

## ***Interactive comment on “Ice hydrometeor profile retrieval algorithm for high frequency microwave radiometers: application to the CoSSIR instrument during TC4” by K. F. Evans et al.***

**K. F. Evans et al.**

evans@nit.colorado.edu

Received and published: 18 August 2012

Author responses to review comments on  
“Ice hydrometeor profile retrieval algorithm for high frequency microwave radiometers: Application to the CoSSIR instrument during TC4”

K. Franklin Evans

C1811

### **Anonymous Referee #3**

I thank this referee for the useful suggestions, all of which I implemented.

*The article describes with a wealth of details all the parts of the algorithm and the assumptions used to develop it. I found this very useful for comprehension and comparability with other retrieval algorithms.*

*A drawback of the manuscript is a far too long introduction. The idea to give an overview of the algorithm at the beginning of the paper to drive the reader in the two following sections and appendices is good, but I would suggest dedicating it a new section after the introduction. Tentatively, the insertion of general information regarding the source of the a priori information into section 1 would make possible to make a new section with the algorithm outline from page 3124, line 21.*

There is now a new section 2 with all five paragraphs of the retrieval algorithm overview from the previous introduction.

*I appreciate the use of the appendices where details regarding specific aspects of the algorithm are given. I would suggest including in an appendix also the discussion on CloudSat reflectivity below -26 dBZ now included in section 2. Furthermore the reading of section 2 would be simplified if it were subdivided into subsections.*

These are especially good suggestions. The method for simulating the CloudSat reflectivity below -26 dBZ is now described in a new appendix. There is still a paragraph in the original section that describes the need for the reflectivity threshold, refers to the procedure described in the appendix, and references Fig. 2.

This section has been divided into six subsections with titles that correspond to the left side of the flowchart (Fig. 1). This really helps organize and clarify the section.

*Technical corrections:*

C1812

Page 3121, line 21: "distribution that IS much closer to log-normal".  
3126, 5: put a reference to appendix B4 for the description of the melting model  
3134, 9: "The relative humidities are converted TO water vapor mixing ratio q"

I made these minor changes.

*Fig. 10: colours used for hail and hexagonal aggregates are not distinguishable*

I changed the colors for the shapes in Figs. 9 and 10 to blue (hexagg), green (sphragg), red (snowagg), and black (hail).

*Fig. 13: add the date in the caption or in the title*

I added the date to the caption.

### **Anonymous Referee #1**

*I would strongly recommend that the authors make the paper shorter and also make extra effort to explain the method in a more didactic way for publication.*

If the retrieval algorithm is to be described completely, it is not possible to make the paper both shorter and to describe the method in a more didactic way for publication. The algorithm is complex and fully developed, and therefore requires many pages to describe it completely. Only essential details have been included, and much other detail has been left out. The paper is designed for readers of all interest levels, so that a reader not interested in the algorithm details can read the overview section, the results section, and the conclusions summary section. The two major appendices describe preparation of the algorithm microphysical a priori inputs and realistic ice particle scattering tables, which are separate from the retrieval algorithm proper. These appendices are useful for some innovations (such as the snowflake aggregation and melting

C1813

models) and as an example of preparation of a priori information for ice cloud retrieval algorithms in general.

*I suspect that my level of understanding is too limited to fully review this paper; this is a long paper which requires an expert in the field to understand this specific retrieval approach. The description of the method is complicated and most of the time we lose the point. I'm wondering why fig 1 is not mentioned more in the text as I am sure that would help the reader to understand.*

The detailed algorithm description sections are designed for those few developers of ice cloud retrieval algorithms, and parts are more comprehensible with knowledge of various Bayesian solution methods.

The subsections in the new sections 3 and 4 organize the material and should help from losing the point in the details. There are now six subsections in the section on generating the a priori CDF/EOF file and four subsections in the retrieval process section. The subsections correspond reasonably well to the flowchart, so it should be easy to see where each subsection is in the flowchart, if a reader so chooses.

*I'm curious about the choice of the database used for building the a priori information. What is the reason for not using cloud properties retrieved by CloudSat-CALIPSO?*

This is partly a matter of philosophy of retrieval algorithm design. I disagree with using other ice cloud remote sensing retrievals, along with their assumptions, as a source of a priori information. I prefer to use the measurements (e.g. CloudSat reflectivity) in a way that is consistent with other a priori information from in situ microphysical measurements. I think using in situ microphysical measurements is a necessary connection to actual ice cloud microphysics, which would be excluded by relying on the output of other remote sensing methods. For a Bayesian method it is also necessary to use a distribution of IWC and  $D_{me}$  that are consistent with the radar reflectivity, not single values from a retrieval. Otherwise, a fixed (or narrow) relationship between IWC and

C1814

$D_{me}$  will be built in.

*Regarding the final evaluation using CRS radar, I think that the lidar was also available and therefore the combined radar-lidar retrievals would be more accurate for evaluating the performance of the algorithm.*

The passive microwave measurements are mostly sensitive to the vertical integral of ice cloud properties. Most of the columns for which CoSSIR is able to retrieve ice water path above its noise level are not penetrated by the lidar. Therefore, combined radar-lidar retrievals would offer little benefit because the lidar cannot sense the region that contains most of the ice mass. Furthermore, I prefer to evaluate a retrieval algorithm with data (as I do by comparing CoSSIR retrieved and CRS reflectivity), rather than assuming another retrieval method is the truth.

*Note that the integrated radar backscatter is mainly dominated by large reflectivity so it might be inadequate to characterise the particle assumption; only the largest particles might be represented.*

There certainly are limitations of using CRS integrated backscattering for validation, one of which is that it has similar particle size sensitivity to the low frequency channels of CoSSIR. On the other hand, lower values of integrated reflectivity where the CoSSIR information comes from the higher frequencies, is a more independent test. An in depth validation of the algorithm (with many independent data sources) is beyond the scope of this already long paper.

*L99 : "A Bayesian pdf is not about how frequently a parameter has a particular value, but instead a pdf specifies how likely the parameter is to have particular values.", I am not sure I understand what you mean here. Maybe the use of "frequently" is misleading me.*

Since the first clause did not help the Bayesian explanation, I simplified this sentence to "A Bayesian pdf specifies how likely the parameter is to have particular values."

C1815

*L145 "Since optimal ... substantially underestimated." I have probably misunderstood but need to use it to estimate errors in your retrieval, is that correct?*

Unlike some optimal estimation implementations, my ice cloud retrieval algorithm does not oversimplify the radiative transfer or fix atmospheric parameters that should be allowed to vary. Nor does my approach require assumptions of Gaussian (or log-normal) probability distributions for the prior pdf, as optimal estimation does. When Monte Carlo integration (MCI) cannot be used, my algorithm uses optimization to maximize the Bayesian posterior pdf. Unfortunately, the only practical way to then obtain uncertainty estimates is to make the local Gaussian assumption of optimal estimation. My concern over the validity of the local Gaussian approximation for error bars led me to implement the Markov chain Monte Carlo (MCMC) method. Fig. 14 does show that MCI agrees much better with MCMC than the local Gaussian approximation does for the IWP uncertainty.

*L 195: "AMSU-B channels", it would be useful to add a reference.*

The research being described is from Seo and Liu (2005), which give a good description of AMSU-B. I did add the five AMSU-B channel frequencies (89, 150, 183.3±1, ±3, ±7 GHz), since that is the most relevant aspect here.

*L 202: "Odin-SMR limb-sounder", any reference?*

A reference about Odin-SMR can be found in Rydberg et al. (2009). The relevant information, that it is a limb sounder and Rydberg used 501 and 544 GHz, was already given.

*L235: "A priori profile information is obtained from CloudSat (Stephens et al. , 2008) project files of radar reflectivity, CALIPSO lidar cloud fraction". Don't you think it is a pity not to use the radar-lidar product instead? Furthermore you have at least 3 products with different assumptions and approaches that you could use to build your CDF/EOF. I agree with the fact that IWP will be strongly dominated by the radar measurement but*

C1816

*in the radar-lidar common region you should have a better retrieval.*

See my comment above on using other retrievals for a priori information.

*L 285 : I know that you give all the details in App B but you could specify the ice particles used in your method in this paragraph.*

This paragraph defines two parameters,  $D_{me}$  and  $D_e$  dispersion, that quantify the ice particle size distributions. The definitions do not depend on ice particle shape, as should be clear from the first sentence “Ice particle size distributions are defined using the particle mass as expressed by the equivalent mass sphere diameter,  $D_e$ .”

*Section 2:*

*Please see my global remark concerning the use of radar-lidar data*

I addressed this before.

*L341-342: You could give some references for these products.*

It would be awkward to put citations in that particular location, which is referring to files in the CloudSat archive. I did have a reference (Mace et al. 2009) for GEOPROF-LIDAR where I describe how it is used. I added a reference to Marchand et al. 2008 for GEOPROF where it is mentioned later in the section. I originally decided not to reference that paper because it is about the cloud detection in GEOPROF, and I do not use the GEOPROF cloud detection.

*L348-349: What do you mean by "interpolated to the layer interfaces"?*

I tried to clarify by adding “interpolated to the specified levels” after indicating that levels are the same as “layer interfaces”, meaning particular heights in the profile and not a range of height.

*L352: I guess you are referring to the clutter contamination? Am I correct?*

C1817

That is correct. If a reader knows what is meant by clutter contamination, then “CloudSat reflectivity within three range gates of the surface elevation is not used”, should be self-explanatory.

*L358: You mention supercooled droplets - where does this information come from? Could it be detected by the radar? Or does it come from the in-situ data?*

The first paragraph of the section states that the secondary source of a priori information is “obtained from in situ aircraft probes that describe relationships between ice cloud parameters, liquid cloud parameters, and relative humidity.” Later this section mentions that the radar is only used for ice and melting/melted particles (droplets are undetectable in the presence of ice particles). The full description of how the supercooled droplet probability distribution is derived is in Appendix A.

*L367: You should mention that you use DDA calculation . . . Once again, I know it is in the appendix.*

This line was “Appendix A describes the analysis of in situ cloud probes from TC4 to generate the a priori information input to the CDF/EOF generation program. This has nothing to do with the scattering calculations, which are mentioned in the next paragraph, which refers to Appendix B.

*L377: I don't understand why you don't use the radar mask available in the 2BGEO-PROF product. Actually CloudSat is more sensitive than what you suggest, see publication from Tanelli et al.*

The original sentence was “Visual inspection shows that a CloudSat reflectivity threshold of  $-26$  dBZ for 500 m thick layers is required to nearly eliminate spurious cloud detections due to receiver noise.” I added “This threshold needs to be higher than the nominal CloudSat sensitivity of  $-30$  dBZ because the probability distribution of the receiver noise power has considerable width.”

C1818

The radar reflectivity simulation method (described in the new appendix) for reflectivity values below the threshold makes the radar cloud mask issue irrelevant.

*L345: Not sure I understand the purpose of Fig 2 (also could you put the x and y axes on the figure, is x axis the latitude, time?), you want to show that you can simulate the clouds not detected by CloudSat? It is difficult to tell whether it works or not as there are no reference values. Maybe you should use the extinction retrieved from the lidar and converted to reflectivity to check if there is a good agreement or not. I know there might be errors in the extinction-reflectivity conversion but it would be a good indicator as you are in the Rayleigh regime (low Z).*

I added the following sentence to motivate the radar reflectivity simulation method: "This allows the prior pdf to have lower values of IWC and  $D_{me}$  than would be produced from the CloudSat reflectivity alone." The figure is simply to show what the simulation method does. The vertical axis is labeled with 0 km and 15 km. The caption explains that there are 1600 CloudSat columns from two separate orbits; so showing latitude or time would be difficult. The ice cloud retrieval algorithm only uses vertical statistics, so each column is treated independently.

It is true that there is no validation of this method. That would be beyond the scope of this paper (I suppose a whole paper could be written about this simulation algorithm). The validation would have to be statistical because the actual values are made up, and only the reflectivity profile statistics matter to the retrieval algorithm.

*L444: Which attenuation are you referring to? Attenuation due to ice, liquid, gas? To me Ice attenuation at 94GHz is very small.*

I changed the description to "(where  $A$  is the ice/melting particle radar attenuation coefficient in  $\text{dB km}^{-1}$ )." Actually, attenuation, which includes scattering and absorption, can be significant at 94 GHz for high IWC and large particles. The paragraph states that the radar reflectivity/temperature table applies to both ice particles ( $T < 273 \text{ K}$ ) and

C1819

melting particles ( $T > 273 \text{ K}$ ).

*L487-488: Are you sure this is mainly a problem in the rain layer? I thought that Battaglia et al. showed that it also happens in convective ice cloud, with a kind of phantom effect in the reflectivity (ie, what you see after a few kilometres below the cloud top is only due multiple scattering).*

Battaglia et al. (2008) was not a good reference for multiple scattering of the CloudSat radar in the ice layer because it focussed on rain retrievals and path integrated attenuation to the ocean. I changed the reference to Battaglia et al. (2011) ("Multiple scattering identification in spaceborne W-band radar measurements of deep convective cores"), which definitely changes my conclusions on multiple scattering. I changed the sentence to the following:

"In deep convection multiple scattering increases CloudSat reflectivity above the single scattering values assumed here. Battaglia et al. (2011) estimated that in tropical deep convective cores, multiple scattering becomes important ( $>3 \text{ dBZ}$ ) below about 9 km altitude. Since correcting for multiple scattering accurately is very difficult, we simply note that the effect is to over-estimate the (single scattering) 94 GHz reflectivity when the reflectivity is high ( $>10 \text{ dBZ}$ ), thereby widening the a priori IWC and  $D_{me}$  distributions."

*L516 to 543: I found this paragraph quite difficult to understand.*

I split this paragraph, now alone in a subsection, into three paragraphs: one on the CDFs, one on the EOFs, and one on the CDF/EOF file. I added equations to help explain the transformations involved in making the correlation matrix from which the EOFs are derived, and I restructured the EOF paragraph.

*L576: Fig 6 is quite difficult to interpret. What are the axes for each block?*

I added more explanation in the figure caption: "Each element of the matrix represents a pair of parameters and layers, with increasing height of the levels/layers within each

C1820

parameter block. The  $T$  and RH levels from the surface up are shown as dots in Fig. 3, while the hydrometeor layer centers are shown as dots in Fig 4.”

*L579: I think that  $D_{me}$  and IWC are by construction highly correlated ( $D_{me}$  is weighted by IWC).*

The manuscript stated “IWC and  $D_{me}$  of the same layer have a reasonably high rank correlation”, and the fact that the correlation matrix shows they are not 100% correlated indicates that the prior pdf allows a range of particle size (and hence number concentration) for a given IWC.

*L582: "Although there seems to be a lot of information in the covariance matrix, and hence the EOFs, it should be noted that there is only one number to represent the relationship between any two variables, which is a tiny fraction of the information contained in a joint probability distribution." Could you explain the nature of the rest of this information and how you can use it, please?*

This request seems beyond the scope of this paper, since the retrieval algorithm only uses the single number (correlation) instead of the hundreds of values it would take to completely describe a discrete joint pdf, e.g.  $p(IWC, D_{me})$ . I think the reader will understand if he knows what a joint probability distribution is, and if not, the point is not important for understanding how the algorithm works.

*Section 3: In this section you should refer more to the Fig 1, that could help. I found it very difficult to understand.*

I added a subsection header, “Atmosphere profile generation”, after the first paragraph. The flowchart in Fig. 1 is high level, so that the subsections fairly obviously relate to the parts of the flowchart. I suspect that it is the details, not the high level structure, that make it difficult to understand.

*L769-782: Why don't you go for an adjoint method instead, then you would need to*  
C1821

*compute the Jacobian?*

I don't understand this comment. The Rodgers  $\mathbf{K}$  matrix is the Jacobian. The manuscript states “The  $\mathbf{K}$  matrix is calculated using the adjoint of the radiative transfer for each channel ( $F_j(\vec{x})$ ) and the adjoint of the a priori function  $\mathbf{G}(\vec{\xi})$  that calculates the geophysical variables from the control vector”, so it should be clear that the adjoint method is used.

*L783: Sorry I am a bit confused here., Is it consistent with what you claimed in the introduction regarding the advantage of not using the optimal estimation?*

The introduction points out some shortcomings of optimal estimation, both in the inherent formulation and in sloppy applications of it. Using the local Gaussian approximation for the retrieval uncertainties when using the optimization solution method is one aspect of the current retrieval algorithm that shares the shortcomings of optimal estimation. I could not think of an efficient method for the retrieval uncertainties apart from the local Gaussian approximation. Most pixels are retrieved with Monte Carlo integration, which does not have this shortcoming. The comparison with the Markov chain Monte Carlo method indicates that the local Gaussian approximation is not as accurate for the IWP uncertainty estimates as MCI is. Other limitations of optimal estimation, such as assuming a Gaussian prior pdf, are overcome with the current approach. Thus, the fact that I sometimes use a part of the optimal estimation framework in the retrieval algorithm is not inconsistent with the analysis of optimal estimation in the introduction.

*L809: What is the purpose of the Markov Chain Monte Carlo solution Method? What is the link with the previous subsection? Sorry to ask this but it is not very clear to me and I actually needed to reach the conclusion to understand their roles.*

I added a sentence at the beginning of this subsection to motivate using the MCMC method: “The Markov chain Monte Carlo (MCMC) technique is developed as an optional solution method to check the accuracy of the optimization/local Gaussian ap-

proximation method.”

*Section 4: L898: "When CRS radar reflectivity is input to the retrieval it is averaged to 20 layers from 5 to 15 km and has a multiplicative uncertainty of 0.4 (about 1.5 dB)." Could you explain the multiplicative uncertainty? What is the original vertical sampling of the CRS radar? L932: Do you assimilate radar reflectivity here?*

A multiplicative uncertainty in radar reflectivity factor (units of  $\text{mm}^6/\text{m}^3$ ) is an additive uncertainty in dBZ. I changed the text to “The profile retrieval algorithm is also used to operate on CRS reflectivity profiles alone or with CoSSIR nadir data. When CRS radar reflectivity is input to the retrieval it is averaged from 75 m resolution to 500 m (20 layers from 5 to 15 km) and has a multiplicative uncertainty of 0.4 (an estimated calibration uncertainty of about 1.5 dB).”

*L939: Is there any reason for choosing this value?*

The sentence is “Somewhat arbitrarily a retrieved probability threshold of 0.95 is chosen to indicate cloud.” I used the words “somewhat arbitrarily” to indicate that I chose the value by eye to avoid having the retrieval dominated by noise. I’ll note that 95% probability is an often used threshold for detection of something real.

*L946: It might be due to the fact you are looking at fractional error. If I am correct:  $\Delta \text{IWC} = (\Delta \text{IWC})/\text{IWC}$ , so if you increase IWC the relative error decreases.*

Yes, that is certainly part of why the retrieved IWP error bars tend to be smaller for larger IWP.

*In fig 7 and 8, maybe you could over plot values retrieved from radar or radar-lidar retrievals. Radar+lidar would give cloud fraction and IWP. Difficult to distinguish the dots from bars.*

The purpose of these figures is to illustrate the output of the algorithm, not for validation.

C1823

For validation, I prefer to compare with measurements, as I do in later figures, rather than assume other retrieval methods are the “truth”.

*L 996: "a burn in fraction of 0.5" what does it mean?*

Burn in fraction is implicitly defined in the MCMC subsection.

*L1025-1030: Are these values computed in dB or linear scales?*

The text states that the vertically integrated 94 GHz backscattering is computed and presented in dB.

*L1046-1048: To me these are very large errors.*

I suspect that the overall profile retrieval errors seem large to an active remote sensing person, but it should be kept in mind that these are profiles retrieved from a passive radiometer. Actually 2 dB rms errors in reflectivity above 9 km is fairly good for any ice cloud retrieval algorithm. How well can lidar alone retrieve radar reflectivity in the narrow overlap region?

*Regarding Fig 17, would it be possible to add a panel with the difference between CRS and CoSSIR? A scatter plot of the difference as a function of Z could be interesting to illustrate the comments in the text.*

I added a panel to Fig. 17 with the difference between the CoSSIR retrieved and the CRS reflectivity. A scatterplot as a function of height is difficult to execute and not a quantitative way to summarize the difference in reflectivity. Instead I added a graph of the rms reflectivity error as a function of height for two CRS reflectivity thresholds (-5 and 5 dBZ) and three averaging layer thicknesses (1, 2, and 4 km).

*Fig 19 and 20, why do you retrieve IWC where the radar can't detect any cloud above ~13 km (see fig 17)? It is not surprising if you only use the microwave but with the*

C1824

*radar, this is weird ... Maybe the vertical scale is wrong. Could you add the y axis please?*

These figures do have a vertical scale showing that the ice cloud properties are retrieved from 4.5 km to 15.0 km (the text earlier states that 21 layers of IWC/ $D_{me}$  are retrieved from 4.5 to 15 km).

Retrieving values below the instrument noise level is not weird to a Bayesian. The Bayesian retrieval algorithm returns a posterior pdf for all specified layers, and the posterior pdf is summarized by its mean and standard deviation. The CRS retrieved IWC along with its error bar is still a valid retrieval, even if the radar reflectivity is dominated by receiver noise.

*L1080: How different is retrieved IWC- $D_{me}$  relationship from the in-situ (a priori)? You could add a plot to illustrate this.*

This sentence is "The IWP and  $D_{me}$  maps have somewhat different patterns, showing that IWP and  $D_{me}$  are partially independent." The a priori pdf is defined for IWC and  $D_{me}$ , not their vertical integrals, and depends on height, temperature, relative humidity, etc. So it is not straightforward to define the IWP -  $D_{me}$  a priori relationship, which, of course, is a 2D joint pdf. There is also not enough CoSSIR pixels to accurately calculate this joint pdf.

*The conclusion is very long and maybe it would be worth reducing the summary.*

The conclusion section is a little over one page in AMT format. I designed the conclusion section for readers to be able to read only the introduction and conclusion sections to be able to skip the details of the algorithm. The retrieval algorithm is summarized in only one paragraph because it refers to the new overview section and the flowchart figure. Two long paragraphs give the pros and cons of the retrieval algorithm, and are therefore not part of the summary. The results section is summarized in two paragraphs (I did shorten the last one a little.)

C1825

*L1134: I guess it should be "updrafts"*

Yes, that was a typo.

*L1199: Maybe a reference here? Are you thinking about TRMM or GPM?*

I was thinking generally. TRMM would be good for precipitation profile a priori information, but it would be much better to have both precipitation and cloud radars looking at the same columns.

---

Interactive comment on Atmos. Meas. Tech. Discuss., 5, 3117, 2012.

C1826