Response to Referee #1

We want to thank the referee for appreciating our work and for the thoughtful comments and suggestions. Most of them have been taken into account to improve the manuscript. We apologize for the difficulties associated with the length of the manuscript and excessively long sentences. We have re-worked the manuscript and addressed each comment. In the text below, the reviewerøs comments are marked in italics blue and our answers are given in normal font.
As the referee has correctly pointed out, the method itself is not new (the first author developed it in the 1990øs for TOVS), and there exist several publications, which are referenced in the article. Indeed it was a difficult task to select what should be presented and what left out, which is reflected in differing opinions of the 2 referees.
Since both referees suggested to shorten the manuscript, we have done our best to do it without losing the message we wanted to deliver to the community. Here is the list of actions performed
1) shortening section 2 õData and methodsö and moving a shortened version of section 3.1 õCollocated AIRS-CALIPSO-CloudSat dataö to this section
2) simplifying Table 2, taking out 5 figures / 22 figure panels (3 figures moved to supplement)
3) taking out the ENSO discussion in section 5 (together with Fig. 16) and
4) revising the remaining applications in section 5

We do not agree with the suggestion of a complete removal of section 5 õApplicationsö, as the presented method is not new and one of the goals of this article was to present scientific applications (as indicated in the title).
Since the results similar to those presented in new Fig. 12 have recently been published for other data sets, it would be difficult to use the presented material in a separate publication. We compare our results to one of them and point out an interesting extension. We plan to work on a more complex analysis to pursue this subject further, but we think itøs important to present already these results in the current publication.

Title: the authors should consider a better title that is punchier and emphasizes the great aspects of using sounders for cloud properties (and not have **õ***weaknesses***ö** *in the title)*
the title was changed to : **Cloud climatologies from the InfraRed Sounders AIRS and IASI: Strengths and Applications**

Abstract: it is pretty long and not very specific. For instance, lines 24-28 has a single long sentence making multiple points about the apparent cloud top/base. Is the correction for co2 really that original and worth advertising in the abstract? On lines 23-24, the **õ***global cloud amount***ö** *is detected clouds, not effective emissivity?*
We have substantially re-worked the abstract, to be more specific.
Rewritten to: The global cloud amount is estimated to 0.67 ï 0.70, for clouds with IR optical depth larger than about 0.1. The spread of 0.03 is associated with ancillary data.

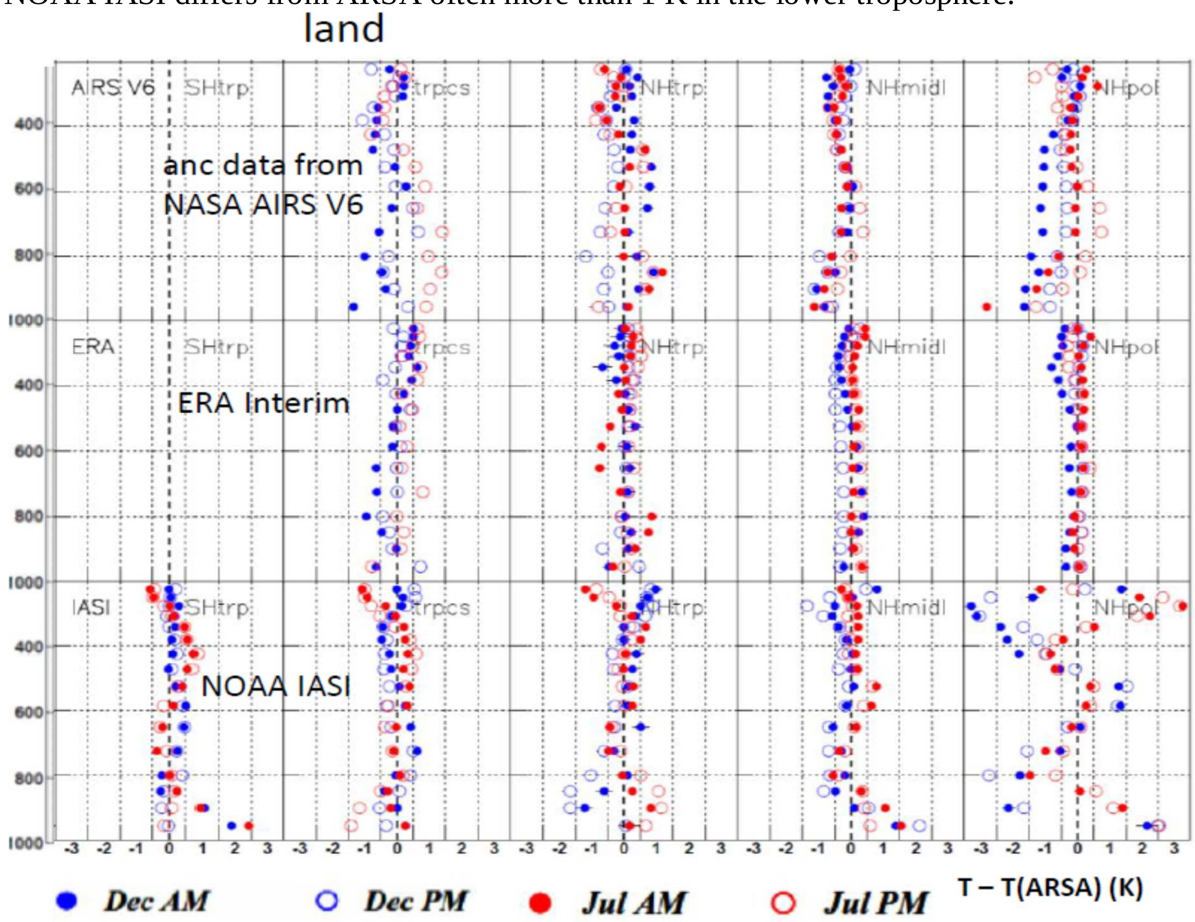
It is really the amount of detected clouds; it is interesting to mention that global effective cloud emissivity of detected clouds is very similar: 0.65-0.66;
This leaves global effective cloud amount (detected clouds weighted by cloud emissivity) to about 0.46-0.48.

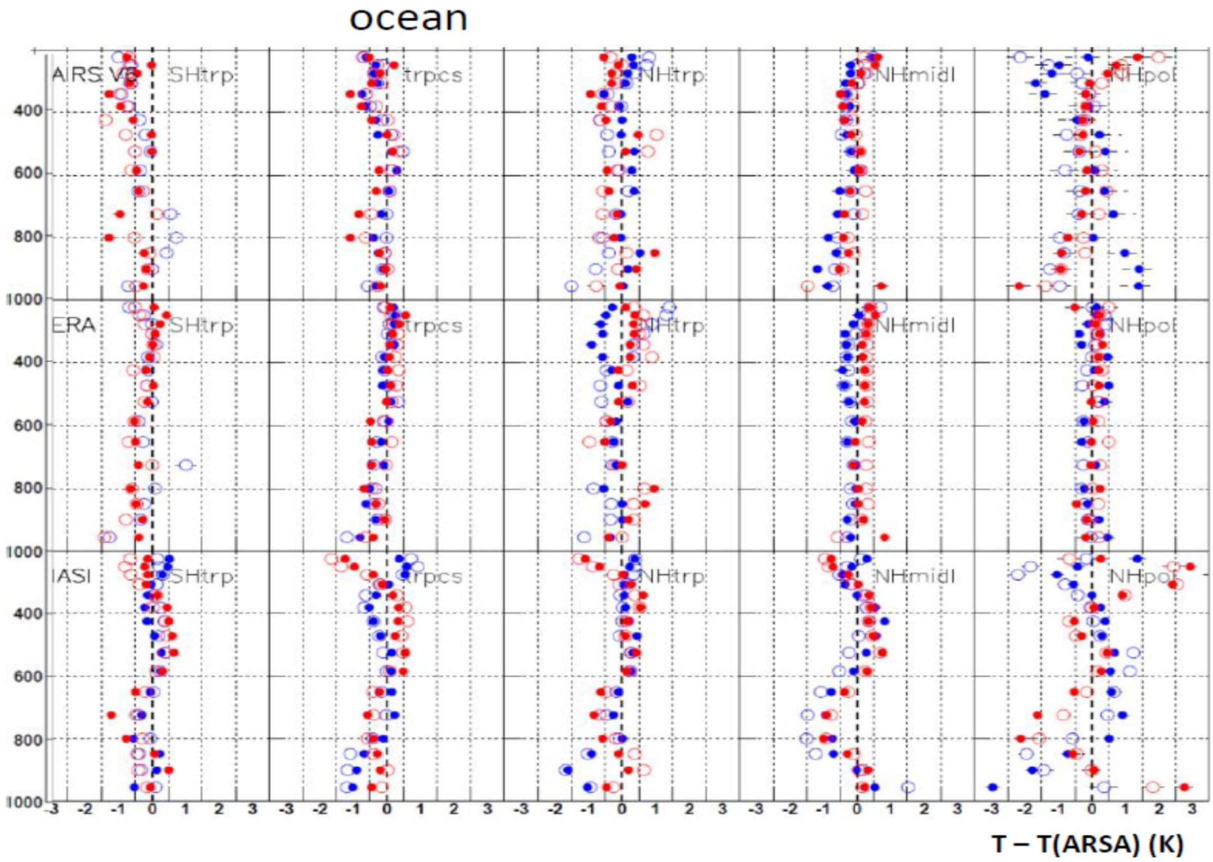
p. 5, lines 20-21: did the authors try (or consider) using a SST data set independent of the IR sounders, say, RTG-SST or the optimal approach using microwave made available at www.remss.com?

There are two philosophies in creating cloud climatologies : 1) ancillary data are also taken from observations, and 2) ancillary data are taken from model forecast or meteorological reanalyses. The advantage of the first is that these climatologies are independent of model input, however the problem is that the ancillary data might have biases due to faults in clear sky detection and due to interpolation when no good quality data are available.
In this article we compare these approaches; for the first one, we preferred to stay with data which include the same instrument (the ancillary data come from a combined IR sounder ï microwave retrieval). For IASI, at the time of the development, the available ancillary data did not have the quality needed. Therefore we switched to the second approach.
A separate SST data set would not help, as we also need surface temperature over land, and both are needed at the satellite observation times. I addition they also should be coherent with the retrieved atmospheric profiles.

p. 5, lines 23-24: óquite differentô is not quantitative and not useful in the context of this discussion. How different were they?
Rewritten to: The comparison with collocated temperature profiles of the Analyzed RadioSoundings Archive (ARSA, available at the French data centre AERIS) has shown that, while AIRS-NASA and ERA-Interim (section 2.3) temperature profiles do agree in general with the ARSA profiles within 1 K, differences between IASI-NOAA and ARSA profiles were often larger than 1 K in the lower troposphere (not shown).

In the following plots we present differences in T profiles between NASA AIRS V6 and ARSA, ERA Interim and ARSA and NOAA IASI and ARSA, separately for different latitude bands over land (above) and ocean (below). Whereas AIRS and ERA agree in general within 1K, NOAA IASI differs from ARSA often more than 1 K in the lower troposphere.





Line 25: the IASI and AIRS sounders will not resolve the diurnal cycle but will capture aspects of it.

We agree that the diurnal cycle is difficult to resolve with data given in temporal intervals with 4h ï 8h ï 4h ï 8h, but as one can see in the cited conference proceeding (publication is under preparation), by using appropriate analysis techniques, both the amplitude and phase of the diurnal cycle of upper tropospheric clouds can be obtained, especially due to the fact that IR sounders provide unbiased day-night results.

Rewritten to :

This brought us to the conclusion, that ancillary data from the same source are necessary to make use of the AIRS ï IASI synergy for exploring cloud diurnal variability in a coherent way.

Lines 26-27: if it is of any help, there is a paper that describes cloud type comparisons between AIRS and ECMWF T/q:
Yue, Q. et al. (2013), Cloud-state dependent sampling in AIRS observations based on CloudSat cloud classification, J. Climate, 26, 8357ï 8377.

Unfortunately this very interesting article refers to NASA AIRS L2 data of Version 5 ;
We have used in our revised cloud climatology NASA AIRS L2 data of Version 6

p. 6, lines 9-12: the variables should be listed here (e.g., T, q, emissivity, sfc T, etc.)

the whole paragraph was taken out, as this issue was already partly discussed in 2.2.

p. 7, lines 11-12: here is a good example of over explaining. Why should ófor which temperature first increases with height before decreasingô be included? This is technically only true if ascending in the atmosphere. Line 12: ómoved to the inversion layerõ is not clear. Is the cloud placed at the base of the inversion? Hopefully not the top because that would be impossible in reality. Line 14: ó. . .about 7 to 15% of the time.ô

Text improvements taken into account.
In the case of an inversion, the cloud height is set to the level at which the temperature starts to decrease with height.

p. 8, lines 20-23: this statement is unclear. How can óclear skyô be ónot too cloudyô?
text in parentheses taken out; the synergy of IR sounder and microwave also leads to retrievals for party cloudy scenes. In that case, a ócloud clearingô is performed before the retrieval.

p. 9, line 7: there is a specific QC approach that filters based on a PBest or PGood pressure level. Was this done on a per profile basis? Or were the Level 3 gridded AIRS Team products used?
We used the quality criteria on a per profile basis, as we work with L2 data ; reference added

p. 9, lines 8-9: there is a paper that describes AIRS surface temperature biases with respect to ship observations:
Dong, S., S. T. Gille, J. Sprintall, and E. J. Fetzer (2010), Assessing the potential of the Atmospheric Infrared Sounder (AIRS) surface temperature and specific humidity in turbulent heat flux estimates in the Southern Ocean, J. Geophys. Res., 115, C05013, doi:10.1029/2009JC005542.
Again, the problem is that this paper refers to AIRS V5 ; we had problems to find published results for the AIRS V6 version, apart from the V6 L2 Performance and Test Report.

p. 9, line 11: with respect to what is the land more complex?
We meant that there was not a clear bias found as over ocean ; clarified in the text to :
Since differences over land might be positive or negative (Fig. 2), we left the AIRS-NASA surface temperature (T_{surf}) values as they are.

p. 10, line 23: is the artifact in cloud amount causing more clouds? Less clouds? Higher clouds? Lower clouds?
Global cloud amount is increasing, when the CO_2 increase is not taken into account in the computation of atmospheric spectral transmissivities (new Fig. 10);
When splitting into low-level and high-level cloud amounts, the artefact led to increasing CAL and slightly decreasing CAH.

p. 11, line 3: base of the inversion?
In the case of atmospheric temperature inversions, the cloud height is moved to the level at which the temperature starts to decrease with height, and ε_{cld} is scaled accordingly.

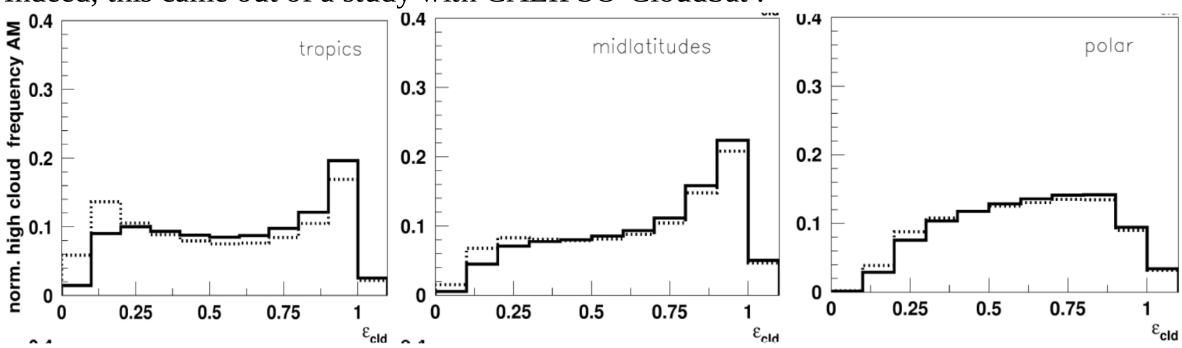
p. 11, line 25: is ónot cloudyô the same as óclearô? or something else?
As the IR sounder footprint size is large, it is difficult to distinguish between completely clear sky and cloudy. Even the evaluation with CALIPSO-CloudSat stays approximate as the sampling is only about 1.5 km x 2.5 km, which corresponds to a sampling of about 2 %.

p. 12, line 10: óexplainableô should be óexplainedô. The paper could use a good thorough editing for clarity of English.
Unfortunately, all authors are non-native English speakers; we tried however to improve the readability of the present version to the best of our abilities.

p. 12, lines 14-16: are there three different sigmas for the three different emissivity_i values? It appears that some of the clear will be selected as cloudy, and vice-versa. Is this correct?

The thresholds were chosen separately for 1) ocean, 2) land and 3) snow/ice, as the distributions in new Fig S1 (original Fig. 2 moved to the supplement) showed slightly different distributions. Indeed, all methods using thresholds include misidentifications. These are difficult to estimate because of the sampling (2% of CALIPSO-CloudSat per AIRS footprint). The cloud detection includes 80 (over ice) to 92% (over ocean) cases for which CloudSat-lidar GEOPROF and CALIPSO at 5 km resolution (excluding subvisible cirrus) have identified at least one cloud layer, and 30% cases for which the samples did not include a cloud layer. The latter might look at first as a large misidentification of clear sky as cloudy, but the very small coverage of the CloudSat-CALIPSO samples (2%) certainly includes partly cloudy fields.
Results in section 3 show that by using these thresholds the overall agreement with CloudSat-CALIPSO is 70% (over ice) to 85% (over ocean), given as hit rates.

p. 12, lines 20-21: the emis < 0.1 threshold is very conservative. The IR sounders will capture a lot of optically thinner clouds than that. Are the authors arguing the point that below that threshold some clear values could leak in? The paper by Kahn et al. (2008) seems to argue that the emis threshold could be lower than that:
Kahn, B. H. et al. (2008), Cloud-type comparisons of AIRS, CloudSat, and CALIPSO cloud height and amount, Atmos. Chem. Phys., 8, 1231ï1248.
Indeed, the AIRS-LMD climatology (Stubenrauch et al. 2010) went down to an ε_{cld} of 0.05. Considering the large footprint and a comparison of ε_{cld} distributions for cloudy and clear sky CloudSat-CALIPSO scenes (see below), we decided to exclude scenes with ε_{cld} < 0.1.
We made the sentence more explicit : To reduce misidentification of clear sky as high-level clouds, only clouds with ε_{cld} Ó0.10 are considered.
Indeed, this came out of a study with CALIPSO-CloudSat :



The above figures present normalized ε_{cld} distributions of high-level clouds, after multi-spectral cloud detection, but leaving clouds with 0.05 < ε_{cld} < 0.10 as clouds, separately for cloudy scenes defined by GEOPROF and CALIPSO (full line) and for all scenes (dotted line). The first bin includes scenes with 0.05 < ε_{cld} < 0.10; in the tropics this bin has more clear sky than high-level clouds. Therefore we have moved the threshold to 0.1. As the contribution of the first bin is small compared to the integral, this seemed a reasonable choice.

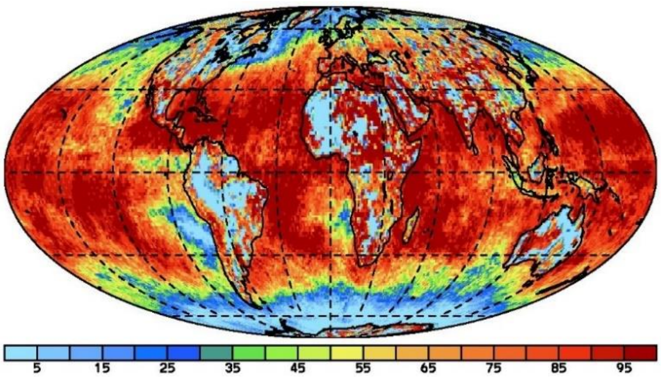
p. 12, section 3.1: this is where the paper starts to be a real grind. Wasnât the methodology of the AIRS and C/C comparison described in a previous paper(s) by the lead author? There must be a way to tighten this up and make it more concise, but I am lacking any good suggestions for that.
Indeed, part of the description of the collocated dataset was already published before, though not the computation of the cloud height corresponding to a specific optical depth. Referee #2 finds that this section is not detailed enough.
We have rewritten this section and moved it to section 2.4, hoping that in this way the paper gains clarity. It also allows the reader who is only interested in the results, directly to go to sections 3-5.

p. 14, start of Section 3.2: it is really nice to see that the level of agreement is very similar to the AIRS Team cloud retrievals in Kahn et al. (2008) with a finer breakdown of surface type and ancillary data.

We donŉt completely agree with the statement about the level of agreement with the AIRS cloud data from NASA V5: one important difference is that while the AIRS NASA V5 cloud data agree well for high-level clouds, they have a very large height bias for low-level clouds. This is stated by the Kahn paper: a bias which reaches about 5 km ! Actually, this was the reason for adapting the χ^2 retrieval method to AIRS. Our comparison with the NASA V5 AIRS cloud height was published (Fig. 12) in 2008. Our goal was to build a cloud climatology which is reliable for all clouds. If this is not the case, there will be many cloud type misidentifications. Though the retrieved properties of low-level clouds might be noisier, it was important that their height is not biased, so that they are not confounded with higher level clouds.
Kahn et al. have published a new version of the NASA AIRS cloud climatology, but as unfortunately the team does not yet participate in the GEWEX cloud assessment (though invited), a direct comparison is difficult.

Is the fact that the percentage is slightly higher over ice/snow indicative of a loss of skill at sounding T/q over these surfaces, and Era-Interim is superior? What is different about these profiles over ice/snow? Better detection of inversions and isothermal layers in ERA-Interim?

The frequency of retrievals with good quality decreases over ice/snow, probably also because clouds over these surfaces are more difficult to detect. In addition, polar regions might oft be covered by clouds (especially in SH ocean). We show a map of relative frequency of good quality retrievals of T_{surf} for December 2007, at 1:30AM LT (criteria described in 2.5.1). When only 10% of the time during a month, data are available and the meteorological situation is very variable during the month, the interpolation gets to its limits, whereas ERA-Interim data are always available. ERA-Interim also detects twice more inversions than AIRS (though we do not know which of the dataset is closer to the reality).



Rel. frequency of good quality Tsurf, Dec 2007, 1:30AM

p. 14-16: section 3.3: this section is extremely long and detailed. A lot of it seems consistent with previous paper by the first author. Around lines 31-32 on p. 15 there is one quite interesting point about opaque clouds and a reduced geometrical thickness. Could this be because the IWC is larger in these clouds and thus leads to a smaller difference between the sounders and CALIOP?
This reminds me of a paper by Sherwood et al. discussing these types of discrepancies: Sherwood, S. C., J.-H. Chae, P. Minnis, and M. McGill (2004), Underestimation of deep convective cloud tops by thermal imagery, Geophys. Res. Lett., 31, L11102, doi:10.1029/2004GL019699.

We have substantially shortened this section, also by taking out Fig. 6 and taking out 3 panels of Fig 4 and moving 3 panels of Fig 5 to the supplement. We also tried to be more concise. Compared to Stubenrauch et al. 2010, the estimation of the height at which the cloud reaches a COD of 0.5 is new, though one has to keep in mind that it depends on several assumptions (section 2.4). Concerning the slight drop in difference between z_{cld} and z_{top} for ε_{cld} close to 1, it probably means that for these clouds opacity is reached within a smaller vertical extent, as for those clouds z_{cld} also corresponds to the mean between top and height at which clouds gets opaque. We cited the Sherwood paper in Stubenrauch et al. 2010, where we had already shown that z_{top} - z_{cld} increases with z_{top} -$z_{(app\ base)}$, reaching up to 3 km.

p. 17-21, Section 4: another really long section with figures 8-14 that have a combined total of over 80 sub-panels. A lot of these figures are known from previous papers or are common knowledge. Some of these panels appear to show some redundant information. I would suggest trying to trim this down as much as possible and try and keep the information to the most interesting and novel bits.

We took out 16 panels of Fig. 10-13 and 6 panels of Fig. 14, which we also moved to the supplement. This leaves 5 Figs, and we shortened the discussion. On the other hand, we want to show the quality of the new climatologies, so we have to show some comparisons, even if they might not be novel.

p. 19, lines 25-27: I donőt see why ówhich might have important consequences on radiative feedbacksô should be there. Since the SW and LW budgets are not shown with respect to the different cloud types described in the paper, this is speculative. I would further emphasize that there are many other interesting things about these particular clouds, including the hydrological cycle, not just radiation and its feedbacks.

We agree with this suggestion so we took this part out and shortened the sentences to:

The independent use of p_{cld} and ε_{cld} made it possible to build a climatology of upper tropospheric cloud systems, using ε_{cld} to distinguish convective core, cirrus anvil and thin cirrus of these systems. These data have revealed for the first time that the ε_{cld} structure of tropical anvils is related to the convective depth (Protopapadaki *et al.*, 2017).

p. 20, lines 27-28: Are the authors suggesting that the global cloud amount should be related to the global surface temperature? Is there a previous reference that argues for this? Most studies show a relationship of the patterns of global cloud distributions, height, types, etc. can change with respect to global averaged surface temperature, but Iőve never seen an argument for an average global cloud amount. Also, another point here regarding surface temperature that it did not increase much. If the authors are referring to the alleged óhiatusô, I think that is basically proven that there was no hiatus (a recent paper by T. Karl at NOAA). http://science.sciencemag.org/content/348/6242/1469

Thank you for the interesting article. We just wanted to make the point that global cloud amount stays stable during this period; we have removed the sentence about surface temperature.

p. 21, lines 28-29: what is the justification to relate infrared derived cloud amount to SW reflected radiation? Are there any previous papers that have shown a correlation? The infrared derived cloud amount saturates around an optical depth of 5 or so, but the SW does not. How can the infrared derived cloud products be used to infer consistency with SW results?

We talk here about total CA, which we have shown in section 4 to be consistent with all other climatologies. Also CAH, CAM and CAL are reliably identified, as all discussions in section 4 have shown ! Indeed the effective cloud emissivity saturates at 1 (corresponding to visible COD of about 10), while VIS COD continues to increase. However, the paper of Stephens et al. 2015 is relating the planetary albedo to cloud amount.

p. 22, lines 1-3: how can the CAH be used as a proxy for precipitation rate? Because the ITCZ is narrower in the CIRS data, one can infer a more intense precipitation rate? Iőm not sure I understand the logic used here.

We understand the InterTropical Convergence Zone as the zone with strong convection which then produces large cirrus anvils. The latter stay longer in the atmosphere than the convective towers themselves. It is also seen in all maps that the ITCZ has a strong occurrence of high-level clouds (which are mostly cirrus anvils, see for example (Protopapadaki *et al.* 2017)). Hence, we assume that the ITCZ can be determined by the latitude with a peak in CAH (new Fig. 8). We have partly rewritten this section and hope that the motivation and analysis are easier to follow.

p. 22, first paragraph of Section 5.2: there is no reason to have a basic tutorial on ENSO in the paper. The authors should just get to the results and describe what is novel and delete that part.
we have taken out the introduction and Figure 16 and its discussion.

Figure 3: numbers are too small and blurry for reading
fixed in new Figure 2

Figure 4: why bother with the right column? Werenât these differences previously described by the lead author?
Right column taken out

Figure 6: three figures in a row describing apparent cloud top and biases with CALIOP. Need to emphasize the novel results and parts of figures that support them. The numbers are overlapping on the x-axis at the edges of the subpanels too.
Figure taken out and added quartiles to Fig 4, so that the width of the distributions are shown together with the medians; this makes the discussion more concise

Figure 13: canât tell the difference between open and closed red circle, red square, and red dashed line
fixed in new Figure 10

Figure 14: the seasonal variability in latitude bands is well understood. What is new in this figure? Are there new insights between different instruments and inferences of the seasonal cycle?
Panels with CAM taken out and Figure moved to supplement (new Figure S4); there is nothing new, it is just to show the quality of the new cloud climatologies, compared to other datasets.