Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2021-2-RC2, 2021

Comment on amt-2021-2

Anonymous Referee #1

Referee comment on "Assessing synergistic radar and radiometer capability in retrieving ice cloud microphysics based on hybrid Bayesian algorithms" by Yuli Liu and Gerald G. Mace, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-2-RC2, 2021

Review for "Assessing synergistic radar and radiometer capability in retrieving ice cloud microphysics based on hybrid Bayesian algorithms" by
Yuli Liu, Gerald G. Mace

The authors made a conceptual study about a synergistic microwave (MW) radar and MW/SubMM(submillimeter)-radiometer retrieval. The topic of the study is similar to the study of Pfreundschuh et al. (2020, AMT). Main differences are the different retrieval methods and that the authors' study is more restricted to a simpler atmospheric setup. The authors compare the synergistic retrieval to radar only and radiometer only retrieval. The synergistic retrieval is a combination of an optimal estimation (OEM) retrieval for the radar and Bayesian Monte Carlo integration for the radiometer. The basic story of the study is clear. The authors first define the hypothetical sensors, then explain the retrievals and retrieval databases, and show and discuss the results.

General comments:

Despite the basically clear general story, the study is partially confusingly written and needs some restructuring. Some examples:

- The retrievals are described in Sect. 3 but the actual values for the different thresholds, minimum numbers etc. are given in Sect. 5. This would make sense, if several retrieval configurations for each of the three retrievals were used, but this is not the case. The same holds for the actual retrieval quantities. They are given in the text of Sect. 5 but even there it is difficult to grasp what are the actual retrieval quantities, for example is log(IWC) or IWC a retrieval quantity.
- In Sect. 4.2, it seems that the retrieval database is described but then the authors suddenly write about ensemble generation.

Update Sect. 3 explicitly stating for each retrieval the retrieval quantities, the measurement quantities, the assumed non-retrieval quantities, and all the retrieval specific a priori assumption, thresholds, minimum numbers etc.. Additionally, please do not mix ensemble generation and retrieval database.

As I understand your paper, the idea is to combine a conical scanning radiometer with nadir pointing radar. This has two implications that you did not address in your study:

- It is very unlikely that both sensors will have the same footprint.
- Due to the different viewing geometry, both sensors have a different view on the atmosphere.

I think, it is not needed to expand your study to include these effects, because focusing on a 'best case' retrieval is still a formidable task. Nonetheless, you should explicitly mention that you neglect any footprint or viewing geometry effects and discuss at least briefly the implications.

I did not check for typos and correct grammar.

Section by section comments:

Section 1 Introduction

p. 3, l. 51: "The retrieval results are obtained through interpolation over the precalculated databases." Neural network retrievals are not an interpolation over a database.

Section 2 Simulated observations

2.1 Remote sensors
Please include information about assumed footprint sizes and write explicitly the assumed viewing geometry of both sensors.

p. 4, l. 77-78: Put the channel description including noise, channel number, main spectral feature (H2O-line, O2-line, window…), etc. into a table and refer to it.

p. 4, l. 80: Please insert a sketch with the viewing geometry of the sensors.

p. 4, l. 79: "Most frequency channels are centered on water vapor absorption lines." Change to something like this: 'The 183 GHz and 380 GHz channels are centered around H2O lines and the other channels are centered around the O2-line or a place within the window region.'

2.2 Reference cloud scenes

p. 5, l. 94-95: "The reason for these simplifications is still to be consistent with the a priori database that will be discussed in section 4." Did you consider the possible errors due to this simplification?

p. 5, l. 97-100: What is the horizontal spacing between each model profile and what is the actual horizontal grid size of the ECCC model?

2.3 Radiative transfer model

Does the radar simulator includes attenuation?

p. 5, l. 106-108: Please add what kind of particle size distribution (PSD) including used parameter and constants you used.

p. 5, l. 109-110: Please include briefly the other forward configurations here and not by just referring to the other paper.

2.4 Simulated observations

p. 5-6, l. 116-119. In the text you wrote cloud ice and in your figure 3, you distinguish between snow and ice but in Sect. 2.2 (p.4, l. 93), you write: "…we do not differentiate the cloud ice and snow…". Please be consistent or explicitly state in the corresponding lines you distinguish between ice and snow. Otherwise it can be confusing.

3 Hybrid Bayesian algorithms

p. 6, l. 135: "…and it[BMCI] is highly efficient since the retrievals are done by interpolating the database cases" BMCI is strictly speaking not an interpolation.

3.1 Radar-only retrievals

p. 7, l. 158: "…where K is the Jacobian matrix to linearize the forward model." Is K the Jacobian matrix of the retrieved state or of the a priori state? Please clarify within the text.

3.2.1 Synergistic radar and radiometer retrievals

p. 8, l. 182-183: What are "standard normalized vectors". Please explain it within the text. Furthermore, do you generate only profiles of IWC and NC or do you generate also temperature and humidity profiles?

3.2.2 Radiometer-only retrievals

p. 9, l. 197-202: I understand that you want to give a brief explanation of your retrieval. Nonetheless, your algorithm is complex. Therefore, I suggest to add a flowchart/algorithm chart especially in the view that you want to give a brief explanation of your retrieval algorithm.

p. 9-10, l. 209-211: "Following this step, the sampling module starts by reselecting the cases according to their posterior value to multiply cases with high weights and kill cases with low weights, and the weights of the selected cases become equivalent again." Please rephrase it. The sentence is hardly understandable.

p. 10, l. 211-214: "The sampling module then adds correlated random noise to the selected cases using the two-point correlation statistics in the covariance matrix. The covariance matrix is computed using the posterior PDF based on Bayesian MCI…" I do not fully understand what do you mean with "cases". Is this an atmospheric state or is this the whole database entry including brightness temperature and full atmospheric state? Is the

mentioned covariance matrix calculated from the full atmospheric state or just from the retrieval quantities? Please clarify within the text.

p. 10, l. 218-221: Please consider to use mathematical formulas and equations.

p. 10, l. 222-223: "…the algorithm evaluates these cases based on the prior PDF and likelihood PDF,…" Please explain, what "likelihood PDF" and "prior PDF" in that context mean? Is "prior PDF" the PDF of the previous iteration? Please clarify within the text.

Section 4 Prior information

I would suggest to rename this section to 'retrieval databases', because this is the actual topic of this section.

4.1 Radar retrieval database

Please add a sentence for what you need this database, because at first view it seems strange to have a retrieval database for an OEM retrieval.

p. 11, l. 254: How and where do you get the temperature information?

4.2 Radiometer retrieval database

p. 13, l. 305-306: According to Eriksson et al. (2020, AMT) and the citation therein "E" is a matrix with each column an eigenvector. Please correct.

p. 14, l. 313-316: Please add the noise quantities to a table with the other channel specific properties, see also my comment about 2.1 Remote sensors.

Section 5 Retrieval simulation experiment and results

p. 14, l. 331: "Similarly, the Gaussian noise of 1K is added to the simulated BT observations in each channel to characterize the measurement accuracy of the submillimeter-wave radiometer,…" Why do you add only 1 K of noise to the simulated observations? The noise values in given in Sect. 4.2 are much higher. This seems to make no sense.

p. 15, l. 336: "For the radiometer-only retrievals, except for the IWC and NC profiles, we retrieve the water vapor profiles as well." The sentence is confusing. Please rephrase.

p. 15, l. 352-353: "The EnPE optimization and the final MCI computations are done directly in the state space, not in the logarithmic space." Some lines above (l. 348) you wrote "The Bayesian MCI computation is also done in logarithmic space". Please explain why do make a difference?

p. 15, l. 356-357: Why is the radiometer retrieval (Figure 8) so noisy?

p. 16, l. 363-365: Do you have any idea, why the synergistic retrieval is noisier than the radar-only retrieval?

p. 16, l. 380-381: What do you mean with "non-Rayleigh effects and attenuation"? Please explain within the text.

p. 16-17, l. 371-396: Please add some discussion about the consequences that you use a combined PSD for snow and ice within your retrieval. For me, it seems, that some of the bias in IWC of the radar retrievals is due to fact that you do not separate between ice and snow.

p. 17, l. 400: What is meant with "retrieval uncertainty" and how do you estimate it?

p. 18, l. 414-415: Please add a sentence explaining what an error of 1 dB corresponds to.

p. 18-19, l. 414-438: Your error unit seems to be wrong. For example, according to Figure

10 (top left) for true IWC of 10^-5 g/m^3 the maximum error ratio is about 10^-4, which corresponds to -4 B or -40 dB. Your error values in dB are off by a factor of 10.

p. 18-19, l. 430-438: Please discuss the CDFs of Figure 13 or remove them.

p. 19-20, l. 439-463: Except for the IQR the plots of Figure 14 seem to show no added value. Remove them or show the added value.

p. 20, l. 459: Please define explicitly and explain the median fractional bias. An equation could be helpful.

Section 6 Summary and conclusions

Please add some (short) comments comparing your results to Pfreundschuh et al. (2020, AMT) and about your main retrieval assumptions (viewing geometry, footprint sizes, no liquid …).

Figures and Tables

Figure 1:
Please mark in the spectrum plot the relevant spectral lines and features. For example, add an ‘O2’ to the 118 GHz line and so on.

Figure 2:
Within the study, you do not distinguish between ice and snow. Therefore, replace “Frozen (Ice + Snow) Water Content” with ‘ice water content’. Furthermore, replace “water content (WC)” with ‘ice water content’ (IWC).

Figure 3:
Combine ice water path and snow water path.

Figure 8:
Replace “Ice + Snow Water Content” and “Ice + Snow Number Concentration” with ‘Ice water content’ and ‘Number concentration’, respectively.

Figure 9:
Unit of the y-axis is wrong.

Figure 10:
IWC and IWP units are wrong.

Figure 11:
IWC and IWP units are wrong.

Figure 12:
IWC and IWP units are wrong. Furthermore, the error unit is wrong. The given numerical values correspond to ‘B’ not ‘dB’.

Figure 13:
The unit for the PDF seems to be missing.

Figure 14:
The unit of the y-axis is missing.