity waves emitted from tropical convection.*

Reply:

Agree. See reply to question C and more detailed explanation in section 2.2.2.

- *p. 29217 line 5: what is "the curly nature" of a temperature profile?*

Reply:

The "curly nature" means the strong curvature of temperature profiles around the cold-point tropopause. This has been rephrased in line 253-255.

- *p. 29220 line 4: "We see slightly drier air in GPS run expected"?*

Reply:

Rephrased.

- *p. 29220 line 20: The sentence "Note that ..." is too long, and it is not clear what is meant by "the two are strongly coupled".*

Reply:

Rephrased. See lines 367-369.

E) p. 29215 lines 24: This is an interesting result, but it is not well discussed. How can these happen? How can MERRA be too cold at model levels (compared to GPS) but too warm in between? MERRA values in between model levels are calculated by linear interpolation and therefore I would expect that a cold bias at the model levels is "transferred" to the layer in between.

Reply:

MERRA doesn't assimilate GPS observations, which makes these two datasets independent from each other. Within the tropopause MERRA model levels are separated ~1.2 km apart, which might miss the temperature variations that could only be captured by data in finer vertical resolutions, such as GPS observations. Therefore, although MERRA is warmer at model levels, it doesn't necessarily mean MERRA should be warmer in-between. This is clearly shown in Fig. R11a-b below.

Moreover, the mean temperature differences depend on the location being examined. For example, if we only consider the deep inner tropics (10^{o} N-S), MERRA shows warm biases throughout the entire tropopause layers (Fig. R11c). Either way, a clear fact is that MERRA is warm biased at the cold-point tropopause (~100-90 hPa).

This discussion has been added in the discussions of Fig. 2.

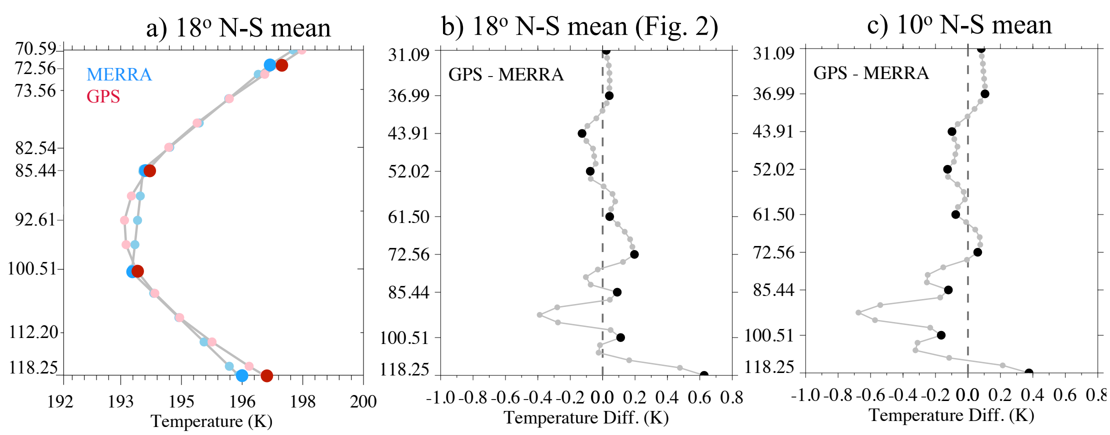

Figure R11. a) MERRA and GPS temperature averaged within 18^{o} N-S in 2007-2013; b) the differences in a); and c) same as b) but averaged within 10^{o} N-S. Clearly average within different latitudes results in different values, but the warm bias in MERRA cold-point tropopause always exists.

F) *In section 3.1 I have a problem in understanding the selection of the trajectories. My impression is that trajectories are selected if they reach the 90-hPa level (this is considered as the entry point in the stratopshere). This is fine with me, but this implies that (during the time period considered) some trajectories maybe don't reach the 60- hPa level. But then you determine FDP statistics up to 60 hPa! Does this not lead to a biased distribution? Should you not select trajectories that reach 60 hPa instead of 90 hPa?*

Reply:

Our trajectory model runs forward, and along time we kept records of any dehydration occurrences. Starting from the initiation level 370-K, parcels ascend to higher altitudes while crossing the tropopause, during which parcels experience multiple dehydrations whenever colder temperatures were encountered. On the other hand, parcel's water vapor is conserved when encountering warmer temperature. To isolate the FDP events, we chose parcels that were already above 90-hPa for at least six months since the last time they were dehydrated (FDP). This guarantees that parcels already crossed the cold-point tropopause (~380 K or ~100-94 hPa) and experienced their final dehydration. This part has been modified accordingly in section 3.1, lines 276-283.

G) *p. 29218 line 19: The bimodal FDP distribution with MERRA data is interesting (Fig. 5). But in principle the distribution should be even more peaked! When using linear interpolation between model levels, then minimum temperature must occur exactly at one of the model levels. So the smearing out of the two peaks is an effect of the temporal resolution of the trajectory output. I assume that you determine the minimum temperature from 6-hourly values along the trajectories. Then of course it can happen that the time when the trajectory reaches the exact pressure of a model level is "hidden" (i.e., in between two times) and therefore the "real" location and value of the minimum temperature is missed. This indicates that the temporal resolution can play an important role, and I suggest that the role of temporal resolution (of the wind fields, of the trajectory output) is discussed in the paper.*

Reply:

This is an excellent point. Along the trajectory integration, FDP is where the coldest temperature is encountered along a parcel's path. This coldest temperature could be found either exactly at MERRA model levels or in-between levels during that step of integration, depending on the trajectory integration intervals. As shown in Fig. R12 below: during two steps of integration (from $t \rightarrow t+\Delta t$, and from $t+\Delta t \rightarrow t+2\Delta t$), the FDP could be found exactly at (Fig. R12a), above (Fig. R12b), or below (Fig. R12c) the MERRA cold-point level (85.4 hPa). Suppose our trajectory integration interval is as small as seconds then at some time steps parcels would inevitably travel to each of the MERRA model levels, and therefore the encountered coldest temperatures would always be at either of the two model levels in MERRA. In another word, the bimodal FDP distribution from MERRA run could be even more peaked when choosing smaller integration step. Two reasons that we didn't choose such small time step: 1) the wind and temperature data are only available 6-hourly or even daily (GPS), so much smaller time step introduces more uncertainties with more interpolation; 2) considering the balance between model running speed and computational resources. This has been addressed in context lines 312-322.

Currently we output trajectory results on daily basis, which is already fine enough to study the evolutions of FDP on monthly or seasonal basis. Besides, due to the domain-filling feature of our model, the FDP results are not sensitive to longer, such as 3-day or 5-day, or shorter, such as hourly, output intervals.

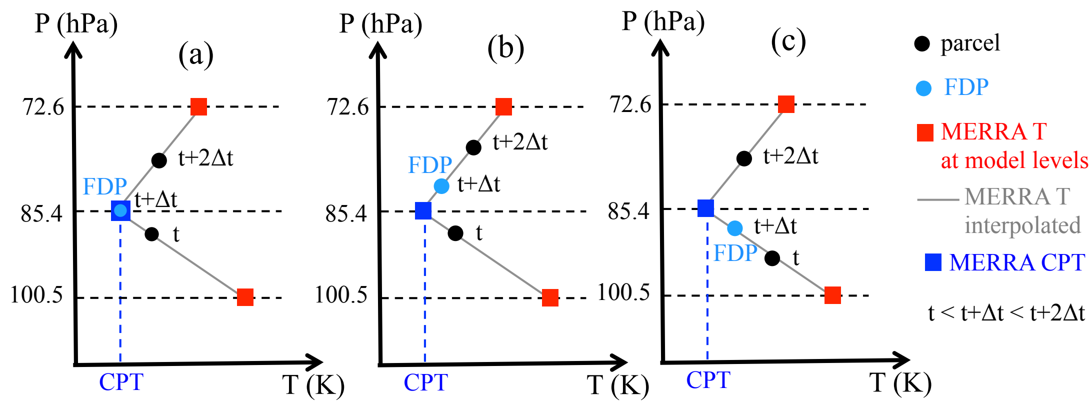

Figure R12. Illustration of the FDP locations in different scenario. Filled squares are MERRA temperatures at model levels, with cold-point tropopause (CPT) marked in blue and others in red. Grey lines are linearly interpolated temperatures in-between model levels. Parcels (black dots) travel from t, to t+Δt, and then to t+2Δt. During this process, FDP (blue dots) could be found exactly at MERRA model levels (a) or in-between MERRA model levels (b, c).

Minor comments:

1) *p. 29212 line 8: maybe a terminology detail: here you write about "resolved but underrepresented waves" - does this (see comment A) also indicate that your study is not mainly about resolution, but more about "effects of gravity(?) waves on the temperature field"?*

   Reply:

   We realized that "resolution" is not appropriate in expressing our objective, so we changed it to "vertical variability". The MER-Twave has more variability than standard MERRA temperatures.

2) *Section 2.2.2 is very difficult to understand. If you keep this MERRA-Twave dataset in your study, this paragraph should become less technical (for the technical aspects the reader can be referred to Kim and Alexander 2013). Here the reader should be able to learn the general concept.*

   Reply:

   Agree. Now we have shortened the technical explanations and replaced with more discussions of waves and temperature variability in section 2.2.2.

3) *Comparing Figs. 5 and 7b: something is probably not correct with the scales of the FDP events. Values in Fig. 7b are about 4 times smaller, but in both cases they should integrate to 100%.*

   Reply:

   For both Fig. 5 and Fig. 7b, the FDP occurrence frequencies are calculated as the ratio of FDP events at each 2-hPa bin relative to total FDP events, regardless of seasons, within the 110-60 hPa range. Therefore, the curves in Fig. 5 represent the mean FDP frequencies averages in all seasons, and the integration of each curve is 100% before being normalized to "%/hPa" (i.e., frequencies divided by 2-hPa). Fig. 7b, however, only shows the frequencies of FDP during SON relative to all season FDP, therefore its magnitude is about ¼ of total frequencies.

   In the updated manuscript, we have changed all normalized FDP frequency unit in Fig. 5 from "%/hPa" to "%", so each PDF profile integrate from bottom to up ends up with 100%.

Editorial comments:

- p. 29213 line 19: "Noted" should read "Note"

Reply:

Corrected.

- p. 29214 line 9: should read "... that use an idealized parameterization of ..."

Reply:

Corrected.

[Reference]

Kim, J.-E., and Alexander, J. M.,: A new wave scheme for trajectory simulations of stratospheric water vapor, Geophys. Res. Lett., 40, 5286–5290, doi:10.1002/grl.50963, 2013.

Kim, J.-E., and Alexander, J. M.,: Direct impacts of waves on tropical cold point tropopause temperature, Geophys. Res. Lett., doi:10.1002/2014GL062737, 2015.