Review of Atmospheric Chemistry and Physics manuscript by Stubenrauch et al. entitled

”Cloud climatologies from the Infrared Sounders AIRS and IASI: Strengths, Weaknesses and Applications"

General impression and recommendations

This manuscript presents in detail a cloud property retrieval method, CIRS (Clouds from IR Sounders), which has been adapted to recent hyperspectral sounding datasets from the AIRS and IASI sounding instruments and which has a capability to be applied also to other existing (e.g., CrIS) and planned sounders. Results have been compared to A-Train data (mainly CALIPSO and CloudSat data) showing very encouraging results. Finally, Level 3 products (monthly, seasonal and yearly means) are demonstrated and compared to other existing climate data records. The study also highlights the necessity to account for changing atmospheric CO_2 concentrations in the use of the method for generation of climate data records.

I have no particularly critical or serious points questioning the core method but there are still issues that need to be further explained and commented. Such issues are e.g. related to

- the role of CALIPSO-CALIOP data for tuning the method
- the exact description of the used CALIPSO dataset for tuning and for evaluation of cloud properties
- the consequence of using some unphysical assumptions in the retrieval
- the balance between finding spectral coherence in the solutions and still maintain physically reasonable emissivity differences
- justification of the statement of achieving successful cloud detection down to IR cloud optical thicknesses of 0.1

More details here (and also several minor issues) are given below in the list of Specific Comments.

Quite a lot of editorial suggestions are given and there is also a need to improve several of the figures to make them easier to understand.

A more serious issue is the length of the manuscript. 52 pages of text (36 pages) and figures (16 pages) appear to be too ambitious here and it will be difficult for readers (and even reviewers!) to digest. I suggest that Section 5 on applications is removed (see specific comment number 25) in an attempt to shorten the paper.

In conclusion, I recommend publication of this manuscript after addressing the detailed comments below and after implementing the suggested shortening of the manuscript.

Specific comments

1. Page 1, Abstract, line 19, "to evaluate":

   The term "to evaluate" should be changed to "to design and evaluate".

   You used A-train data to find your 'a posteriori' cloud masking thresholds, right? Then you should be clear in your description that A-train data is not completely independent from your data/method. This is important for the reader to know.

2. Page 1, Abstract, line 23, "coincides":

   To use the term "coincides" here is a too strong conclusion from your results. Figure 6 (lower right panel) clearly shows a rather broad distribution of results where frequencies at the two extremes (0 and 1) are still about 20-25 % of the frequency for the value 0.5 (representing the middle of the defined layer).Therefore you can possibly only state that the cloud height can be "approximated" by the middle of the defined layer. Also "middle" could possibly be replaced by "the mean layer height" to make the description scientifically stricter.

3. Page 1, Abstract, line 27, "apparent vertical cloud extent":

   The explanation here is confusing, indicating that upper level clouds generally have higher cloud emissivities than lower level clouds. This cannot be true. I guess the authors mean something else. Please clarify!

4. Page 2, Abstract, lines 5-8, "response to climate change" + Page 3, Section 1, lines 23-25 and the entire section 5:

   The last sentence in the abstract, the sentence about Section 5 in Section 1 and the entire section 5 could possibly be removed for shortening the paper (see also comment 25!).

5. Page 2, Section 1, line 11, "70 % cloud cover":

   Although this is a widely used and accepted figure for global cloudiness, I would like to point out that a value of global cloud cover cannot be stated without first defining what you mean by a cloud. The figure 70 % is kind of representing clouds which have a significant impact on radiation budgets and it could possibly be relevant if you define that clouds should have at least a cloud optical thickness of approximately 0.2. But if including also the thinnest clouds (often called sub-visible clouds and so far

only observed by high sensitive instruments like CALIPSO-CALIOP) the figure may increase to values well above 80 %.
I think it would be appropriate to at least make a short statement on what clouds are considered when stating that global cloudiness is about 70 %.

6. Page 3, Section 1, line 3: “optical depth less than 3”

   My impression is that the capability is better than that, i.e., the capability of having reasonable cloud optical depth estimations from CALIOP data covers the interval 0-5. Please check that the value of 3 is really justified.

7. Page 7, Section 2.4, line 4, “emissivities larger than 1”:

   I must say that it is quite disturbing to “be forced” to use unphysical values in the retrieval. I understand that uncertainties can lead to this but I am not sure that this is then the best way of handling these uncertainties. Why not restrict emissivities to 1 in the optimization/minimization process when knowing that this is physically correct? I can’t see why your present method gives better uncertainty descriptions of the retrieved cloud pressures than when using a restricted emissivity value. Don’t inconsistencies give rise to new inconsistencies? Please explain and motivate.

8. Page 7, Section 2.4, lines 22-28, “a posteriori cloud detection”:

   The “a posteriori cloud detection” has already been briefly introduced (page 4, lines 7-11). Why repeating this information here? Delete these lines or move part of this to the relevant section 2.5.

9. Page 9, Section 2.4.1, lines 18-20, “ocean cloud amounts larger during night”:

   To find larger ocean cloud amounts at night than during day is found in many regions (e.g. over marine stratocumulus areas). What made you think this was a problem specifically for ERA-Interim? Please explain.

10. Page 10, Section 2.4.2:

    The CO_2 correction appears to be a very relevant change (also visualized nicely in Figure 13. This appears to be one of the most important improvements of the methodology. Should become mandatory in all sounding-based retrievals for climate datasets, in my opinion.

11. Page 11, Section 2.5, general comment on the “a posteriori cloud detection”:

    The methodology appears a bit awkward compared to many other cloud retrieval methods in that cloud properties are first derived and then a determination whether a FOV is cloudy is carried out as a second step. Most common otherwise is that a cloud screening is done first and then followed by a cloud property retrieval. So, could you confirm that after having performed the cloud property retrieval, all FOVs are still assumed to be cloudy? Does it mean that you will always find a solution to Equation 2? You have already mentioned some problems in finding a distinct minimum for low-level clouds (page 7, lines 2-3) but what happens in obviously cloud-free situations?

12. Page 11, Section 2.5, line 16 + lines 20-21, “meaning of spectral coherence”:

    I am a bit concerned about the concept indicating that, for a cloud to be identified, the differences between emissivities in the six infrared channels should be small. In this wavelength region we know that the refractive indices of water and ice, respectively, varies considerably. For example, this is one of the fundamental properties that allows separating water clouds from ice clouds in passive imagery (e.g. as introduced by Pavolonis et al., 2005, J. Appl. Meteorol.). This fact would also certainly introduce considerable differences in cloud emissivities depending on if it is a water or ice cloud in addition to variations in optical thickness or partial coverage within each FOV. So, isn’t there a risk that the demand on spectral coherence is in conflict with reality? Or are you able to find a balanced and optimized method based on reference observations from CALIPSO-CALIOP data and still retain reasonable resulting emissivity differences? I guess that the access to CALIPSO-CALIOP data here is essential since it would be difficult otherwise (e.g. through detailed cloud model simulations) to find an optimal way here. Please comment.

13. Page 11, Section 2.5, line 25, “standard deviation”:

    How do you calculate the standard deviation here? Do you use all values in the “AIRS golf ball” (i.e., 9 values) for the calculation for each wavelength? The current description is not clear enough on this.

14. Page 11, Section 2.5, line 27, “CALIPSO samples”:

    Unfortunately, here you introduce the use of CALIPSO data without having described what data you actually used (this description comes later in Section 3.1). More clearly, it is not obvious to the reader that you will get three CALIPSO samples in the AIRS golf ball. For this, you need to know that you use 5 km CALIPSO data. Because of the importance of A-train data for your method and study, I am of the opinion that you should have introduced them already in Section 2 on “Data and Methods”. Can you consider changing this?

15. Page 12, Section 2.5, lines 18-19, “minimum optical depth”:

    In the introduction section you mention that with IR vertical sounding data “reliable detection of cirrus with IR optical depths as low as 0.1” is possible indicating that this is much better than what can be achieved from other sensors (except from active sensors). I wonder what this restriction in order “to reduce noise” means in this context? Have you estimated further the minimum cloud optical depths being detected after introducing this restriction? CALIPSO-CALIOP offers the possibility to do such in-depth studies.

16. Page 13, Section 3.1, lines 16-19, “CALIPSO <u>and</u> CloudSat data”:

    This requirement should mean (?) that you require that both CloudSat and CALIPSO say it is cloudy. But what about the fact that CALIPSO sees much more of the very thin cirrus clouds being available? Does it mean that these cirrus cases are not included in your evaluation study despite the fact that you several times have emphasized the capability of your method to detect very thin cirrus? Or is it different for studies of cloud amount (as indicated by description in lines 7-15) and cloud top height? Please comment!

17. Page 13, Section 3.1, line 23, “underestimated COD”:

    Just for your information: The latest version of the CALIPSO-CALIOP dataset (version 4.1) gives indeed higher CODs. This change can possibly be connected to what you write here (currently I do not know the details behind this change).

18. Page 14, Section 3.2, lines 2-3, “agreement”:

    I have to ask you to specify better what you mean by “agreement”. There are so many skill scores around so you’d better be strict in describing exactly the measure you use. I guess you refer to what is normally called “Hit Rate” which is the number of correct cloudy AND clear cases divided by the total number of cases.

19. Page 14, Section 3.3, generally on results in Figure 4 (Page 40):

    First, please revise the wording of the caption of this figure. The first sentence here is too complicated and the description should possibly be made more clear (the same is actually true for Figure 5). Also make clear (in all figures) what you mean by “1:30 LT” (AM or PM??).
    The question raised in the previous comment 16 remains: Are thin cirrus detected by CALIPSO but not by CloudSat part of this study or not? If not, what can be said about the quality of these retrieved cloud heights (as compared to CALIPSO data alone)?

20. Page 15, Section 3.3, line 9, “coincides”:

    See previous comment 2.

21. Page 16, Section 3.3, lines 5-24, Figure 7:

    Very interesting and impressive results shown here! Results for medium and high clouds are probably quite superior to those being presented from passive imagery in other CDRs. Only for low-level clouds we still see quite some discrepancies which is understandable for several reasons. This indicates that the best representation of the true vertical distribution of cloudiness in a climate sense could be a combination of sounding and passive imagery data. Do you agree? Maybe you should mention this. Interesting is that problems for low clouds for sounding applications is not showing up very clearly later in Figure 9, except possibly during night for the land-ocean difference. Maybe you should explain why?

22. Page 16, Section 3.3, line 32, “coincides”:

    See previous comment 2.

23. Page 18, Section 4, lines 15-16, “sensitivity of lidars”:

You write that “active lidar is the most sensitive”. Quite true but you haven’t explained whether CALIPSO results in Figure 9 are already “filtered” (so that the thinnest clouds as given by the original CALIOP CLAY product are removed) or not. Has there been any filtering of ‘sub-visible clouds’ (I assume there has)?
This is a relevant question to ask also for the statement in the Conclusions section on page 25, line 25. We need to know exactly what is the used CALIPSO dataset used as reference!

24. Page 21, Section 4, line 4, Figure 14, “Seasonal cycle of cloud temperatures”:

    How come there is a rather large consensus between different methods when studying cloud temperatures for the polar areas (leftmost and rightmost columns) when the spread is very large when it comes to cloud amount (top row of the same columns)? I suspect it is an indication of that cloud temperatures and surface temperatures are very similar here. This implies (in my opinion) that the separation of cloudy and cloud-free areas is indeed not very accurate. So, where is really the truth as regards polar cloudiness?
    Apart from this reflection, I consider Figure 14 as a very nice compilation of global cloudiness and its variation.

25. Pages 21-24, Section 5, “beyond scope??”

    In my opinion, Section 5 feels like out of scope of this study. Although introducing highly interesting topics (especially section 5.2), this work would benefit from being presented as a separate (or companion) publication.
    This manuscript is very, very long and it will put the readers (as it truly has for reviewers!) to a real test when digesting it. I would say that especially section 5.2 on the ENSO effects and its coupling to cloud/radiation feedbacks also requires a different category of expertize for reviewing it with more focus on modelling and studies of climate change and climate feedback effects. Consequently, I have not provided specific comments on this section and I suggest that it is removed for the shortening of this paper.

26. Page 26, Section 6, line 1, “coincides”:

    See previous comment 2.

27. Page 24-27, Section 6, general comment:

    A very comprehensive and good summary of the content of the paper. However, it could be shortened (page 26, lines 14-32) as a consequence of comment 25 above.

Technical corrections

1. Page 1, Abstract, line 11-14:

   The current introductory sentences assumes that the reader already knows about the LMD cloud retrieval scheme. I suggest a slight reformulation to make it less unclear, e.g. like the following

"The Laboratoire de Météorologie Dynamique (LMD) cloud retrieval scheme CIRS (Clouds from IR Sounders) has been adapted to cope with any Infrared (IR) sounding instrument. This has been accomplished by applying improved radiative transfer calculations as well as by introducing an original method accounting for atmospheric spectral transmissivity changes associated with varying CO_2 concentrations".

2. Page 2, Abstract, line 3, "5 % asymmetry":

   Please clarify better what you mean with asymmetry. Does it mean that there is generally 5 % more high clouds in the Northern Hemisphere? I assume this is what you mean (supported also by Figure 10) but you should make it crystal clear for the reader in the Abstract!

3. Page 2, Section 1, line 17, "properties":

   Do you really mean "properties"? I would rather say "cloud detection".

4. Page 2, Section 1, line 32, "determine":

   Like the previous comment, I am not sure about the correct wording here. The word "determine" is very strong and almost indicates that the CALIPSO and CloudSat satellites together are creating/defining the clouds (☺). Rather, you should express that they "are capable of observing the cloud vertical structure".

5. Page 3, Section 1, line 5, "the cloud retrieval method":

   Be a bit more specific, e.g. write "the evolution of the original cloud retrieval method".

6. Page 3, Section 1, line 9, "radiative transfer":

   I think you should write "radiative transfer calculations" or "radiative transfer modelling". To only write "radiative transfer" is too general and (I guess) just a shortening of more correct terms.

7. Page 3, Section 1, line 11, "initial": See 5 above (consider using same notation).

8. Page 3, Section 1, line 11, "radiative transfer": See 6 above (consider using same notation).

9. Page 4, Section 2.1, line 11, "The NASA Science team….":

   I would recommend to start a new paragraph here to increase the readability.

10. Page 4, Section 2.1, line 15, "Susskind et al, 2003":

    I see inconsequent reference formulations on several places in the manuscript. When you make a direct reference to other publications directly in the text (like here) you should (according to my experience) preferably write:

"The methodology is essentially unchanged from that described in Susskind et al. (2003)."

You have done this correctly in other places (e.g., Page 5, line 27). I think you should be consistent here. Use the formulation above when specifically discussing a publication and use reference in parenthesis when not making a direct statement of the referred publication (a "softer" reference).

Check also the following references for the same reason:

- Page 4, line 27
- Page 6, line 5

11. Page 4, Section 2.1, line 20, "shortwave window channels":

    Please write "shortwave infrared window channels" since "shortwave" most often is reserved to define visible channels.

12. Page 4, Section 2.1, line 22, "partial cloud cover":

    A better formulation is probably "under partially cloudy conditions".

13. Page 4, Section 2.1, line 24, "snow or ice":

    Maybe a better formulation is "…snow or ice covered surfaces also provided by NASA L2 data".

14. Page 4, Section 2.1, line 26, "ideology":

    I would suggest using the term "concept" rather than "ideology".

15. Page 4, Section 2.1, line 27, "and allow":

    I suggest replacing this with "which allows".

16. Page 5, Section 2.2, line 1, "12 km":

    Is the 12 km valid for each individual footprint or the 2x2 array?

17. Page 5, Section 2.2, line 9, "the cloud retrieval":

    You should write "the CIRS cloud retrieval".

18. Page 5, Section 2.2, lines 9-10, "retrieved atmospheric profiles":

    Be more specific. You should write "IASI-retrieved atmospheric profiles".

19. Page 5, Section 2.2, line 15, “Therefore”:

    You should not start a new paragraph here if you refer directly to what was written in the previous sentences. Make it also very clear that you never (well, not in time for your development) got access to EUMETSAT Version 6 data otherwise this statement appears rather strange.

20. Page 5, Section 2.2, line 21, “same source”:

    I guess you rather mean a “less instrument-dependent source”?

21. Page 6, Section 2.3, line 1, “proxy”:

    I don’t like the word “proxy” in this context. It indicates that it is a kind of simulation or approximation of the real vertical velocity. The vertical pressure velocity ω is just another formulation of the vertical velocity which arises when you use pressure as your vertical coordinate instead of the standard geometrical height in meters. So, to my knowledge, it’s the “real thing” and not a “proxy”.
    But I guess you refer to the fact that the direct calculation of ω is difficult without making approximations. The most common here is the geostrophic assumption leading to the so-called “ω-equation”. In this sense, I guess you may be correct in interpreting it as an approximation. But still, present day NWP models are capable of calculating ω so I just wonder what value you are using here? On the other hand, the approximated value at the 500 hPa level is probably quite accurate anyway (conditions here are largely quasi-geostrophic on the large scale) so perhaps this discussion is less important. Anyway, give it a thought.

22. Page 7, Section 2.4, line 12, “arise”:

    Maybe reformulate to “these cases occur in about 7 to 15 % of all cases”?

23. Page 8, Section 2.4.1, line 14, “less than ..?..”:

    Strange formulation. You’d better write “0.99 for wavelengths less than 10 µm and 0.98 for wavelengths larger than 10 µm”.

24. Page 13, Section 3.1, line 6, “spatial resolution CALIPSO”:

    Shouldn’t it be “5 km x 0.3 km”? I thought the basic FOV of CALIOP was 300 meter.

25. Page 15, Section 3.3, Figure 5 (Page 41):

    I suggest that you try to include some additional explanatory features or legends in the figure (e.g., legend with the three coloured dots explained). To look for all explanations in the caption is not very reader-friendly. Try to speed up the correct interpretation of figures with the use of more graphical legends or marks. This remark is probably valid for many other figures in the manuscript.

26. Page 15, Section 3.3, line 27, “Considering…”:

I suggest starting a new paragraph here in order to avoid too long chunks of text (unnecessary tiring for the reader).

27. Page 15, Section 3.3, line 28; Figure 6 (Page 42):

    In the caption you describe one of the curves as “broken line”. I am not sure whether this is the most common way of describing such a curve. More often the term “dashed line” is used. Consider changing to “dashed”. This suggestion is valid for many other figures in the manuscript.

28. Page 16, Section 3.3, lines 28-29, “height of COD”:

    Semantically, it sounds strange (or even incorrect) to express COD as representing a height. Of course, I understand what you mean but it can actually be misinterpreted. Since you have already defined $z_{COD0.5}$ why not use this terminology here, e.g. “the retrieved cloud height exceeds $z_{COD0.5}$ for optically thin clouds while it is lower than $z_{COD0.5}$ for optically thick clouds”.

29. Page 20, Section 4, line 17, “three CIRS datasets?”

    It is not obvious what three datasets you mean (not explained in text)! Please clarify.