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. The authors response or text from the article is given in blue.

1 Response to authors response to my review

- This is only part of the conclusion, or the implications of the importance of this work to improve future satellite retrievals and/or model cloud physics scheme development. The major conclusion is that upper-troposphere ice cloud in the nature has a systematic meridional tilt in the tropics, which has never been studied before. This paper is a science and observation oriented work, and the scientific discoveries are the real focus.

  OK. I understand that point. It is a good conclusion.

- 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 observation. 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 $80\,\mathrm{g\,m^{-2}}$ 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.

  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 $-20\,°\mathrm{C}$ 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 $-20\,°\mathrm{C}$ 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.

- 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.

- 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.

- 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.

- 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.

- parallax ...

  I think the parallax explanations are clearer now

- 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 didnt 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.

- 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]).

- Moreover, this is the first finding that could tilt could be impactful to cloud retrievals. Its 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.

2 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.

- In my opinion, substantial is a strong word for the measurement errors considering there are several even more serious problems. (see above)

Introduction

- 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 '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"

- difficult to be resolved to "difficult to resolve"

- 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.

UT cloud tilt in the tropics

- I would change differencing to "finding the differences in the"

- 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?

- 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.

- upright and upwind. what does this mean? is this updraught?

- Swap The entire idea to "This"

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?

References

R. T. Austin, A. J. Heymsfield, and G. L. Stephens. Retrieval of ice cloud microphysical parameters using the CloudSat millimeter-wave radar and temperature. 114:D00A23, 2009. doi: 10.1029/2008JD010049.

A. Bodas-Salcedo, M. J. Webb, S. Bony, H. Chepfer, J.-L. Dufresne, S. A. Klein, Y. Zhang, R. Marchand, J. M. Haynes, R. Pincus, and V. O. John. COSP: satellite simulation software for model assessment. 92:1023–1043, 2011. doi: 10.1175/2011BAMS2856.1.

G. Cesana and H. Chepfer. How well do climate models simulate cloud vertical structure?a comparison between calipso-goccp satellite observationsand cmip5 models. *GRL*, 39:L20803, 2012. doi: 10.1029/2012GL053153.

S. J. Cooper, T. S. L'Ecuyer, P. Gabriel, A. J. Baran, and G. L. Stephens. Performance assessment of a five-channel estimation-based ice cloud retrieval scheme for use over the global oceans. 112:D04207, 2007. doi: 10.1029/2006JD007122.

A. Devasthale and M. A. Thomas. Sensitivity of cloud liquid water content estimates to the temperature dependent thermodynamic phase: a global study using CloudSAT data. 2012. doi: 10.1175/JCLI-D-11-00521.1.

S. Eliasson, G. Holl, S. A. Buehler, T. Kuhn, M. Stengel, F. Iturbe-Sanchez, and M. Johnston. Systematic and random errors between collocated satellite ice water path observations. 118:1–14, 2013. doi: 10.1029/2012JD018381.

A. J. Heymsfield, A. Protat, R. Austin, D. Bouniol, R. Hogan, J. Delanoë, H. Okamoto, K. Sato, G.-J. van Zadelhoff, D. Donovan, and Z. Wang. Testing IWC retrieval methods using radar and ancillary measurements with in situ data. 47(1):135–163, Jan. 2008. doi: 10.1175/2007JAMC1606.1.

G. M. McFarquhar and A. J. Heymsfield. The definition and significance of an effective radius for ice clouds. 55:2039–2052, 1998.

D. J. Posselt, T. S. L'Ecuyer, and G. L. Stephens. Exploring the error characteristics of thin ice cloud property retrievals using a Markov chain Monte Carlo algorithm. 113:D24206, 2008. doi: 10.1029/2008JD010832.

A. Protat, D. Bouniol, J. Delanoe, P. T. May, A. Plana-Fattori, A. Hasson, E. OConnor, U. Gorsdorf, and A. J. Heymsfield. Assessment of Cloudsat reflectivity measurements and ice cloud properties using ground-based and airborne cloud radar observations. 26:1717–1741, Sept. 2009. doi: 10.1175/2009JTECHA1246.1.

D. Rosenfeld and W. L. Woodley. Deep convective clouds with sustained supercooled liquid water down to -37.5°c. *N*, 405:440–442, 2000. doi: 10.1038/35013030.

T. H. M. Stein, J. Delanoë, and R. J. Hogan. A comparison among four different retrieval methods for ice-cloud properties using data from CloudSat, CALIPSO, and MODIS. 50:1952–1969, 2011. doi: 10.1175/2011JAMC2646.1.

G. L. Stephens. Radiation profiles in extended water clouds. II: Parameterization schemes. 35:2123–2132, 1978.

Z. Zhang, S. Platnick, P. Yang, A. K. Heidinger, and J. M. Comstock. Effects of ice particle size vertical inhomogeneity on the passive remote sensing of ice clouds. 115:D17203, 2010. doi: 10.1029/2010JD013835.