Review of « Retrieval of microphysical cloud parameters from EM-FTIR spectra measured in Arctic summer 2017», by Philipp Richter et al.

General comments

This study presents a retrieval algorithm for cloud physical properties (liquid and ice optical depths, liquid and ice effective radius) based on Fourier Transform InfraRed spectra in the range 770.9 – 1163.4 cm^{-1}, called TCWret. From these parameters, cloud water content, total effective radius and total optical depth are computed. The algorithm is based on radiative transfer simulations performed with LBLRTM and DISORT, coupled by LBLDIS. It also makes use of Mie calculations for single scattering properties of spherical particles, and of the database of Yang et al. (2013) for ice crystals single scattering properties. The algorithm is tested on synthetic spectra from Cox et al. (2016) that are perturbed in various ways to estimate the robustness of the retrieval. It is also applied on spectra measured on the Polarstern in the Arctic during summer 2017. These retrievals are compared to Cloudnet measurements acquired simultaneously on the Polarstern, that combine, lidar, radar and microwave radiometer measurements. The authors also demonstrate that the algorithm behaves as well as the previously published CLARRA algorithm, and allows to retrieve cloud optical thickness for thin clouds ($\tau<6$), thus complementing other instruments that are more sensitive to optically thicker clouds. The retrieval of the cloud phase is not very reliable, hence the focus is on total condensed water. Likewise, the ice crystal habit is not retrieved by the algorithm. The algorithm also perfoms worse when a cloud is spread vertically compared to when it's geometrically thin, because of the assumption of geometrically thin cloud in the algorithm.

The manuscript is easy to read at first sight. However many details are missing when trying to better understand what the authors have done. Also, it is not particularly original in the sense that TCWret is very similar to existing algorithms. The sensitivity study is interesting but lacks of scientific perspectives. The observed spectra are not sufficiently explored to demonstrate the capability of such new instrument/algorithm. More generally the authors should stress more clearly why they developed a new algorithm while they refer to existing ones which seem to produce very similar results. Only with significant improvements on the method description, additional context and perspectives the paper might be worth being published in AMT.

Specific comments

1) Although the methodology is standard, many details are missing, which either points to deficiencies in the method, or prevents the reader from reproducing the algorithm. This is extensively detailed in the technical corrections but the most critical points are : poor treatment of ancillary data uncertainties in the retrieval, insufficient details about S_a and S_y, limited usage of the matrix A (presented but not used further) and of the retrieval uncertainties.

2) To expand on the previous point, there is not sufficient discussion about the uncertainties in general. From the ancillary data used in the algorithm, to those on the retrieved parameters. Likewise, many figures are given for correlation coefficients, without any hint to their meaning or statistical significance. As a consequence, the reader does not really know what to expect from the numerous sensitivity tests, neither how to interpret the results. Statements like « are in agreement » do not help much.

3) The algorithm, which is the core of the study (given there is nearly no interpretation of the data acquired onboard Polarstern), is probably not sufficiently original, or different from existing algorithms (primarily CLARRA) to justify a dedicated paper. Unless the authors more strongly point why using CLARRA on their data would not have been possible or would have been worse than developing a dedicated tool.

4) More generally, the lack of physical interpretation of the results, in particular when it comes to real data, is frustrating. I understand that AMT may publish papers on pure methodology, but still some extended interpretation of the strengths and limits of the present algorithm would have been appreciated.

Technical corrections

title : the title suggests that the focus of the study is about the retrieval onboard Polarstern, while the paper is more about the algorithm validation. To be more consistent with the title, the paper should discuss in more details the retrievals performed in the Arctic. Conversely the title could be updated. Also, EM from “EM-FTIR” is never described.

l.2 : this « therefore » suggests that the reason for the study was the fact that FTIR is sensitive to clouds. I’d suggest to turn this into a more geophysical objective.

l.6 : CWP should not be italics I guess, neither “total” under “r”

l.7 : it is not clear here, neither in the text, why you specify that radiances under 600 cm^{-1} are not used. Do you mean that usually such retrieval requires these longer wavelengths ? Does the FTIR acquire something outside the range 770.9 – 1163 cm^{-1} defined earlier ? Please clarify

l.8 : is the validation the comparison to Clounet products or to synthetic testcases ? The structure of the abstract is not very clear in that sense

l.12: a correlation is not sufficient to describe a retrieval capability, because you could have a scaling factor in between. An RMSE or a relative error would be more relevant. Also it is not clear if these testcases are meaningful or irrealistic (no noise, very simplistic cloud), making these figures poorly informative.

l.14 : same comment for the correlation coefficient

l.16 : I believe the interest is for night conditions, otherwise powerful algorithms work well in the solar spectrum. If this is the case this should be emphasized in the abstract. More generally, the methodology should be put in a broader context of instrumentation and inversion methods to demonstrate its added-value.

l.19 : radiation budget of what ? Surface, atmosphere, Earth system ?

l.20 : “ with respect” sounds awkward

l.22 : radiative forcing is loosely defined (surface, TOA, upward or downward convention). Maybe say clearly “ are assumed to cool the Earth”

l.24 : would it be useful to include ice clouds here instead of focusing on LWP ?

l.26 : LW flux is loosely defined. Do you mean broadband, or at any wavenumber ?

l.27 : please clarify where up- and downwelling fluxes are defined (below the cloud downward and at cloud top upward ?). Multiple scattering between what and what ? Inside cloud ? Between cloud and surface ? Also, the limits should be in terms of optical thickness, not in terms of LWP, no ? So maybe state that these LWP ranges are approximative and correspond to average clouds.

l.30 : larger → large ?

l.35 : quantify “high range”. operate continuously

l.36: LWP < 15 g m^{-2} ?

l.37 : maybe tell why MWR are only sensitive to liquid water. Maybe precise the frequency of such MWR.

l.38 : there exists a variety of instruments to study clouds, so maybe give more motivation to use FTIR instead of active instrumentation, VIS-NIR radiometers etc.

l.45 : maybe give a relevant reference for cloud studies with AERI. What do you mean by “built in particular” ? What’s the difference with absorption FTIR ? If relevant, stress the importance of performing emission observations in your case.

l.46 : maybe specify both spectral resolutions

l.49 : is the algorithm related to the campaign, or to the instrument ? Could this algorithm be applied to other observations, potentially taken with another FTIR at a distinct resolution ?

l.54 : please give some insight into the differences between MIXCRA and CLARRA, here or later. Do both algorithms use far-infrared channels ? Also, if relevant, emphasize more clearly that you achieve similar results without far-infrared channels.

Figure 1 : green~~s~~ dots

l.69 : FTIR acronym used earlier

l.71 : is this resolution before or after apodization ? What apodization is performed ? “ so that a maximum...”

l.72 : “aperture”. Not clear what you mean. There will always be contribution from the instrument itself. Maybe clarify that the field of view is not limited by the aperture but by the detector itself, if true. Maybe specify the field of view and focal length.

l.73 : “detector”. What is MCT-detector ? Is there a window on the container to isolate the instrument but let the radiation get in ?

l.76 : the blackbodies. Please detail the absolute error of the instrument, or at least the temperature stability of the blackbodies.

l.77 : How long does it take to acquire an interferogram ?

l.84 : figures 6, 7, 8 do not have this unit ($^{-1}$ missing)

l.92 : detail what are aerosols and cloud properties. Only LWP or CWP from OCEANET? Do not mention LWP if ultimately you get vertical profiles of ice water content and liquid water content. Is effective radius retrieved for liquid and ice as well from Cloudnet approach ?

l.95 : acronym already defined. “inspired” is vague. Please review the main differences with those tools, to highlight that you’re really developing something new. This is crucial for the relevance of this paper in AMT.

l.104 : does this mean that no uncertainty is attached to these gaseous profiles?

l.105 : “inhomogeneous”. What does this mean ? Vertically inhomogeneous ?

l.107 : I guess DISORT does not handle effective radii, but single scattering properties (phase function and g). Can you clarify this.

l.109 : the role of LBLDIS is not clear

l.110 : please detail the characteristics of this distribution, and support the choice made with references or observations. Do you work with phase functions or simply asymmetry parameters ? Please clarify all these technical details.

l.112-116 : it is not clear what you take from LBLDIS and what you have computed on your own. Do the details on refractive indices refer to your work or to that of Turner (2014) ? I don’t understand if you used the 1975 or 2005 database.

l.117 : ice water droplets is misleading for crystals that are practically not spherical…

l.118 : why not using the updated Warren (2008) database ?

l.119 : please detail how you computed SSPs from Yang et al. (2013) database. Using the same size distribution as for liquid clouds ? Also, please state here what choice is made in the inversion for ice crystal habit.

l.120-121 : not clear what “chosen … and modified” really means.

l.121 : it is not clear what offset, slope, and curvature are. Maybe provide a figure to illustrate this quantity on a given microwindow. Please explicit your observation vector. Does it comprise radiances, or these 3 parameters for each microwindow ?

l.123 : not clear. What do you mean regarding the trace gases ? That they are absent within the microwindow, or that they can still alter the retrieval in your case ?

l.124 : the information on cloud boundaries, as it does not appear in the state vector, should be provided beforehand. Also at which wavelength do the optical thicknesses correspond ?

l.131 : check the “n” indices, some are missing. Also $F(x_i)$ are radiances, while I would expect the 3 parameters defined above. This should be clarified if radiances or offset/slope/curvature are used.

l.132 : the S_y matrix only includes observational errors, but no impact of the uncertainties on the ancillary information. How do you justify this ? What would be the impact of adding such uncertainties in S_y, as is commonly done in such retrieval techniques ? Also, how do you argue that S_y is diagonal ? There is very likely correlation among channels of the instrument. That should be accounted for.

l.133 : how is S_a built ? These details are needed.

l.140 : maybe precise how the averaging kernel matrix should be to be perfect, or how you can quantify the quality of the retrieval from this matrix. Is this matrix used further ? If not remove this information, although it can be informative if you add extra information on this.

l.150 : what is D_n ? Diagonal unitary matrix ? Then why index n?

l.152 : what about the initial guess for r_{eff} ? How many different optical depths ? What increment in this initial guess research ?

l.158 : what's the underlying assumption on extinction cross section of Q_{ext} ? Is it justified ?

l.161 : not sure 3 decimals are necessary for a quantity that may vary with temperature

l.163 : "ice droplets of any shape" sounds awkward because droplets generally means spherical

l.170 : please detail why these total water quantities, which are built from quantities expected to be non reliable, should be more reliable.

l.178 : the covariance errors between retrieved state parameters is not discussed, while it is a very useful information

l.179 : please detail this error propagation method as it is not trivial, in particular because of the covariance errors.

l.179 : what residuum ? This alternative method is not clear.

l.180 : what residuum are used to compute S_y ? It cannot be residuum at the end of the retrieval because this quantity is needed at the beginning of the iterations

l.183 : what about the radiative transfer code used to compute these spectra ? Is it consistent with your retrieval algorithm and the assumptions made on cloud particles SSPs ? What's the resolution of these spectra, how are they converted into your instrument resolution, assuming which spectral response function ?

l.188 : "inhomogeneous"

l.189 : does it mean that cases above are only pure ice clouds ?

l.191 : why setting spheres, which are for sure far from reality, and scatter way too much forward? Don't you think assuming another shape would be *a priori* more representative ?

l.193 : what's the rationale for this value of 0.2 radiance units ? How is this noise distributed across channels ? With any spectral correlation ?

l.194 : what temperature profile, the one assumed in the retrieval ? Is the actual one provided by Cox et al. (2016) ? More generally what information is provided by these latter authors along with the radiances. Why not recomputing the radiances on your own, because here we can't say what's the impact of having different codes for the radiance computation and the retrieval algorithm

l.195 : again, why such offset ? Is it based on experimental characteristics of the instrument/calibration unit ? Please give some insight into this value, maybe converting it into blackbody temperature offset.

l.198 : how do you observe such noise ?

l.200 : issue with section/subsection

l.204 : this information on cloud boundaries should have been provided earlier

l.205 : why not adding cloud top in the state vector, although the uncertainty would be large ? This would still be better than assuming a geometrically thin cloud

l.206 : again, droplets – spherical

l.207 : no information is provided about the uncertainty of the retrieved values, while this is critical here

l.210 : how are thin clouds selected ? What's the threshold ? Is there a flag when a retrieval does not work (for instance for optically thick clouds) ?

l.215 : it seems that there is absolutely no difference with CLARRA. Can you extend on this. Then, what's the point of developing a novel tool ? What differences would you expect ? What are the retrieved quantities for CLARRA ?

l.218 : what's the difference between CLARRA and TCWret for cloud position. In CLARRA it is in a layer (forced by an initial profile ?), but what about TCWret ? This is not clear.

l.220 : Do you mean that CLARRA was originally developed for an instrument covering these microwindows as well ? And you use those because they are available in the theoretical radiances ? Stress more strongly this difference in terms of far-infrared partial coverage.

l.223 : Why is there a difference between liquid water **path** and ice water **content** ? This is not clear. Do you compare only CWP or also optical depth and r_{total} with Cloudnet ?

l.226 : this formula is useless, this is trivial

l.230 : how do you handle these errors when integrating ?

l.231 : where do these flags come from ? Can you explicit them to make it more understandable. Does it mean that you reject specific flags ?

l.234 : likewise, can you explain these flags. Why "No liquid water" in this subsection ?

l.236 : I don't understand. Where do you get LWC ? How can its integral be different from LWP ?

l.245 : when fitting this line, you always force it to pass through the origin ?

l.246 : this section is difficult to interpret, because we don't know to which extent the retrieval should be perfect, or what the residual errors come from. Inconsistencies between radiative codes or cloud assumptions, cloud position etc. The motivation for performing such unperturbed retrieval should be emphasized.

Figure 5 : would error bars on the retrieved value help interpret the figure ?

l.247 : why changing the τ limit from 8 to 6 ?

l.251 : can be retrieved

l.252 : it's very hard to comment on these correlation coefficients. What would you expect ?

l.255 : remind the reader what could be the inconsistencies there. I would actually expect something even worse if you assume spherical ice crystals

l.253 : « a closer view » is not very explicit

l.256 : it is not clear what the subset shown in Figure 9 is. What about the shape in Cox et al. (2016) database ? What's the difference between Figs. 9 and 5b?

l.262 : correlation coefficient r is not altered by the addition of noise ? Can you explain this ? In Fig. 6a it seems that optical depth exceeds 6 while you said (l.249) that you were excluding values larger than 6. Can you clarify.

l.266 : larger than what ? There was no underestimation so far.

l.276 : warning : a temperature offset won't have the same impact at all wavenumbers, so it's not equivalent

l.276 – 277 : consider rephrasing because this is not clear.

l.280 : at some point it would be helpful to understand how perturbations drive over- and underestimation of retrievals, providing some physical insight into how spectral radiance changes with the state parameters. Maybe showing the Jacobians at some point (earlier in the paper) would be useful.

l.283 : uniformly

l.292 : please clarify "different uncertainties in the spectral radiances"

l.300 : this is the first time error bars are shown (Fig. 11). Can you remind the reader where they come from. These should be discussed in more details throughout the paper

l.302 : the error bars are so large that points can hardly be out of this range

l.304 : I suspect in most cases the cloud extends over some height. Could you show the distribution of cloud altitudes or thickness according to Cloudnet as a complement ?

l.308 : I suspect assuming any other shape you've been working with would reduce the uncertainty due to ice crystal shape. Spheres might be the worst choice

l.310-315 : not sure this is useful because it only repeats previous parts of the paper. This is more generally the case for all this section, especially for such a short article. Consider removing This section 6.

l.318 : ~~c~~correlation ; retr**ie**ved

l.320 : leads

l.322 : again, explain why this criterion on r_{liq} is working. What's the physical meaning ?

l.327 – 328 : already said

l.328 – 329 : this is a commonplace without further suggestions regarding the meaning of "careful"

l.329 : representation → assumption

l.330 : an error

l.339 – 340 : already said

l.350 : detail further why Cloudnet does not work well on thin clouds

l.352 : only one blackbody ?