Response to Review Comment 1, acp-2021-154

General: The authors have chosen a topic, modification of the thermodynamic structure of the tropopause by organized convective systems, that is both timely and interesting. This is an important topic because an accurate understanding of diabatic convective processes at the tropopause in the tropics and subtropics could impact the modeling of climate response to increased convection caused by rising surface temperatures. The analysis uses retrievals of temperature and water vapor profiles from AIRS, combined with a Radar-Lidar estimation of cloud ice water content to address an old and on-going debate about whether overshooting convective plumes can hydrate the lowermost stratosphere. It is refreshing that the authors have provided an alternative to the very coarse-resolution MLS water vapor profiles to address this question, with the strong benefit of having a co-located retrieved temperature profile. The cyclone-centered coordinates are a welcome way to organize the observations. I am recommending that this paper be accepted with some minor changes that are listed below.

Thank you for your general assessments of this manuscript and constructive comments that help us improve. Following your suggestions, we reorganized the figures, improved the Section 2.1 by adding detailed introduction to each instrument and dataset, and included comparisons of MLS v5 and the synergistic retrieval in the Conclusion section.

In addition, vertical resolution around the tropopause for each data set/retrieval needs to be stated explicitly,
Done.

some references ought to be consulted and updated to reflect the most current thinking and the development/uncertainty of satellite algorithms, for DARDAR-cloud and for the brightness temperatures.

Done. Following this comment, Section 2.1 has been improved and expanded.

Uncertainties and potential sampling biases need to be discussed.

We account for uncertainties caused by sampling differences mainly by increasing the uncertainty range estimated by different products in the optimal estimation method. It is detailly discussed in the Appendix. Following this comment, we also add relevant information to Section 2.1.

Are the retrievals representative?

The retrievals are representative of atmospheric states above thick high clouds surrounding tropical cyclone centers and the composite is representative of each cloud category.

However, there might not be sufficient samplings in every radius bin. As a result, the composite constructed with retrieved values as a function of radial distances may not be representative of the geographical pattern associated with tropical cyclone events.

How many AIRS-DARDAR combined profiles did not converge (I am only finding the number that did converge).

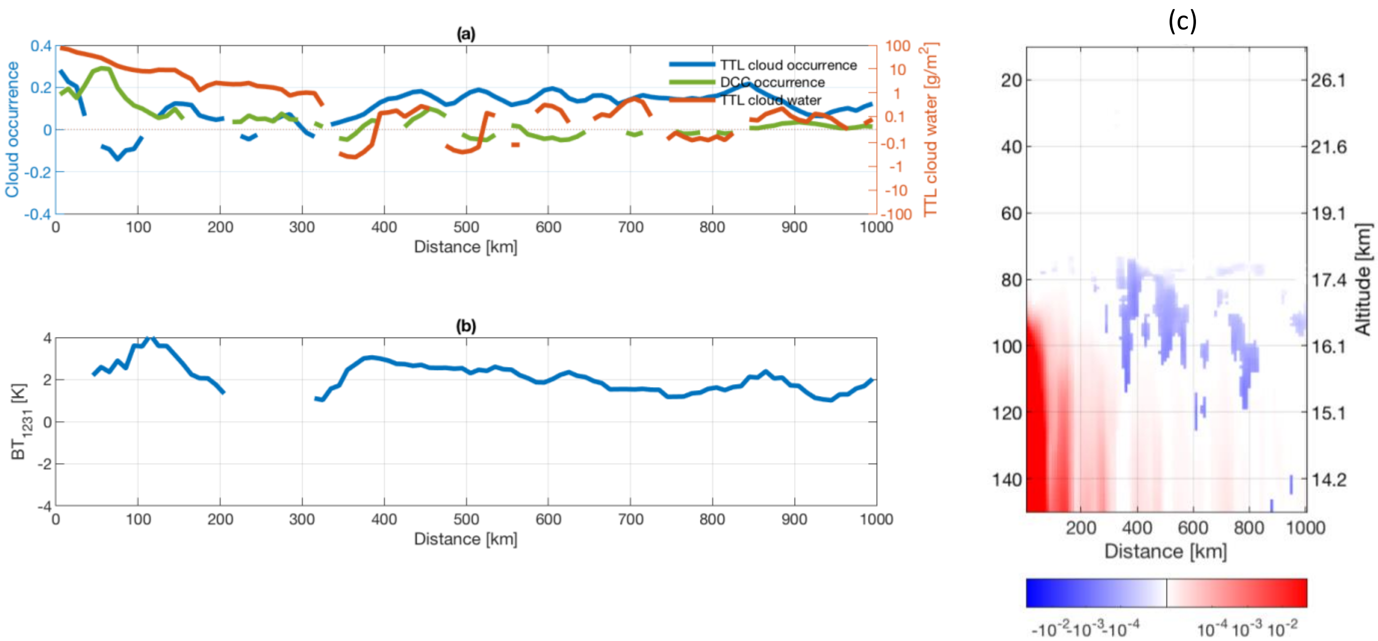
Done.
3475 FOVs meet the criterion of 1) cold scenes 2) CloudSat footprints within 6.5 km from the center of FOV. 2735 converges. 740 FOVs do not converge. A typical situation for these rejected FOVs is that the radiance residual at the initial time step is too large (i.e., > 20K). It happens when cloud amount among a FOV is not uniform so that there is a large difference in cloud states between CloudSat (1.4 x 1.8 km) and AIRS footprint (13.5 x 13.5 km). It may also happen when the optical depth of the topmost cloud layer is less than 1 (in CI and MIX category). We assume that spectral optical properties with respect to cloud ice mass are uniform through vertical layers of an atmospheric column; this assumption fails when the topmost cloud layer does not effectively attenuate infrared radiation.

What do the authors think about diurnal changes? Are these relevant to their results?

We do notice statistically significant day-night differences. A-Train satellite overpass the same region two times a day, once in ascending nodes and the other in descending nodes. The attached figure shows nighttime minus daytime in (a) TTL cloud occurrence, DCC occurrence, and TTL cloud water, (b) BT1231 anomaly, and (c) ice water content, computed as the difference between descending nodes (nighttime) minus ascending nodes (daytime), based on the DARDAR-Cloud (a,c) and the AIRS L1B products (b). Sample densities are displayed in Fig.1 (b,d). Only statistically significant (99%) differences are shown. Note there is a lack of nighttime observation operated by CloudSat after the year 2011, therefore, only overpasses between 2006-2011 are used for generating this figure. Due to a lack of sampling, day-night contrast in temperature and water vapor during this period retrieved by the joint AIRS-DARDAR method does not pass the significant test.

What we see here is that the nighttime overpasses are associated with higher TTL cloud occurrence frequency but lower IWC, especially in regions away from cyclone centers; higher DCC occurrence frequency and higher TTL cloud ice mass near cyclone centers. The BT1231 is warmer in the nighttime.

Although the day-night differences pass the significant test, we are not sure whether it is partly caused by differences in observations and retrieval process, for example, CALIPSO is calibrated using nighttime profiles and daytime calibration are interpolated from the nighttime, and 532-nm channel might be more sensitive to clouds during nighttime (Winker et al., 2009) so more nighttime TTL clouds could be a result of the higher sensitivity. We are not sure whether the contrast between ascending and descending nodes is truly representative of the diurnal cycle because the local solar time of overpasses is fixed at roughly 1:30 and 13:30. We find it curious that while there is a higher occurrence frequency of TTL clouds in the nighttime, the BT1231 is warmer (indicates lower cloud top altitude or less cloud cover) and IWC is lower away from cyclone centers. We do not find a rational way to relate the day-night difference with other results presented in this study, for example, the radiative heating with or without the shortwave heating. Therefore, we chose not to elaborate on the diurnal differences in this manuscript.



The choice of 16 km as a threshold for "overshoots" seems arbitrary and creates awkwardness for the interpretation because it includes the cold point. Many tropical cyclones extend much higher than this at their cores, sometimes to 18 km. I'm not recommending that the authors redo the analysis, but acknowledgement that "overshoots" may instead be "cloud tops" ought to be included.

We chose 16 km as a threshold because the averaged LZRH is located at 15.7 km and the 380 K potential temperature is located at 15.9 km, at the selected region (northern part of the west pacific). We focus on the potential temperature and LZRH, instead of the level neutral buoyancy, for identifying the threshold for 'overshoot', due to the focus on diabatic transports across isentropic surfaces in this study. We do not directly use potential temperature as a threshold because this variable is derived from ECMWF while one of the arguments of this study is that the reanalysis product may not be reliable in the TTL when deep convective storms occur.

You are right that using 'overshoot' can be misleading, as it refers to energetic convective events penetrate the level of neutral buoyancy, if not otherwise stated. In this manuscript, we use 'overshoot' to describe the continuous convective clouds that extend above 16 km (~ 380 K), which is the lower bound of the TTL as defined in this study. We also reword the definition of 'overshoot' to avoid confusion.

Presentation of the new AIRS retrieval technique adds information and shows good promise, and it would be good to also see the authors present and understand limitations of this technique.

Done. We also summarize and compare the limitation of the synergistic AIRS-based retrieval and MLS. The most noticeable drawbacks of the joint-AIRS-DARDAR retrieval method are 1) very limited samples as we only retrieve very thick high clouds currently, making further analysis of the convective impacts difficult; this can be overcome in the future, considering DeSouza-Machado et al. (2018) and Irion et al. (2018) have made some progresses in

incorporating passive sensors to retrieve single-FOVs in all-sky conditions. 2) strong smearing effect limited by spectroscopy of mid-infrared sensors that are not very sensitive to the dry stratosphere. 3) due to the limited precision (0.31 K for temeprature and 0.36 ppmv water vapor), we do not recommend to results from joint AIRS-DARDAR to analyze relative humidity.

Specific suggestions:
The reference to Jensen 2007 should be updated to include more recent Ueyama et al. (2020) and Schoeberl et al. (2018) references in JGR.
Done.

The authors need to list version numbers for each of the data sources, including DARDAR. CloudSat does not observe small ice particles, and so cirrus cloud anvil edges with small effective particle size (< 40-50 microns) won't be included in the CloudSat data, but will be included in DARDAR.

Done. Yes, radar is more sensitive to the largest particles within a cloud volume. To avoid this inconsistency, we use IWC from DARDAR and fluxes, heating rates, and cloud classifications from CloudSat 2B-*-LIDAR products (not the radar-only products), where CALIPSO observations are used as well. McErlich et al., 2020 compared cloud occurrences between 2B-CLDCLASS-LIDAR R05 and DARDAR v2.1.1 and found that DARDAR tends to have less cloud occurrence than 2B-CLDCLASS-LIDAR R05 at high altitudes. They found that DARDAR results agree better with ground-based observations from an AWARE campaign in Antarctica (although such ground-based campaign may underestimate high clouds). For the tropical high altitude in this study, we choose to use 2B-CLDCLASS-LIDAR R05 for cloud classification and DARDAR-Cloud v2.1.1 for cloud occurrence, as described in Section 3.1. We note that for tropical high clouds near tropical cyclone events, using 2B-CLDCLASS-LIDAR directly (without combing DARDAR data for adjustment of cloud occurrence) does not visually impact the result presented in the manuscript (Fig. 2, 3, 9, and 10.).

17 km would likely be a better proxy for the tropopause at many locations in the tropics. Check Tseng and Fu (2017) in JGR, for example and also for a discussion of the positive relationship between tropopause height and deep convection. Cloud tops higher than 16 km are likely not all really overshooting into the stratosphere, and in many places might include a local Ci maximum near the cold point tropopause.

We agree that 16 km may not be high enough for the lower boundary of the stratosphere. However, we do not intend to use it to identify overshoots into the stratosphere. Instead, we use 16 km conservatively to identify the upper-troposphere (lower part of TTL). As stated earlier in this response, the 16 km threshold is set considering the average altitude of LZRH and 380 K for regions within 1000km to cyclone centers over the northern part of the West Pacific, based on CloudSat-2B-FLXHR-LIDAR and CloudSat-ECMWF-AUX. The 380 K isentropic surface may be located at a higher altitude when perturbations in thermodynamic conditions are large, but we do not have a comprehensive temperature dataset with high quality to identify it.

Following this comment, we expand Section 3.1 to clarify the source of CIs. It can be formed by 1) transport of cloud ice via convective outflow or propagating wave activities, and by 2) local cooling that condensates supersaturated vapor that is found in Tseng and Fu (2017) in JGR, where the cooling can be a result of dynamical or radiative processes. This study does

not intend to attribute the source of CI above 16km to convective overshoot. Instead, we break down the TTL cloud ice from DCC-OTs and CIs separately and try to highlight that during tropical cyclone events the total mass of cloud ice above 16 km in CI type is negligible compared to those directly in the top of deep convective clouds (which we phrase as 'overshoot'), despite the frequent occurrences of CIs.

Daytime and Nighttime differences in cloud top height for TTL CI in DARDAR could be substantial because of day/night differences in the Lidar observations. A large majority of TTL CI are only observed by the Lidar due to relatively small effective particle size. Do the authors mention whether they are analyzing daytime, nighttime or both?

See previous responses. We include both. To clarify, we only use data when both radar and lidar observations are available so that the construct between ascend-descend is not a result of the lack of radar observations. Despite the diurnal differences mentioned earlier, we do not notice eyeballing differences in the composite by including/excluding data between 2011 to 2016, where only daytime is available.

To understand the difference between CI and MIX one needs to see Figure 4, so a recommendation is to reference this and to place it earlier in the paper.
Done.

What is the anvil CI above DCC-NOT shown in this drawing? It appears to be part of the anvil, so would those profiles be CI or DCC-NOT?
These profiles will be DCC-NOT because the 'CI' in our classification refers to high cloud-only above a clear troposphere. Any cloud column containing DCC and other cloud types is classified as DCC-NOT.

Figures:
The Figure 1 caption is confusing. The sample density is for CALIPSO and CloudSat, both of which are used for DARDAR. If the sample density is measured at about 1x1 degree resolution (~ 100 km), how can it then be shown at higher resolutions? The text is more clear on this point, but a better caption would allow the figure to be understood better.

Thanks for pointing it out. Take AIRS as an example, we count the number of samples at each 20 x 20 km first and then convert it to number per 100km x 100km so that the number density at different resolutions can be comparable (i.e., number density in Figure 1 (b-d) is in the same unit). We add the definition of the sample density and how it is calculated in the text.

When discussing sampling it is appropriate to say CloudSat/CALIPSO because before 2015 they were both flying information in the A-Train, and both data sets are used in the DARDAR extinction retrieval and subsequent IWC estimation.
Done.

What is the vertical resolution of the combined AIRS-DARDAR temperature profile? What vertical resolution is the AIRS L2 and the MLS data converted to? It would be useful to know this in pressure, but also in equivalent geometric altitude.

We added this information into the manuscript. The vertical resolution is 3.2 km for temperature and 5.8 km for water vapor. Please note that the retrieved values are provided roughly every 0.4 km around the tropopause; the vertical resolution is the FWHM of the vertical averaging kernel that the joint AIRS-DARDAR method can resolve.