Comment on amt-2022-22
Anonymous Referee #3


Review of "Can state-of-the-art infrared satellite sounders and reanalyses detect moisture inversions in the Arctic?", Chellini et Ebell, AMTD 2022.

The manuscript presents an important systematic work in view of radiative budget and climate studies. The work analyses and documents the strengths and limitations of numerical models and of satellite products complementarily to sondes which, thought extremely accurate, have very sparse coverage and cannot support global climate studies.

I find the manuscript very well written: the structure is clear and the explanation are sufficiently detailed, with concise but yet complete illustrations.

I recommend the publication of the manuscript with minor revisions, considering the general and specific comments below.

General comments:
The hyperspectral IR Nadir sounders don't have the resolving power necessary to accurately represent the inversions like sondes do, especially surface-based and on short vertical scales. However, sounding products often contain at least inflexions where not-too-small temperature inversions occur. If observed with humidity too, even though not accurately describing the inversions, could this be interesting information for climate studies? I feel that this discussion is missing in the paper and could be more interesting than making the (expected) experience that those sounders lack vertical resolution, especially near surface.
Also, it would be useful to be specific that the assessment is really with IASI and AIRS products of a given version - as opposed to IASI and AIRS as a mission. It seems a bit pedantic, but the measurements may or may not be exploited at their best, especially in the context of polar sounding with humidity inversions. Even the title should perhaps be explicit about an assessment of IASI and AIRS L2 (versions TBC). Recent work by Prange et al, seem to indicate that IASI sounding could detect elevated moisture inversions in some instances.

Confirm whether the IASI L2 products have been used correctly: all-sky sounding and quality information seem to have been overlooked. Same with ignoring scene-dependent or at least typical averaging kernels to convolve with radiosonde profiles. This could be important to this study.

Evaluate and discuss if the numerical model is as skilled away from sonde stations as it is in Ny Alesund in representing humidity inversions. At least from a statistical standpoint, do we get numbers close to those of Ny-Alesund?

Specific comments:

L9-L11: I understand the point, but I feel it could advantageously be presented from a different perspective. The point is not so much the vertical grid sampling the retrieval (and analyses) may have been performed on, but more the vertical resolution intrinsic to the passive IR sounders - assuming the measurements are exploited to the maximum extent. The actual vertical resolution being usually significantly smaller than the vertical grid sampling (which is true for numerical models too). So the question is more about what are the vertical scales that matter for climate and radiative budget studies -which obviously super-resolved radiosonde profiles can represent- wrt how much of it can be captured by the passive satellite IR sounders. Perhaps the title could be revised a bit along these lines.

You are finding that, as expected, IR sounders can catch some (stronger) inversions but usually miss most (medium and small) and underestimate their intensity. Similar findings were made regarding surface-base temperature inversion in Antarctica. See paper, which I suggest to include in the reference.

L14-15: consistent with Boylan et al for temperature.
L17-19: It is essential to state whether the radiosondes used in your assessment have been assimilated in the ERA-5 reanalyses. Also to note somewhere if you happened to have used radiosonding away from Ny-Alesund. Indeed, because of the regular synoptic sounding from this station, it can be expected that the numerical model is particularly well constrained in that area. I suggest testing the model abilities away from the regular radiosonding stations, e.g. by computing similar statistics on inversions and see if number are close to those found in Ny Alesund.

Fig.1 and throughout: Use "Metop" official case after EUMETSAT standards (not MetOp)

L93: yes, IASI measures through a broader spectral range than AIRS, but more importantly for water-vapour sounding, with higher spectral resolution (a little over twice higher in the WV band).

L97: same as before, important to mention somewhere (the earlier the better I think) whether these sondes have been assimilated by the numerical models.

L113-114: well in fact, very often the grids are oversampled, so really we should be smoothing the radiosondes with the averaging kernels of the retrievals, which are normally provided along with the products, or with proxy AK (e.g. static seasonal AK typical for this region). This would be more of a validation exercise of the given IASI and AIRS products. But it remains more interesting to explore what these instruments can contribute (or not) to the monitoring of humidity inversions given their intrinsic vertical resolution/sensitivity. E.g. what Fig.4 shows, would that be sufficient?

L155: strictly speaking, NWP assimilates radiances, not profiles. I suggest rephrasing: "is to provide such _thermodynamic information_ with high resolution..."

L156: "the nominal accuracy" ==> the 1K/1km was the end-user requirements spelled at the beginning of the mission. IASI L2 Temperature do actually perform better than these objectives in the mid troposphere.

L159: the actual resolution of the IASI L1C products (radiances) after homogeneisation (i.e. removing instrument signature and wavelength dependency) is actually 0.5cm-1 (sampling is 0.25cm-1).

L159: MetOp ==> Metop + suggest reference Klaes et al. on the EPS programme. "The EUMETSAT Polar System - 13+ Successful Years of Global Observations for Operational
L185 and L186 is a confusion: OEM attempted in clear-sky as per cloud mask. Double-check.

L217: Actually, AIRS V6 includes a combined MW+IR retrieval scheme, which was the nominal strategy. However it became not functional after the loss of AMSU-A2 in 2016 (https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/AIRS/3.3_ScienceDataProductDocumentation/3.3.5_ProductQuality/V6_Test_Report_Supplement_Performance_of_AIRS+AMSU_vs_AIRS-Only_Retrievals.pdf)

It also includes a statistical first retrieval, implementing ANN - first guess to the physical retrieval, similarly to IASI. (https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/AIRS/3.3_ScienceDataProductDocumentation/3.3.4_ProductGenerationAlgorithms/V6_L2_Product_User_Guide.pdf)

L238-243: "IASI retrieval under specific sky conditions" and "No quality control is provided" ==> this is wrong. IASI L2 from EUMETSAT include all-sky sounding, nominally exploiting the microwave in synergy with IASI in the statistical retrievals (this is actually the regional service). The OEM usually only improves little over the first retrieval. The operational products also include quality indicators and error estimates, which have been used in other studies. The cloud mask in v6.5 (2020-) showed improvements over v6.4 (used here), which could play a role in the results here. Why have all-sky and QC info been discarded? I suggest liaising with the products developers if that has just been omitted in your study.

L265: yes, indeed.

L372: "out of scope": why has the utilisation of AK been excluded from this study? Out of simplicity? I think it is limiting somewhat the conclusiveness and should be explained.

L372: "We argue...": the syntax reads weird, not sure to get the point here.

L375: I see the rationale, but the 2.5-7.2 hinge points might be placed differently in the vertical. Also, I would expect DoF to be on the lower side in the Arctic anyway, I suggest commenting on quick typical DoF figures in the Arctic.
L405-408: you would get more match-ups with IASI all-sky retrievals.

L411: a word of cautious with the footprint size: IASI 12km is at Nadir. It goes up to 40km on swath edge. Similarly AIRS 50km at Nadir would flirt with 200km on swath edge. Would consequence could it have on humidity inversion statistics and in particular compared to point-measurements with radiosondes?

L412-413: Here again, IASI L2 contain all-sky retrievals.

L414-420: I find the difference between IASI-A and IASI-B puzzling/interesting. Big enough to be noted and it seems statistically significant. This was not expected, why is that?

L446: interesting result if genuine retrieved information, which would suggest that this should be at reach for IASI - but not yet in this product version.

L450: yes, indeed. For example, for IASI, a parallel to the 1K/1km figure for temperature was 10%RH/2km for humidity (in good conditions, how many DoF over Polar regimes?)... I think this rough figure should be accounted for in the preliminary discussions and in introducing the objective of this work.

4.4: I think we're hitting the same concept as with AIRS/IASI: the vertical sampling of ERA-5 is higher than the actual vertical resolution of the physical processes modelled and information assimilated in general. Are there Kernel functions for the models? that notion could be recalled in this paragraph too.

L497: yes, the radiosonde assimilated on this very point could well explain the good fit. The ERA-5 mid-range forecasts before an inversion occurs could be looked into to test this assumption. Also checking the frequency of low-level inversion in other places empty of radiosonde to constrain the model near surface, compared to the stats in Ny-Alesund is felt missing and would be very informative in this discussion.

L545: the difference IASI-A and IASI-B is unexplained while the IR+MW measurements are very comparable in nature (Metop-A having slightly less MW channels than -B at that time). It is certainly worth highlighting, feed-back to the products developpers and good information to other products users overall.

L576: yes, why hasn't it be done in this study? "mostly based on MW" ==> that is not
correct. The first guess exploits MW and "constraints" the OEM. But MW only really add value over IASI for the part of the atmosphere affected by clouds.