Reply to Reviewer 2

We thank the reviewer for his/her constructive comments which helped to improve the quality of the paper. In the following reviewer comments are marked in italics.

The paper addresses the very interesting and hot topic of water vapour concentration and distribution in the arctic regions with different satellite instruments and retrieval algorithms. Both L2 and L3 products are considered, which are compared with reference data derived from various re-analyses, radiosonde and GNSS observations. I have found the paper well written and complete; therefore, my remarks are minor. Here a few comments to improve the reading and understanding of results.

1. *For IASI and MIRS, the L2-product accuracy is not given while it has to be provided. It could explain in part the diverse performance of the two products. Based on the paper results, it seems that the bias of MIRS has not been well assessed previously.*

The quality of the IASI L2 sounding products (and also for MIRS) varies with the scenes. Better performances are expected in clear sky and with larger thermal constrasts. Assessments against correlative radiosondings and intercomparison to numerical models stratified in quality ranges as per IASI L2 quality indicators confirmed that biases (systematic errors) are close to zero. The precision (random component of the uncertainties) is better than 1K for temperature in the free troposphere in the best retrieval classes. The sounding precision decreases nearer to the surface. The exact characterisation of IASI L2 precision is more difficult as the collocation uncertainties with radiosondes (who most of the time are distant by at least 1 to 2h in time) are expected to play a larger role in the IASI-sonde departure budget. Precision of about 1.5K are typically expected. Same considerations apply to humidity, but the uncertainties in absolute and relative terms vary with the actual moisture load. Thus, it is not possible to give an easy answer. Note, that we cite the results of the intensive intercomparison of IASI and AIRS to GNSS measurements by Roman et al., 2016.

2. *The AR event should be better defined for the benefit of the reader*

We also modified some text for clarity and added the following definition: ” ARs have been also associated with several cyclones helping the moisture supply within the same AR structure (Sodemann and Stohl, 2013). In the Polar regions ARs are often associated with moisture inversions showing maxima in specific humidity between 800 and 900 hPa (Gorodetskaya et al., 2020). Here we choose the AR event from 6 June 2017 at 12 UTC to illustrate the capabilities of the different products (Fig. 4). Note, that this is only one out of three AR events that occurred during ACLOUD/PASCAL documented in detail by Viceto et al, (submitted).”

3. *The IASI retrievals come from a combination of IR and MW instruments. The product simply as IASI is confusing; maybe the authors could use the acronym IASI/AMSU/MHS, which could even be shortened to IAM.*

This concern was also addressed by reviewer 2. We followed his/her naming suggestion.

4. *Given the high cloud coverage of the arctic, IASI (the IR) is expected to yield a very poor contribution in comparison to AMSU and MHS; therefore, the large performance difference between IASI and MIRS (based again on AMSU and MHS) is not understood and would deserve a bit of more explanation.*

   As shown by Knudsen et al. (2018) average cloud top height during ACLOUD is about 700 hPa showing the dominance of low level clouds. Specific to the Arctic are also humidity inversions leading to high moisture amounts in the mid-troposphere which can feed clouds from above. In fact, during the AR on 6 June specific humidity is maximum at around 800 hPa. Therefore, IASI

still provides important information on mid-level moisture which due to its more pronounced weighting functions at these altitudes provides additional information compared to microwave only retrievals. Thus, the synergistic exploitation of microwave observations from AMSU and MHS together with the hyperspectral sounding from IASI enable quasi nominal sounding down to the cloud top and to preserve relatively good performances in most cloud-effected pixels. The assessments carried out with reference radiosonding confirmed that the quality indicator (= uncertainty estimates in the lower tropo) is a reliable. User can rely on this indicator to perform quality control tailored to their needs. A short discussion was added in the conclusion.

5. *On the same subject, what were the sky conditions during the ice camp in June? Clear Sky? Clear sky, in fact, would in part explain the better performance of IASI in comparison to MIRS.*

   The monthly mean cloud fraction at Ny-Alesund for May 2017 is 75% and in June 83% (Nomokonova et al., 2019, figure 9). For the ice camp of Polarstern it is 88 %. The information has been added to the paper.

6. *By the way, with reference to Fig.6 (6 June 2017), the authors say that the same bias pattern seen in IASI and MIRS is because they both make use of MW units. But this is conflicting with the MIRS large bias seen in Fig. 2. Once again, the MIRS bias could be better explained in light of the IASI good performance.*

   In Fig. 6 the difference seen between the different products ("biases") occurs mostly due to synoptic systems which are shifted in the different products due to their spatio-temporal sampling. This "weather noise" is much higher than the systematic long-term effects – compare Figure 6 and 9. Because the METOP satellites fly IASI and two of the MIRS instruments their sampling is rather similar leading to the same spatial difference patterns. The MIRS bias seen in Fig.2 especially in June is on the order of a few kgm-2 while the daily means in Fig.6 have differences between +50 and – 20 % (for IASI) which is up to 10 kgm-2. In fact, if you look at the area of highest relative differences close to Svalbard in Fig.6 you can see that IASI reaches up to 50 % while for MIRS it is up to 20-30 %. Therefore we do not see any contradiction.

7. *Tab A1 Skill scores. Why is the GOMe r for Ammasalik 0.07? This is really an outlier with respect to the r for other stations. Is that correct*

   The low correlation for this station is mainly due to a few measurement points where the GOME-2 WV is considerably lower than the RS WV. Possible reasons for this are the influence of clouds and changing surface albedo due to e.g. changing ice cover.