Comment on egusphere-2022-755

Nadia Smith (Referee)

Referee comment on "How adequately are elevated moist layers represented in reanalysis and satellite observations?" by Marc Prange et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-755-RC1, 2022

Review of paper entitled "How adequately are elevated moist layers represented in reanalysis and satellite observations" by M. Prange et al.

The authors present a study where they evaluate how well three datasets capture elevated moist layers (EMLs) in the extra-tropics, specifically at Manus Island, in the Western Pacific. With these three datasets they aim to characterize the difference in capability between a reanalysis model (ERA5) and satellite sounding products (IASI L2 and CLIMCAPS-AIRS). I think their experimental framework is sound and their scientific goal is relevant to ongoing research and product/instrument design efforts. I want to commend them on a paper that is well written and organized. It was easy to follow and made for an engaging read. I especially appreciate their effort to communicate their decisions clearly, which makes this paper a valuable scientific document.

Overall, I think this paper is nearly ready to publish and I have only a few comments for the authors to consider.

Line 3: should be "significantly affect" not effect

Line 14: "compared to the three other datasets…" This creates confusion because the first sentence of the abstract lists three datasets in total. It was only when I read the rest of the paper that I learned the authors meant GRUAN, IASI L2 and ERA5. I suggest either listing GRUAN in the first sentence or removing the word "three" from this sentence.

Line 15: "moist layer height of about 1.3 km…" I wonder if 1.3 km can realistically be called a "significant" bias (Line 14) given that the satellite sounding retrievals have a vertical resolution between 1 and 4 km, depending on pressure. Can the authors elaborate on this?

Line 55: "We address this gap in this study." While I loosely agree with the authors that the study of EMLs are underrepresented in hyperspectral IR product evaluations, I think it prudent to add a qualifier to this statement to suggest that the study of EMLs go beyond what the authors present in this paper. I suggest one of following edits: "We take steps to start addressing this gap" or "We partially address this gap". EMLs are three dimensional features spanning hundreds of kilometers, can last a day (many hours!) and are associated with deep convection globally and especially in the extra-tropics. In short, these are large features on a global scale. In this paper, the authors do not do a 3-D analysis, nor do a global evaluation. Instead, they use a point source dataset (GRUAN radiosondes) at one single location as the reference set, against which all other datasets are evaluated. At best the authors can conclude that at a specific site and for a specific location within an EML feature (3-D blob), their results hold true. Would this not be more accurate? Or do the authors feel confident that their results can be extrapolated globally? If so, kindly motivate.

Line 109: "The also available purely operational IASI L2 retrieval data…" Confusing sentence. Rephrase.

Line 109: "jumps…" A more appropriate word is probably "discontinuities"

Section 2.4, Lines 129-143: The authors opted to use the relative humidity field that is reported in the CLIMCAPS L2 file on 66 pressure levels. This field is derived from the water vapor column density field [molec/cm2] retrieved directly from the IR radiances and reported on 100 pressure layers. It is possible that the vertical bias reported here is due to a shift from pressure layers (air_pres_lay) to levels (air_pres_h2o) when converting to relative humidity. Another issue, and one that is entirely the fault of the product team, is that the relative humidity field already has the boundary layer adjustment applied but this is not communicated in the technical documents (I discover to my dismay). The authors, therefore, didn't need to do this adjustment. I commend them, however, for following the science guides to a fault. In future I will be curious to learn if the authors report a similar bias when starting their analysis with the column density field instead (mol_lay/h2o_vap_mol_lay).

Lines 154-156: "…the IASI product attempts retrieval through the clouds, CLIMCAPS…represent the atmospheric state around the clouds…". Does the IASI L2 product really retrieve through clouds? Can the authors explain this algorithm component in a sentence or two? Thinking out loud, I wonder if IASI L2 uses the collocated AVHRR cloud fractions to determine which regression coefficients to apply. But even then, the cloudy regression retrieval would not represent the atmosphere through the cloud. Infrared radiance is highly sensitive to clouds and does not transmit through opaque clouds. The IR radiances, therefore, do not contain information within and under such

clouds. Can the authors elaborate on this distinction they're drawing here? This will help the reader better understand the results. As it is written and laid out currently, it appears that the authors say that there is no difference in EML detection between an algorithm scheme performing cloud clearing (aggregate footprints) and one retrieving through clouds (usually single footprint). But the IASI fields are also on aggregate footprints… I'm confused.

Lines 170-173: "As spatial and temporal collocation criteria we use 50 km and 30 min. These criteria are…conservative since the EMLs…have lifetimes of about a day." Given this, the authors could easily justify collocating the CLIMCAPS profiles to GRUAN sondes. Can the authors explain their adoption of this conservative approach? Do their results change when they adjust these criteria?

Lines 175-176: "In these cases [where multiple ERA5 pixels match up within an IASI FOV], we randomly select one of the matching pixels to assure that datapoints are only used once." I have two questions:

- Can the authors clarify what they mean by using a datapoint only once? I struggle to understand under which conditions an ERA5 pixel will be used twice. The IASI/AIRS FOVs do not overlap and therefore would not contain sets of collocated ERA pixels that share members.
- Can the authors justify their choice against averaging the ERA5 pixels within each satellite sounding FOV? The authors demonstrate how the comparison between ERA5 and GRUAN sondes can be improved by vertically smoothing the sondes, which have higher resolution. Do the authors think that their comparisons between IASI/CLIMCAPS and ERA5 can be improved by spatial "smoothing" (averaging) of the higher resolution ERA5 data?

Lines 177-178: "Applying these collocations criteria…we obtain…2500 AIRS/ERA5 collocations." I find the discrepancy in total number of data pairs confusing. It will help the reader if the authors can explain these numbers here. Also, the total number of 2500 AIRS/ERA5 collocations looks like a rounded-off number.

Lines 317-319: "Nonetheless, the number of moisture anomalies in the AIRS CLIMCAPS retrieval speaks [of] a good capability…to capture vertical moisture capability." This is a positive result as far as CLIMCAPS goes and surprised me. From the abstract and introduction, I expected only negative results for CLIMCAPS. I wonder if the authors can update their abstract to reflect the value in different retrieval approaches, as far as EMLs go.

Do the authors think that their results apply to reanalysis models in general, or to ERA5 specifically? CLIMCAPS uses MERRA-2 as a-priori for its water vapor column density retrievals and it will be interesting to know how much CLIMCAPS follows or deviates from

the MERRA-2 fields, especially since it uses an optimal estimation scheme that gives it the ability to adjust a-priori fields based on scene-specific information content from the measurements. In future the authors could include an evaluation of the averaging kernels to help make sense of this.