Reviewer#1:
The authors did a good job at answering my questions and improving the manuscript. In particular, the domain map is helpful and the maps overall are more readable now. However, I need a few more clarification in the text, in particular around the domain of the study — tropical or tropical+midlatitude. See below for comments.

Thanks to the reviewer again for your acknowledgement of our effort on improving the quality of the manuscript, and on pointing out some weaknesses, especially on the figure suggestion. The responses to your comments are marked below in blue.

Page 5, line 1:

The authors cite Austin et al. (2009) for a 40% error estimate. Austin et al. are wrong, the errors are larger. Although there may not be an authoritative source to cite on this, the intercomparison by Eliasson et al. (2003, JGR, doi:10.1029/2012JD018381) should be of interest.
Thanks for providing the reference. Based on Fig. 2 of Eliasson et al. [2013], the IWP uncertainty for DARDAR remains below 50% when IWP>10 g/m^2. In our paper, the outmost contours of mean IWP in Fig. 2 and Fig. 3 are ~ 20 g/m^2, so the 40% error estimate is a very reasonable estimate for the situation in our paper. Eliasson et al. [2013] also pointed out that the uncertainty may be bigger for mixed-phase cloud. We fully agree with that, but mixed-phase situation rarely occur at the upper troposphere. For mid-level cloud, that may bring some impact.

We now changed the main text of P5 to "*Austin et al. (2009) estimated this IWC product uncertainty of up to ~40%,while Eliasson et al. (2013)pointed out that the error could be much larger in mixed-phase clouds as well as in thin ice cloud. Since the mass of upper-level tropical ice cloud is the main focus of this paper, results would be least impacted by the large uncertainty associated with mixed-phase cloud and thin ice cloud.*"

Page 5, line 10 onward:

I still think there is a scope issue with this paper. Either the authors want to extend the analysis to mid-latitude, or they don't. In the answer to reviewers, the authors state that "this paper is focused on the tropics", as is indeed indicated in the title. Then why do Figures 1(b) and 2(b) include a swath at 50°S, and Figures 3 and 4 use a domain up to ±45°? Either the authors need to use proper analysis (e.g. different layer settings) for the mid-latitudes and make clear in title and abstract that the paper considers tropics and mid-latitudes, or they should show results only from the tropics.
Cloud vertical extension in the "summer hemisphere" would have the same range with those in the tropics, so it would be fair to apply the same layer criteria in the

summer hemisphere up to ±40°. The mean IWP contour also touches ±40° in the summer hemisphere as shown in Fig. 3 & 4.
We've mentioned in the text (Page 9, Line 1-5) that similar "diverging features" are discernable in some mid-latitude active convective regions.

Page 5, line 13:

Ok, I see what the authors mean now.

Page 6, line 4:

The authors still need to the text that for their reasoning they mean "uniformly randomly distributed in the horizontal direction". Of course, this is a hypothetical construction, but the reasoning does require the assumption of a uniform distribution (it is easy to come up with other distributions where the conclusion fails that the two ways of integrating IWP would show no difference).
I think the current presentation still causes some misunderstandings. Say, at a given place, the cloud occurring frequency at a given altitude is X%. Within the area of (17-11)*4=24 km away from this place, the cloud occurring frequency at the same altitude is very likely still X%. Therefore, the probability of cloud occurrence at any time at any altitude should be the same within this 24*24 km box. This is what the "uniformly randomly distributed in the horizontal direction" means.

Hence, it's not necessary to change the current writing: "If the ice cloud density is randomly distributed along the horizontal direction or homogeneous inside a cloud,…". Since cloud occurring frequency nearby any given location can be rationally assumed the same, cloud can pop up randomly at any place within the neighborhood of a given location and altitude.

Page 7, line 1: See my comment for page 5, line 10
Please refer to my reply to page 5, line 10.

Page 21, Figure 1c: in the illustration, the northward view and southward view appears a factor of sqrt(2) longer than the nadir view, which is rather a factor sec(77°)=4.4 longer. Of course, as calculated by the authors, this does not mean a factor 4.4 more IWP. To avoid giving the impression that the slant path is longer than the vertical path, I would indeed delete the two slantwise line-of-sight arrow lines, and possibly replace this by a small vertical arrow in each of the individual squares that make up the slant path, to illustrate the mock slantwise path.
Thanks for your suggestion. The two slantwise paths are now removed and replaced by staggering arrows to illustrate the integration process in Fig. 1c and Fig. 7.

Also on Figure 1c: I didn't notice the 77° indication earlier. It is good that this is now mentioned in the caption. Perhaps the authors could indicate this with an arrow (probably at the northward rather view because the space above the southward view is already taken by the arrow indicating the cloudsat motion)

Without the slantwise path, a short dashed line is added to help illustrate the 77° angle.

==

Reviewer #2:

Overall

The article has improved considerably in my opinion due to the authors efforts to clarify many statements and generally improve the wording. However, overall I think more attention should be paid to the uncertainties. Granted, that to cover all the uncertainties is outside the scope of the paper, but I feel many statements need to be revised baring them in mind. Following is a response to the authors response to the first review.

We thank the reviewer again for your acknowledgement of our effort on improving the quality of the manuscript, and on pointing out some weaknesses. The responses to your comments (black) are marked below in blue, and the previous round of response is marked in grey.

Response to authors response to my review

• Firstly, DARDAR is a joint retrieval of CloudSat, Calipso and MODIS measurements. It contains much more ample ice cloud details compared with CloudSat. DARDAR is indeed not an independent obser- vation. Some conclusions (e.g., their consistencies validate ....) have been tuned down in the revised manuscript. Secondly, DARDAR and CloudSat disagree with each other on the tilt direction of the lower level ice clouds, while WRF simulation results support the CloudSat observation.

Yes, DARDAR is not identical to CloudSat, but in terms of IWP, for clouds with a reasonable amount of IWP from approximately 80gm−2 or more, nearly all of the information content in the DARDAR retrieval comes from the CloudSat's CPR measurements [e.g., Figs. 4 and 5 in Eliasson et al., 2013]. As also mentioned in the manuscript, the main differences between DARDAR and CloudSat are when the clouds are thin.

Thanks for pointing out the difference and commons between these two datasets. We didn't state in the manuscript that they were two independent dataset. Rather, we treat DARDAR as a complement to CloudSat as it has information from CALIPSO. Now we explicitly added your comment in the text:

"Since DARDAR retrieval is dominated by CloudSat input when IWP exceeds 80g/m2, we do not expect the two results would be significantly different. DARDAR is used as a complement in this study rather than an independent evidence."

On your second point, for me it is somewhat surprising that these two datasets are not more similar. However, there is a big difference in how the datasets handle the transition from ice water content to liquid water content, which may have a large influence on the IWP retrievals. All clouds below −20◦C are considered to be

completely ice for CloudSat 2B-RO, yet DARDAR uses the detection of supercooled liquid identified by a strong lidar backscatter signal to find the phase transition [Stein et al., 2011] (if the lidar has not already attenuated). However, where the authors assess clouds above 9 km, (which is around −20 ◦C in the tropics), I'm not sure if the different phase assumptions have an impact on these results or not (but I don't think it can be obviously ruled out). However, for the studies of IWP tilt below this altitude the uncertainties due to the cloud phase can be quite big.
Thanks for pointing out the mixed-phase cloud situation. We didn't pay attention to such a difference until you and the other reviewer pointed it out. However, based on the standard tropical atmosphere temperature profile (e.g., Fig. R1 below), the atmosphere should be cooler than -20°C above 8 km (using diabatic lapse rate), and our upper troposphere is defined between 11 and 17 km.

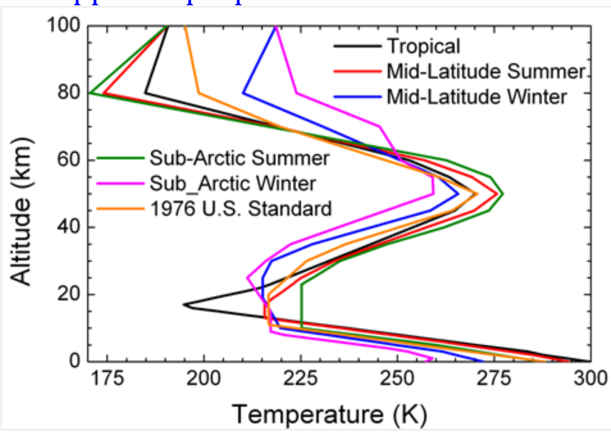

Figure R1: Standard atmosphere temperature profile (source: modtran5.com/faqs/index.html).

We now changed the main text of P5 to "*Austin et al. (2009) estimated this IWC product uncertainty of up to ~40%, while Eliasson et al. (2013)pointed out that the error could be much larger in mixed-phase clouds as well as in thin ice cloud. Since the mass of upper-level tropical ice cloud is the main focus of this paper, results would be least impacted by the large uncertainty associated with mixed-phase cloud and thin ice cloud.*"

• By definition, plane-parallel atmosphere means that the atmospheric compositions (e.g., CO2, Ozone, H2O, cloud) and characteristics (e.g., Temperature) should be horizontally homogeneous. In such case, any property from nadir-view can be immediately converted to any slantwise view by simply multiplying the cosine of zenith angle. Thats apparently not a good assumption for highly inhomogeneous cloud field. In case the reviewer is not familiar with the plane-parallel bias in cloud property retrievals, the reviewer is kindly referred to Cahalan, R. and his colleagues publications on such a topic, some of which are listed in the reference list appended with this comment response.
Yes, I am aware of the plane parallel assumption. It has many problems for cloud retrievals, and as described in this article, cloud tilt contributes to this problem for sure. My comment was rather to point out that this paper is not the one that

"invalidates" the plane parallel assumption, since problems are well known with this assumption. There are much bigger problems to the plane-parallel assumption than tilt in my opinion. I will list some:

– The inhomogeneous distribution of particle shapes is a big problem, and, for instance, the choice of particle shapes made in the retrieval can lead to a factor 2 difference in IWP retrievals alone [Posselt et al., 2008].

– The "real" vertical gradient in cloud particle size is also a big problem (there are mostly small particles at the cloud top and larger particles futher down). Passive sensors such as MODIS can only retrieve the effective cloud radius from about the top 3-4 optical depths, (depending on the cloud and the measurement frequency) [McFarquhar and Heymsfield, 1998], and this is supposed to represent the mean effective radius of the whole column (so the thicker the cloud, the larger the systematic bias). This fact alone leads to large uncertainties in IWP retrievals (e.g., Zhang et al. [2010]).

– Also, the existence of multilayer clouds can also be a large source of uncertainty [e.g., Cooper et al., 2007].

– Beam filling, i.e. measurements that contain information both from the surface and a cloud (fractional clouds, and semi-transparent clouds) is also a potentially a large source of errors, especially at the edge of larger cloud systems.

– Cloud phase is a big problem too. Before a cloud water path or effective radius retrieval is made, the cloud phase must be determined. The cloud column is then assumed to be completely that phase, and if the phase is wrong the retrieval will be very wrong.

– The list goes on It is quite tricky to assign a total uncertainty to IWP retrievals since they are situation dependent, and we don't fully understand the size of all the uncertainties, but overall the uncertainties are large. However, the reason we retrieve using the plane parallel assumption is because, by this assumption, IWP is a linear function of optical depth and effective radius [Stephens, 1978]. Also, not much else can be done since passive retrievals are based on measurements which are the combined contributions from the atmospheric column. I believe as long as we are honest with these uncertainties, and understand them, the plane parallel assumption can used. Active sensors, such as cloud radar and lidar from CloudSat and Calipso are particularly valuable, since they can provide the information on the vertical structure of the cloud that passive (down-looking) sensors cannot.

Thanks for your review and personal inputs of the plane-parallel assumption. In our previous version of revised manuscript, we have completely deleted such a strong statement, and didn't discuss about this assumption at all. Hope that would resolve some of the arguments that would be potentially brought up.

- Firstly, Austin et al. [2009] claimed that CloudSat IWC retrieval error was at most 40 %, not at least. Therefore, 5 % to 20 % is an alarming value to raise concern. Secondly, this is the first research that shows that ice cloud tilt is systematic rather than random.

It is noted that you have found a systematic uncertainty, which is especially helpful since uncertainties are often only presented as random. Systematic uncertainties are not adressed that much in other literature that I'm aware of. To your first point, it is good to quote the uncertainty provided by Austin et al. [2009], but it is still a lower bound of the total uncertainty at best. The 40 % estimate was found from simulated reflectivity measurements using a microphysical database based on in situ measurements from 4 campaigns in Heymsfield et al. [2008]. But, there are more uncertainties that propagate into the retrieval, that should be taken into account when talking about the real world uncertainties.

– The uncertainties in the auxiliary model data

– The uncertainty in assuming any ice particle model

– The uncertainties arising from assuming globally valid temperature dependent cloud ice particle distributions and other microphysics properties when making the CloudSat retrievals.

– The errors introduced by large particles. Even a relatively small number of large particles that violate the Rayleigh criterion may cause significant errors in the retrieval [Austin et al., 2009]

– The errors from the lack of an independent way to determine the phase of the cloud [Devasthale and Thomas, 2012]

– The uncertainty is also certainly not constant. The error is generally larger for IWC retrievals of small amounts than large [Eliasson et al., 2013, e.g., Fig 2 in] (true for DARDAR and surely true for CloudSat).

– The uncertainty due to beam-filling. This means, the uncertainty due to not knowing if, but assuming that the measurement is completely filled or completely not filled by a cloud.

– The list goes on. Many, if not all the above are hard to put a number on, and are also situation and regionally dependent, but we must make assumptions in order to make a retrieval at all. I believe saying that the uncertainty in our best IWP retrievals in the real world (probably, DARDAR/ CloudSat 2C-ICE) is around 100 % (and that's just the random part) is not an exaggeration (but the numbers are hard to cite). I asked about the uncertainty in these retrievals to Andy Heymsfield personally a couple of years ago and this was also his view.

Thanks for listing so many error sources. Indeed, we can only evaluate the percentage of uncertainty only from very limited sources. But based on the "random" occurrence assumption, any uncertainty would be beat down to -> 0 when we add a large number of obs. together. That is why the sysemmatic uncertainty discovered by our paper has a particular meaning.

The detailed cloud vertical structures are difficult to be resolved in passive satellite observations. Subsequently, they are either neglected or significantly simplified in GCMs.

This sentence is better than before but I still have some comments. Firstly, as mentioned above, there is virtually no information on the vertical structure of clouds from passive down-looking instruments, and I still think it is too harsh to say “the models are neglecting the vertical structure of clouds”, and this opinion should be toned down (along with similar strong statements in the article). Since the introduction of CloudSat and the existence of satellite simulators which simulate the observations from climate models [Bodas-Salcedo et al., 2011], many efforts have been made to validate the vertical distribution of clouds using active sensors [e.g., Cesana and Chepfer, 2012], up the CloudSat and Calipso era, the models did not have had much to go on.

As for your first point, passive sensor with hyper-spectral resolution (e.g., AIRS) can surely provide vertical cloud information to some extent.

Austin et al. [2009] (see added reference) claimed that the uncertainty is less than 40 %. Could you give us a reference that explicitly claims the official uncertainty level? Thanks. A sentence has been added in the paragraph to mention the uncertainty in the retrieval.

I touched on this above. There are quite a few additional sources of uncertainty that are tricky to quantify. It is definitely good to at least mention something about the error as you do now, although as my comment implies the number 40% is not enough if you are talking about the total uncertainty. It’s just good to bear in mind the total uncertainties when making conjectures and formulating the text.

Please see my response to your previous question and comments about uncertainty sources. We’ve added some sentences mainly on Page 5 to acknowledge possible larger uncertainty ranges. But again, as along as these uncertainties can be assumed “randomly occur”, they will become small enough to be neglect when the sample size is large enough. Only when the uncertainties are biased systematically and REGIONALLY that may contribute to the systematic “tilt” signal we interpreted in the paper. This is unlikely to occur in my view.

From the first review: The broad consistency between CloudSat and DARDAR analysis results validate the robustness of our findings. As mentioned earlier, these datasets are not independent. We agree with the reviewer that DARDAR and CloudSat are not independent. Please notice that we didn’t claim anywhere in the paper that they were independent.

My point was not that the authors claim that DARDAR IWP and CloudSat RO IWP are independent but I was rather hinting that since the datasets are dependant, and statements about the robustness of a particular finding cannot be found by comparing to the other dependent dataset.

In our previous version of the revised manuscript, “validate” has been replaced by “show”. The statement tone has been tuned down.

Yes. At lower level, CloudSat result indicates that the ice clouds should tilt inward while DARDAR result is contradictory. Although we know that CloudSat has saturation issue with heavily precipitating cloud, WRF simulation agrees with CloudSat analysis result nevertheless.

This is puzzling. Can WRF be fully trusted here? CloudSat RO doesn’t just have a saturation issue, in particular it has a cloud phase issue (which is also an attenuation

issue) which DARDAR doesn't have (as mentioned above). I wonder what DARDAR is doing with the same, but more, information when doing its retrieval which makes it worse than CloudSat's 2B-RO product with its simple linear cloud phase assumptions? I think DARDAR is more likely to be right since it makes use of an independent measurement (lidar) to detect if supercooled water is present (Supercooled water droplets can exist in temperature lower than −30 ◦C [Rosenfeld and Woodley, 2000, Protat et al., 2009]).
I don't get the point completely. Since lidar would be more easier saturated than radar, in heavily precipitating scenes, how could DARDAR (with adding in CALIOP information) reveal the structure at lower level?
WRF simulation has its own limitation, mainly because the highest resolution used here (3 km) is still too coarse to resolve the full tilted structure. An LES model run with domain as large as the WRF simulation here could be more helpful but unfortunately not realistic for this work. Hence, we open the lower-level tilt structure and stated that larger uncertainties exist and the conclusion is not solid for lower level.

Moreover, this is the first finding that could tilt could be impactful to cloud retrievals. It is worth mentioning the potential issue of retrieval algorithm that ignores this effect, even the impact would be small compared with some other dominant factors.
I agree that the tilt found in this study systematically adds to the error in some regions, and is important. Yet, I don't remember reading in this paper about the existence of the other more dominate errors that the cloud retrievals have to deal with. I feel this context is missing.
The retrieval errors have been mentioned when introducing the datasets (page 4-5).

Technical comments
In the abstract

- unrealistic/unconsidered. OK, it may be more or less unrealistically parametrised, but I would not go so far as to say they don't even consider cloud vertical structure. When this strong statement is repeated in the introduction, there is still no reference backing this up. I suggest toning down and removing the word unconsidered.
  The word "unconsidered" has been deleted.
- In my opinion, substantial is a strong word for the measurement errors considering there are several even more serious problems. (see above)
  Introduction
  "Substantial" has been replaced by "considerable".
- the bulk outlook is visibly irregular. The word "outlook" can't be used here. Do you mean that clouds are "visibly irregular" or just "the clouds have an irregular appearance"?
  The sentence has been replaced as "cloud is visibly irregular".
- The 'detailed' cloud vertical ....

It depends on where you draw the line for what is considered “detailed”. Passive instruments have next to no capability of retrieving the vertical structure of clouds. I suggest removing “detailed”
“Detailed” is removed.

- difficult to be resolved to “difficult to resolve”
  Changed as suggested.
- either ’neglected’ ....
  Again, the vertical structure is hardly neglected. Assuming max-random overlap in order to emulate the vertical cloud distributions is often a decent compromise in a global sense.
  “either neglected” has been deleted.

UT cloud tilt in the tropics

- I would change differencing to “finding the differences in the”
  The current wording is more concise, and follows the mathematical convention, so we decide not to change.
- The broad consistency between CloudSat and DARDAR analysis results show the robustness of our findings.
  I still feel very little can be said about the robustness of the findings from one dataset using another dataset that is heavily dependent. Maybe leave that sentence out?
  In the upper troposphere, CloudSat and DARDAR are different, not only because of abundant CALIPSO information, but also because of the retrieval strategy is quite different. Therefore, I think the statement is appropriate here.
- Ice cloud tilt in the middle troposphere (5-11 km) still has some ambiguities due to large un- certainties embedded in IWC retrievals below 9 km for heavily precipitating cases.
  As mentioned above, that is just one big problem in this altitude region in the Tropics. Arguably the simplistic cloud phase handling in the CloudSat RO dataset causes even bigger problems. Devasthale and Thomas [2012] addressed this uncertainty extensively. For instance they found a difference of up to 20 % to 40 % in liquid water path (LWP) in the Tropics depending on how you choose to parametrise the relationship between temperate and cloud phase (cloudSat uses a linear transition of ice fraction = 0 at 0 ◦C to ice fraction=1 at −20 ◦C). The uncertainty in IWP is in the same ball park as LWP.
  The mixed-phase situation is now added in the discussion accordingly, which reads “The simplified assumption of IWC/Liquid Water Content (LWC) partitioning of this dataset on mixed-phase situation contributes another big source of uncertainty.”
- upright and upwind.
  what does this mean? is this updraught?
  No, upwind means against the wind shear.
- Swap The entire idea to “This”
  Changed.

Conclusion

- Moreover, both CloudSat and WRF simulations suggest a mid-level (5 - 11 km) cloud mass.

  Are the uncertainties just too high to make this statement?

  A sentence has been added afterwards to clarify that this is not a solid conclusion yet: “The mid-level tilt is still debatable due to large uncertainties associated with the limitation of W-band radar in precipitating scenes and mixed-phase scenes, and the coarse resolution of WRF simulations.”