Response to Reviewers

Reviewer 1

We would like to sincerely thank the reviewers for their professional comments and helpful suggestions. We believe they help us to improve the manuscript significantly and provide many useful ideas to our work. This manuscript has been revised by native English editors. We have revised the manuscript according to the reviewer's comments and answered the reviewer's question point by point below.

Reviewer comments are in italic blue and our responses in black, the manuscript changes are in red.

Reply to General comments

In terms of focus, I think the authors need to make their conclusions more specific. Currently, their main conclusion is that adding the 150 GHz channels to the retrieval improves its accuracy, which is not very informative. I therefore suggest to include the retrievals which use only channels around 183 GHz and retrieval using Ch. 2-5 in the results in section 4.3. This would allow the authors to quantify the effects of using one or two 150 GHz channels on the instantaneous accuracy (using the tropical cyclone case) as well as the climatological accuracy (using the comparison of the yearly means). This would put the results into a more practical context and provide the novelty required for publication in AMT. I would also like to encourage the authors to publish their code and include a reference to it in the manuscript.

Reply: Thanks for the great suggestions. The manuscript has been revised and rewritten in response to reviewers' comments. The comparisons of adding 150 GHz channel and only using 183 GHz channels in tropical cyclone and year maps have been added. In addition, the code and test data has been uploaded to Zenodo and mentioned in the manuscript.

Reply to specific comments

• l. 32: get microphycial of clouds -> determine the bulk and microphysical properties of clouds

Reply: Thanks for the comment. It has been corrected.

• l. 48: microphysical -> microphysical properties

Reply: Thanks for the comment. It has been corrected.

• l. 67: The name of the sensor is Cloud Profiling Radar. I added the capitalization just to emphasize the difference. I am sorry if this caused confusion.

Reply: Sorry, I misunderstood your comment before. It has been corrected now. Thanks again.

• l. 74: . . . based on the deep neural network -> . . . based on a deep neural network

Reply: Thanks for the comment. It has been corrected.

• l. 128: Ice clouds are . . .

Reply: Thanks for the comment. It has been corrected.

• l. 129: I would suggest using phenomenon or component instead of parameter

Reply: Thanks for the comment. It has been corrected.

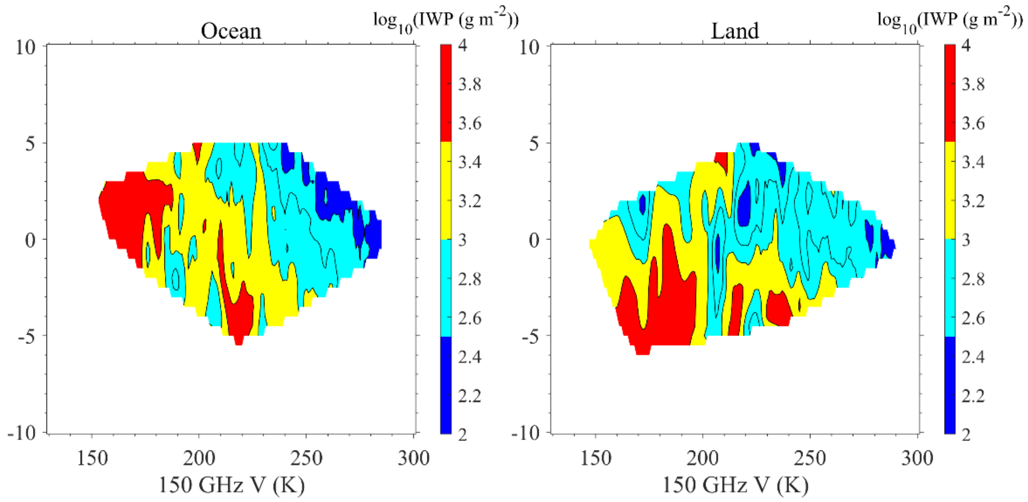
• l. 177: Why would the PD at similar incidence angles as that of conical scanners be lower? Isn't that rather an effect of the much larger footprint?

Reply: Thanks for the comment. I mean the PD of the quasi-polarization channels is lower than the real PD. The much larger footprint of MWHS is of course an important reason. The changes are as:

"However, due to the quasi-polarization mode and the much larger footprint, the PD of MWHS is much lower than that of conical scanners (e.g. GMI)."

• Fig. 4 (c) and (d): Please reduce the bin size of for these plots or smooth the results to make the contours less noisy.

Reply: Thanks for the comment. The figure has been changed as:



• l. 193: . . . a nonlinear mapping from the input to the output data

Reply: Thanks for the comment. It has been corrected.

• l. 250: If I understand you correctly, the conclusion that Ch. 4 is the best for cloud detection is based on it adding the largest improvement in Tab. 2. I don't think that this is a valid conclusion as it may just be that only Ch. 5 and Ch. 3 together work better than Ch. 4 - 5.

Reply: Thanks for the comment. The conclusion is corrected and the Table 2 has been revised. The changes are as:

"The cloud filtering performance for different channel combinations is listed in Table 2. The results showed that all three 183 GHz channels have cloud identification capability, and the addition of one 150 GHz channel enhances the POD of the network, while the two 150 GHz channels do not yield additional information."

Table 2. Errors of cloud filtering using different channels

	AC	FAR	POD	F1	CSI
1. CH.1-5	0.91	0.31	0.61	0.65	0.48
2. CH.2-5	0.91	0.31	0.61	0.65	0.48
3. CH.3-5	0.91	0.31	0.54	0.60	0.43

4. CH.3&4	0.90	0.30	0.52	0.59	0.42
5. CH.3&5	0.90	0.31	0.50	0.58	0.41
6. CH.4&5	0.91	0.29	0.54	0.61	0.44
7. CH.3	0.88	0.42	0.37	0.45	0.29
8. CH.4	0.90	0.26	0.41	0.52	0.35
9. CH.5	0.89	0.33	0.35	0.46	0.30

• *l. 259: . . . five channels are all used -> . . . all five channels are used*

Reply: Thanks for the comment. It has been corrected.

• *l. 276: . . . focused on here.*

Reply: Thanks for the comment. It has been corrected.

• *l. 309: The PDs at . . . : Some text seems to be missing here.*

Reply: I am sorry for the mistake. The description has been revised.

• *l. 309: The PDs do not share the distribution characteristics of the low TBs. That would be in contradiction of PDs occurring mostly in warmer ice clouds and outflow regions.*

Reply: Thanks for the comment. The description has been revised. The changes are as:

"The distribution characteristics of PDs at 150 GHz (TB_V-TB_H) are similar to the structure of the tropical cyclone, but significant PDs occur mainly in the warm ice clouds at approximately 200-250 K."

• *l. 379: Refine conclusions to assess effects of a single 150 GHz channels as well as both 150 Ghz in terms of instantaneous estimates as well as annual means.*

Reply: Thanks for the suggestion. The changes are as:

"Since 2C-ICE was used to train the networks, MWHS IWP is certainly approaching the 2C-ICE and similar to the IWP maps in Duncan and Eriksson (2018). There is no significant difference between the results of the three MWHS channel combinations on the map, but the IWP result using only the 183 GHz channels is lower at middle latitudes than the IWP results with the addition of the 150 GHz channels."

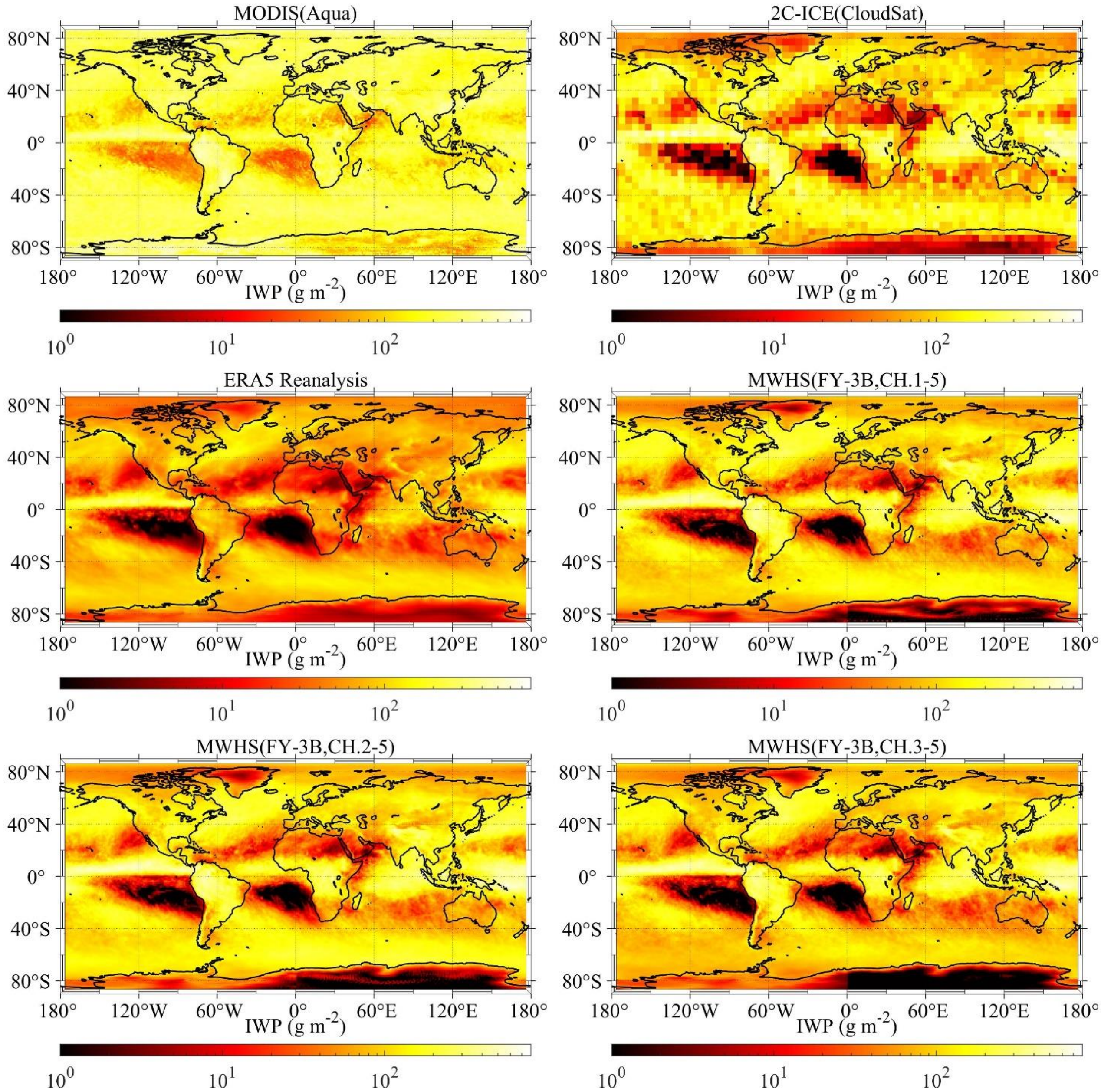


Figure 11. Global mean IWP maps for 2015 from MODIS, 2C-ICE, ERA5 and different channel combinations from MWHS. 2C-ICE is gridded on a 5° grid, while the other products are gridded on a 1° grid.

• Fig. 10: Please change the comparison to 2C-ICE to a line plot include this as an additional figure. I suggest you also add results from a retrieval using only channels around 183 as well as a retrieval using Ch. 2 - 5.

Reply: Thanks for the suggestion. The changes are as:

“For tropical cyclone retrieval, the addition of the 150 GHz channel does not have a significant impact on the accuracy. The RMSE and CC of the three retrievals are similar. Although there are differences between MAPE and BIAS, the differences are not significant.”

Table 6. Errors of the tropical cyclone retrieval

	RMSE (g m^{-2})	MAPE (%)	BIAS (g m^{-2})	CC
CH. 1-5	1191.3	77.69	82.07	0.73
CH. 2-5	1197.3	82.98	18.22	0.72

CH. 3-5	1174.1	79.71	-113.67	0.73

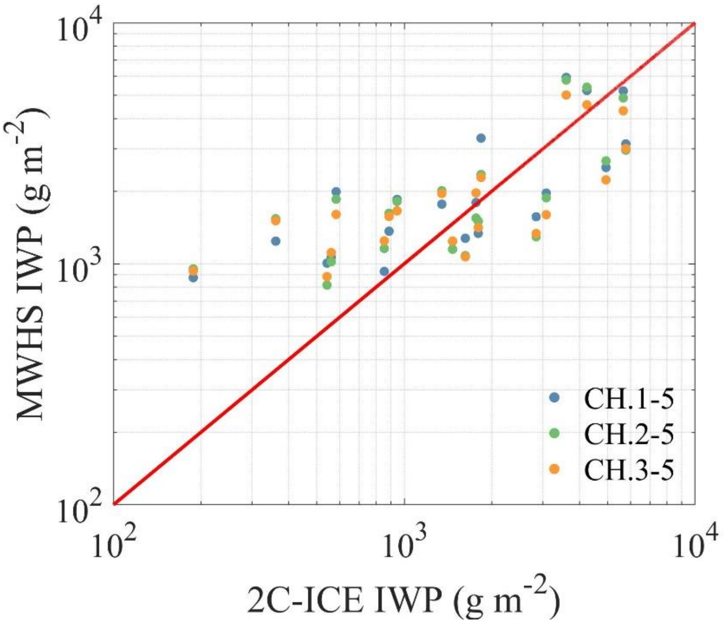


Figure 10. IWP comparison of MWHS and 2C-ICE at the tropical cyclone Bansi.

• *Fig. 12: Here I suggest you also include results from a retrieval using only the channels around 183 as well as a retrieval using Ch. 2 - 5.*

Reply: Thanks for the suggestion. The changes are as follows:

"The IWP from MWHS is generally close to 2C-ICE, and the result without the 150 GHz channel is significantly lower than 2C-ICE between 30°S - 60°S in the Northern Hemisphere and 20°N - 60°N in the Southern Hemisphere. There is an improvement after adding the 150 GHz channel (little difference between using 1 or 2 150 GHz channels), and the IWP in the Northern Hemisphere is basically the same as the 2C-ICE, while it is still lower in the Southern Hemisphere."

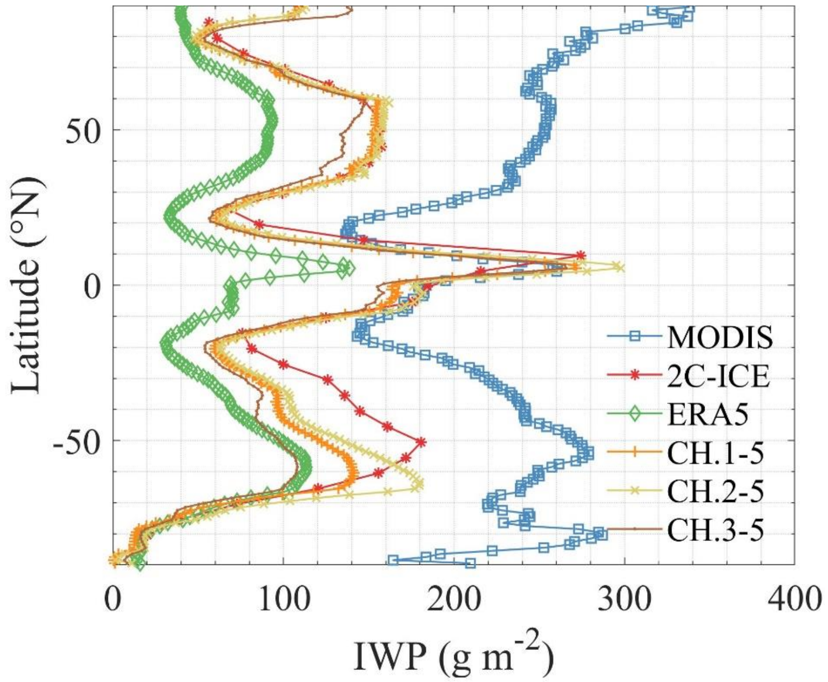


Figure 12. Zonal means of IWP for 2015 from MODIS, 2C-ICE, ERA5 and different channel combinations from MWHS. 2C-ICE is gridded on a 5° grid, while the other products are gridded on a 1° grid.

• l. 361: If you mention quantile regression neural networks, please cite Pfreundschuh et al. 2018.

Reply: I am sorry for the mistake. The reference has been added.

• l. 376: See comment above on impact of Ch. 4

Reply: Thanks for the comment. It has been corrected.

• l. 379: Refine conclusions to assess effects of a single 150 GHz channels as well as both 150 Ghz in terms of instantaneous estimates as well as annual means.

Reply: Thanks for the comment. It has been corrected.

Reviewer 2

We would like to sincerely thank the reviewers for their professional comments and helpful suggestions. We believe they help us to improve the manuscript significantly and provide many useful ideas to our work. This manuscript has been revised by native English editors. We have revised the manuscript according to the reviewer's comments and answered the reviewer's question point by point below.

Reviewer comments are in italic blue and our responses in black, the manuscript changes are in red.

Reply to comments

1. More statistical plots of the retrieval database are suggested to provide. An interesting figure should be the two-dimensional histogram between IWP and TB for different channels. Since the IWP-TB relationship has been revealed by both forward model simulations and remote sensing measurements, these plots could help to verify if the collocation database captures the right statistics. Also, the correlations between TB/PD and the scanning angle in the database are worth to be investigated. The results in Table 3 show that adding scanning angle as auxiliary information cannot improve the retrieval accuracies. Since TB and PD are both significantly influenced by the incidence angle, this geometry viewing information should be able to better constrain IWP in the state space. Additional plots in this aspect could make the discussion clearer.

Reply: Thank you very much for the comments. We agree with the criticism and two statistical figures have been added. The changes are as:

"The statistical information of TB and IWP for different channels (CH.2 – CH.5) in the collocation database is given in Fig. 4. The TBs for CH.3 and CH.4 were mainly concentrated at approximately 250 K, indicating small sensitivity to ice clouds. For CH.2 and CH.5, the TBs had a larger range of variation, which is due to the larger contribution of near-surface information to the "window" channels. However, it can be seen that in the presence of ice clouds (IWP >100 g m-2), the surface information is blocked by clouds, making the TB range significantly smaller as the IWP becomes larger. The statistical relationship between the 150 GHz TB and IWP at different scan angles is given in Fig. 5. It can be found that there is a significant decrease in the measured TB with increasing IWP for large scan angles. As the scan angle decreases, especially in the case of nadir observations, there are many low TBs appearing in clear-sky scenes because nadir observations have a very large number of collocation scenes in the polar regions (see Fig. 2b), where the surfaces lower the measured TB. In contrast, collocation scenes with large scan angles are mainly located in the tropics, which makes the TB-IWP relationship very significant."

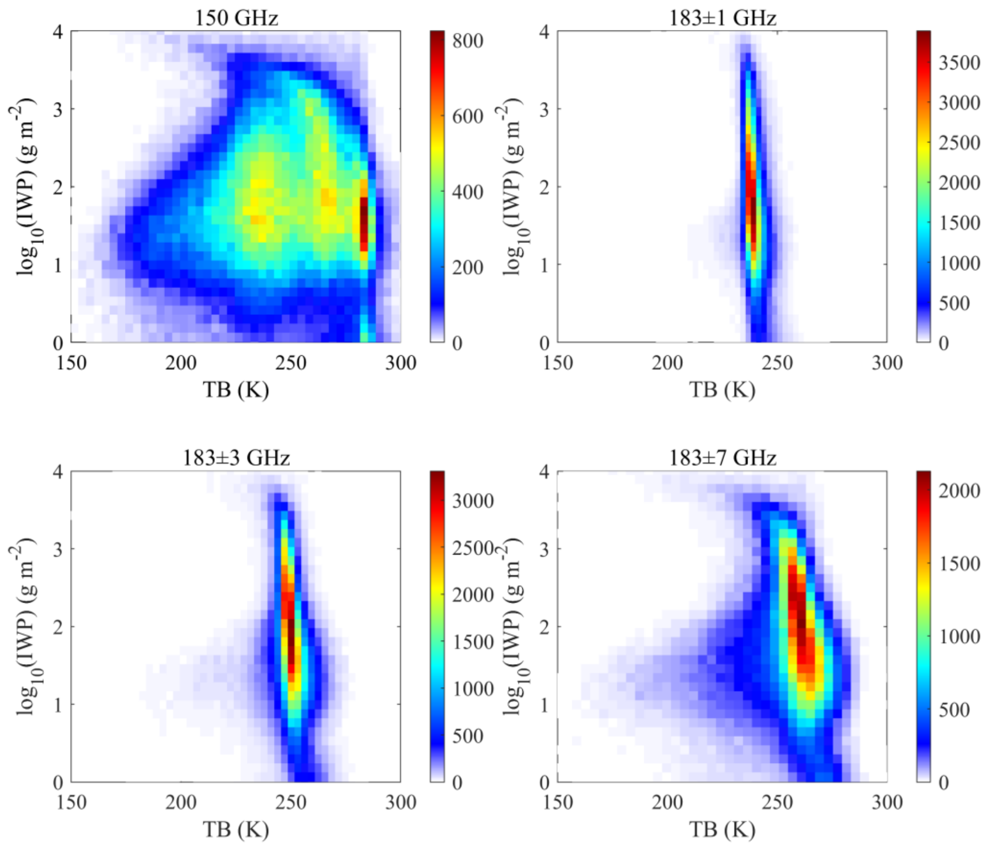


Figure 4. Statistical information of TB and IWP for different channels.

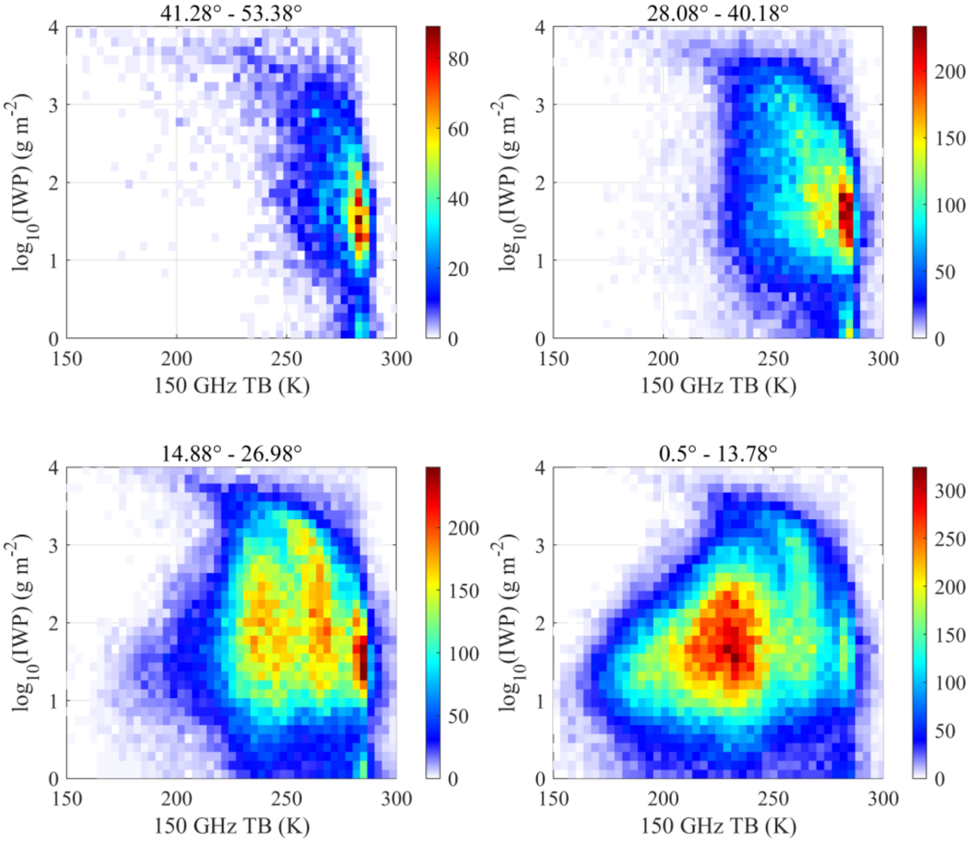


Figure 5. Statistical information of IWP and 150 GHz TB for different scan angles.

The incident angle and observation geometry information should be able to better constrain the IWP in state space, which is why Holl et al. (2014) used local azimuth and zenith angles in training. Here we just used the scan angle of the satellite, which does not fully represent the geometric information, so it is of little help for the inversion, but combining it with the latitude information can be of better use.

*2. I suggest investigating the sparsity of the retrieval database in the measurement space when conducting retrievals, at least for the case study in section 4.3.1. You can plot the TB of each channel against that of all the other channels for the database and measurements, respectively, as the fig.2 in Brath et al., 2018 shows. Alternatively, you can calculate the \chi^2=(y_{obs}-y_{db})^T*S_y^{-1}*(y_{obs}-y_{db}) and check the \chi^2 values for the several best cases. If the training database does not well cover the full range of measurements, the NN retrievals cannot be expected to produce accurate results. This investigation could help to find the limitation of the retrieval database and better explain the retrieval results.*

Reply: Thanks for the comments. The changes are as:

"Figure 9 gives the TB of each channel against that of all the other channels for the training dataset (blue) and the validation dataset (red). Overall, the training dataset covers the full range of the validation dataset, which means that the neural network is well representative."

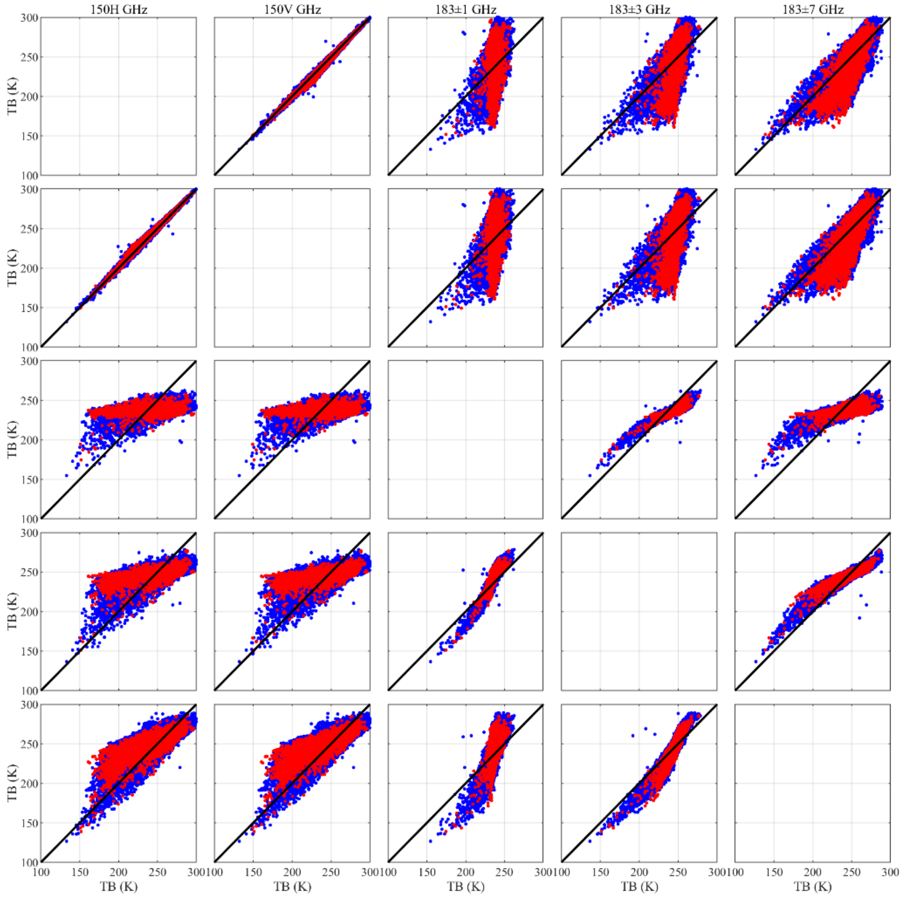


Figure 9. Measurements comparison from different channels of train data set (blue) and valid data set (red).

3. Since very similar work has already been done by Holl et al., 2014, you should explicitly discuss whether your results are consistent with the results in Holl et al., 2014. Either direct or relative comparison is fine, and more statements regarding the improvements and extra findings of this work should be added. By the way, could you comment on why the biases in Tables 3-5 are always negative? Shouldn't the retrievals be unbiased?

Reply: Thanks for the comments. The changes are as follows:

"In terms of the retrieval using the neural network, the results of this paper are basically consistent with Holl et al., (2014). The error between the retrieval results and 2C-ICE is approximately 100%. The latitude and ocean/land mask are important auxiliary information for DNN retrieval. Holl et al., (2014) used angle information that contains geometric observations of the local zenith and azimuth and showed a significant improvement. However, the results in Table 3 show that the scan angle is of limited help for retrieval, due to the fact that the scan angle is not fully representative of the geometry of the observed radiance, and it works better when used in conjunction with the latitude and land/sea mask."

The main improvements and findings of this work is the use of polarized 150 GHz channels. We have further supplemented the relevant channel combination comparison (with only using 183 GHz channels) in accordance with the comments of Reviewer 1 (see Figs 13, 14, 15). Related statements have been mentioned in the discussion and conclusion.

In the NN training, we chose the logarithm of the IWP as the training target and used the absolute percentage error as one of the main references for selecting the model, which make the model results underestimating the true value, as mentioned previously by Reviewer 1. In addition, due to the difference in sensitivity of MWHS and CloudSat to ice cloud particles, it will lead to biased results.