1 Comments from reviewer 2

We would like to thank the reviewer for reading our manuscript and providing helpful feedback.

1.1 Specific comments

Reviewer comment 1

Lines 69-71. It may be helpful to add the time periods of these campaigns in the text or in the table.

Author response:

We will include the requested information in the revised version of the manuscript.

Changes in manuscript

- The sentence introducing the B984 flight will be modified to:

> **Changes starting in line 66:**
>
> The first considered flight, designated B984, ~~took place~~ was performed on 14 October 2016 ~~as part of~~ during the North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX~~, Schäfler et al. (2018))~~), which took place during September and October 2016 (Schäfler et al., 2018).

- The sentence introducing the C159 and C161 flights will be modified to:

> **Changes starting in line 71:**
>
> The two other flights, designated C159 and C161, ~~took place in March 2019 as~~ were part of the PIKNMIX-F campaign~~.~~, which took place in March 2019.

Reviewer comment 2

Line 100. It seems that you do not have the same measurements to perform the retrievals for different flights. How is the missing information handled in the retrieval method?

Author response:

To handle the varying availability of channels (and sensors) across the different flights, the retrieval implementation is adaptive in the sense that it can be run with arbitrary sensor configurations. Channels that are sensitive to surface emission and thus only used over Ocean are disabled by setting their assumed uncertainties to very high values.

Changes in manuscript:

- To make this clearer, we will reformulate the sentence that describes the adaption of the forward model and retrieval.

> **Changes starting in line 134:**
>
> The ~~retrieval forward model has been~~ forward model and retrieval were made adaptive so that the ingested observations can be easily adapted to the ~~sensors~~ different sensors and channels that were available for each flight. Low frequency channels that are used only over Ocean surfaces are deactivated over land by setting the corresponding channel uncertainty to 10^6 K.

Reviewer comment 3

Figure 10. The discrepancies are quite large for 243 GHz channel for flight C161 compared to the other two flights. Could you comment on that?

Author response:

As we explain in the manuscript, we suspect that the remaining discrepancies in the 243 GHz channel for flights C159 and C161 are mainly caused by precipitation that may be observed differently by the different sensors due to co-location issues. This reasoning is based on the observation that similar residuals are observed in the same location in other passive channels that are sensitive to the lower atmosphere as well as the radar. For the region where the largest biases are observed for flight C161 the temporal delay between radiometer and radar observations is about 30 minutes during which the structure of the cloud has likely changed thus leading to inconsistencies between the radar and radiometer observations.

The initial version of the manuscript has not mentioned the temporal co-location of the observations that differs significantly between the flights. For the revised manuscript we propose to add a new figure that displays the time delay between the radar and radiometer observations for the different flights. We will also extend the discussion of the residuals.

Changes in manuscript:

- We will add the figure shown in Fig. 1.1 together with the description shown below to the the section the presents the radar observations.

> **Changes starting in line 90:**
>
> While the radar observations for flight B984 come from an airborne radar, the observations for flights C159 and C161 stem from a spaceborne sensor. The high velocity of the spaceborne sensor causes significant temporal delay between co-located observations from the radiometers and the radar. Figure 1.1 displays the delay between co-located radar and radiometer observations with respect to the along-track distance for the three flight scenes. So while the delays for flight B984 remain mostly within 5 minutes, they reach values exceeding 30 minutes for the two other flights.

- The presentation of the retrieval residuals will be extend as follows.

> **Changes starting in line 188:**
>
> Radiometer residuals for flight B984 are mostly within ±5 K ~~. For the two other flightsthe residuals are larger. Differences up to and~~ but larger for flights C159 and C161. For these two flights, residuals exceeding 10 K are observed ~~at~~ in the window channels up to 243 GHz as well as ~~in~~ the outermost channels around the absorption lines at 118 GHz and 183 GHz. Since these ~~correspond to profiles in which residuals of opposite sign are present~~ occur in profiles where precipitation is present and in which similar residuals can be observed in the radar observations, a likely explanation is that they are caused by ~~small-scale precipitation events that are missed by one of the~~ precipitation that is not observed by all sensors due to spatial and temporal co-location issues. Especially the large residuals in the 243 GHz channel for flight C161 at around 100 km along track distance may well be caused by the evolution of the convective cloud during the delay of almost 30 minutes that separates the radiometer and radar observations.

Reviewer comment 4

Lines 220-222. As you also mentioned, the largest uncertainties correspond to the higher-level clouds where we have smaller ice particles. Are these uncertainties also related to the lack of representativeness of the particle shape/type used in the models?

Author response:

As we meant to express in the paragraph starting in l. 285, our hypothesis is that the underestimation of IWC high up in the cloud that is observed across all tested particle models is rather due to the mismatch in shape between the assumed and observed PSD. Since none of the sensors used in this study has any significant sensitivity to particles with diameters smaller than 200 μm, the contribution of those particles to the total

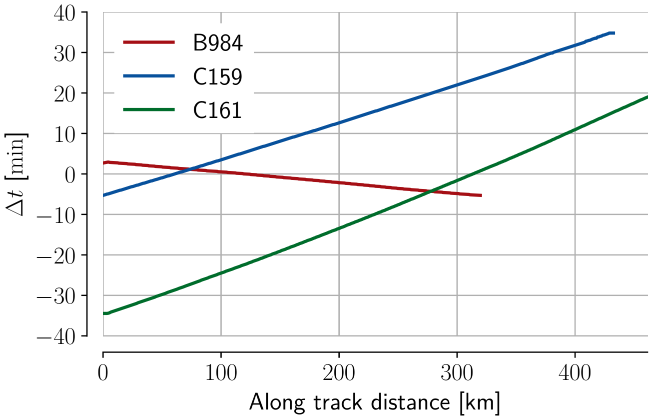

Figure 1.1: Delays between the co-located observations from radar and radiometers for the three flights.

IWC is essentially inferred through the assumed shape of the PSD. Since the assumed PSD drastically underestimates the amount of those particles, this may explain why the retrieved IWC is lower than the true IWC.

Changes in manuscript:

To make this point clearer, we will reformulate the paragraph that discusses the deviations from the in situ measurements.

> **Changes starting in line 288:**
>
> This indicates that the ~~observed deviations are~~ concentrations of these larger particles may be retrieved correctly but that the total IWC is underestimated due to the ~~presence of a large number of small ice particles that the microwave observations are not sensitive to~~mismatch between assumed and actual PSD shape, the former of which lacks the very high concentration of small particles that are present in the in situ measurements. Although O'Shea et al. (2021) and O'Shea et al. (2019) show that the ~~occurence~~ occurrence of high particle concentrations at sizes below 200 μm may be due to measurement inaccuracies of the CIP-15 probe, the measured PSDs correctly reproduce the measured IWC at these altitudes when the corresponding water content is calculated using any of the tested particle habits (Fig. 8).

Bibliography

O'Shea, S., Crosier, J., Dorsey, J., Gallagher, L., Schledewitz, W., Bower, K., Schlenczek, O., Borrmann, S., Cotton, R., Westbrook, C., et al.: Characterising optical array particle imaging probes: implications for small-ice-crystal observations, Atmospheric Measurement Techniques, 14, 1917–1939, 2021.

O'Shea, S. J., Crosier, J., Dorsey, J., Schledewitz, W., Crawford, I., Borrmann, S., Cotton, R., and Bansemer, A.: Revisiting particle sizing using greyscale optical array probes: evaluation using laboratory experiments and synthetic data, Atmospheric Measurement Techniques, 12, 3067–3079, 2019.

Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J. D., McTaggart-Cowan, R., Methven, J., Rivière, G., Ament, F., Boettcher, M., Bramberger, M., Cazenave, Q., Cotton, R., Crewell, S., Delanoë, J., Dörnbrack, A., Ehrlich, A., Ewald, F., Fix, A., Grams, C. M., Gray, S. L., Grob, H., Groß, S., Hagen, M., Harvey, B., Hirsch, L., Jacob, M., Kölling, T., Konow, H., Lemmerz, C., Lux, O., Magnusson, L., Mayer, B., Mech, M., Moore, R., Pelon, J., Quinting, J., Rahm, S., Rapp, M., Rautenhaus, M., Reitebuch, O., Reynolds, C. A., Sodemann, H., Spengler, T., Vaughan, G., Wendisch, M., Wirth, M., Witschas, B., Wolf, K., and Zinner, T.: The North Atlantic Waveguide and Downstream Impact Experiment, Bull. Amer. Met. Soc., 99, 1607–1637, https://doi.org/10.1175/BAMS-D-17-0003.1, 2018.