We thank Anonymous Reviewer #2 for taking the time to review our manuscript, engaging in collegial discussion, and providing constructive comments and suggestions. These have been helpful in improving the manuscript. The manuscript is revised as described below. For clarity, reviewer's original comments are included in black, our responses are written in blue, and the revision in the updated manuscript is marked with green.

Author response to Anonymous Reviewer #2

This study inter-compared the polar-orbiting satellite UTH datasets by including four participating datasets, two of which are new datasets and two of which have updated versions. Further authors have presented case studies for the El Nino period in terms of time series and spatial anomalies analysis, and the authors also try to conclude that during both El Niño and La Niña events, the values of the spatial anomalies in the infrared dataset appear smaller than those in microwave datasets and the spatial patterns of the four datasets are generally consistent over the deep tropics. This information is very informative but readers of the remote sensing community or users of these datasets might be interested to know about the quality of these datasets for their further utilization. It is suggested that include the quality of the datasets as per your results and also mention the finding of any systematic biases over different geographical regions in the manuscript. In that respect the present study has potential for publication after incorporation of the comments/suggestions as given below:

Thank you for the comments and the helpful suggestions. In this study the quality of the data is assessed in terms of consistency with other UTH datasets. To improve on presenting the information, we added descriptions of changes over different geographical regions, and added more quantified discussions in the text and in the conclusions of the revised manuscript.

Comments:

The abstract is not informative enough.

More informative statements with quantified values are added to the Abstract in the revision:

During a major El Niño event, UTH had higher correlations with the coincident precipitation (0.60 – 0.75) and with 200 hPa velocity potential (-0.42 - -0.64) than with SST (0.37 - 0.49). Due to differences in retrieval definitions and gridding procedures, there can be a difference of 3-5% UTH between datasets on average, and more significant anomaly values are usually observed in the microwave UTH data. Nevertheless, the tropical-averaged anomalies of the datasets are close to each other with their differences being mostly less than 0.5% over tropical domain average, and more importantly the phases of the time series are generally consistent for variability studies.

The conclusion is not informative enough.

More informative statements with quantified values are added to the Conclusions in the revision:

The infrared UTH dataset exhibits the largest proportions of dry areas at the peak of El Niño and La Niña events (more than 4% larger ratio of dry areas compared to those of MW datasets). The MW datasets have larger proportion of humid measurements during El Niño events, while during a major La Niña such as the 2010-11 event, the ratios of humid areas are close to each other among three UTH dataset (differences less than 1%), except the CMSAF dataset which overall has larger humid areas.

During a major El Niño, the spatial correlations between UTH and SST are not high, with the correlation values in the range of 0.37-0.49. In the meantime, the spatial correlations between UTH and precipitation are higher, ranging in 0.60-0.75. The infrared dataset has lower correlation values (usually more than 0.1 smaller) with SST, precipitation, and 200 hPa velocity potential compared to those for the MW UTH datasets due to the smaller magnitude of anomalies in the infrared dataset.

Though there are apparent and expected differences in the values of total UTH due to differences in the definition and in the gridding procedure, the tropical-averaged anomalies of the datasets are close to each other (mostly within ±0.5% over tropical domain average), and more importantly the phases of the time series are generally consistent for variability studies.

No discussion about insitu measurements as well as reanalysis data that provides UTH in the introduction section (e.g. Radiosonde, etc.).

The following are added to the introduction section (in different paragraphs):

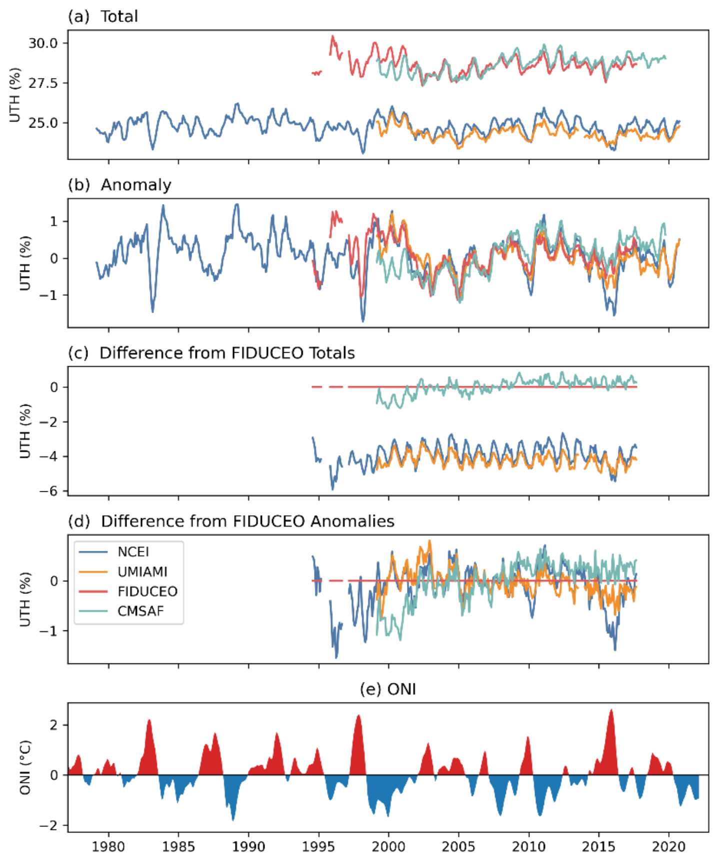
Measurement of UTH has traditionally been obtained from global radiosonde observations as part of the atmospheric water vapor profiles (e.g., Durre et al. (2018), Ferreira et al. (2019), Brogniez et al. (2015)).

The UTH datasets facilitated the evaluation of climate models and contributed to a better understanding of large-scale atmospheric processes (Allan et al., 2003; Soden et al., 2005; Chung et al., 2016; Allan et al., 2022; John et al., 2021).

The four datasets appear to be separated with two groups of similar UTH values. The values of CMSAF and FIDUCEO UTH are larger than the values of NCEI and UMIAMI UTH.

Why have authors taken the UMIAMI dataset as a reference for intercomparison? kindly add a magnitude of bias and also add some earlier analysis or findings with references.

The UMIAMI was used as a reference because it has the longest time period of AMSU-B and MHS. The conclusion on the relative differences of the datasets would be the same if a different dataset were used as the reference. The following shows a sample plot using the FIDUCEO UTH as a reference. Note that the SSM/T-2 UTH at the beginning of the FIDUCEO time series is not as stable as the AMSU-B/MHS data due to frequent missing data.



A reference that analyzed the relative differences among channel brightness temperatures of several sounders sensing the upper tropospheric humidity is added. The following is the revised text to clarify the choice of using UMIAMI as a reference:

To quantify the differences between datasets, the relative differences are calculated. Note that any of the four datasets can be used as a reference for this purpose. Among the MW UTH datasets, the UMIAMI dataset has the lengthiest time period of AMSU-B and MHS to allow for the longest MW comparison with others, and it is used as the relative reference in the calculation. Figure 1d shows that the anomaly values are mostly within ±0.5% UTH of each other, with the exceptions during El Niño events when the anomaly differences can be larger. Chung et al. (2016) analyzed the relative differences among the brightness temperatures of the channels sensing upper tropospheric humidity from HIRS, AMSU-B/MHS, and AIRS. The brightness temperature differences between the HIRS and AMSU-B/MHS were mostly within ±0.2 K.

Why did CMSAF show a negative anomaly during the year 2000 as compared to the other three datasets? Kindly explain.

This should be a result of remaining inter-satellite discontinuity in the channel brightness temperature data used for deriving UTH values. The resulting higher change rates during the common period was discussed later in the article as "The largest change rates are found in the CMSAF image, with positive changes covering most of the areas, consistent with the trend in Figure 1d. An earlier study (Lang et al.,

2020) plotted the time series of individual satellite's UTH from NOAA-15 to Metop-B for both FIDUCEO and CMSAF datasets (Figure 6 in that article). Their figure 6b showed that offsets between the UTH time series from consecutive satellite missions in the CMSAF record tend to be positive over time. When all the satellites are merged into one time series this may lead to a positive trend."

Line160: adds some suitable references.

This sentence is an observation of the time series patterns and is meant to say that the infrared clear-sky dataset has better sampling for dry regions than humid regions. It is revised to:

This indicates that the infrared clear-sky dataset may not fully capture the increase of water vapor during El Niño events due to the exclusion of very humid pixels associated with clouds, and tends to have better sampling of the dry regions.

Why does the NCEI UTH trend underestimate higher magnitude over the Tibetan Plateau as compared to the other three datasets?

The following discussion is added to the revision:

The clear-sky measurement excludes some high humidity data due to removal of cloudy pixels compared to MW datasets. The Jacobian of less-humid data has a lower peak in the atmosphere, and the lower tail of the Jacobian profile is closer to the surface (e.g., see Figure 1 in Brogniez et al. (2006)). Over a high elevation, increasing surface effect can be included in the observation radiances. A warming at the surface may contribute more to an infrared dataset due to larger portion of less-humid data.

Line304 Over the Indian Ocean, decreased UTH centered over the equatorial central Indian Ocean is surrounded with increased UTH in all datasets.

Such types of features are not seen in the CMSAF data plot and inconsistent trends are seen in the CMSAF data plot (Fig8). kindly explain.

We made a correction of the UTH changing rate description regarding the CMSAF data as below. Some explanation was given following the changing rate description as "The largest change rates are found in the CMSAF image, with positive changes covering most of the areas, consistent with the trend in Figure 1d. An earlier study (Lang et al., 2020) plotted the time series of individual satellite's UTH from NOAA-15 to Metop-B for both FIDUCEO and CMSAF datasets (Figure 6 in that article). Their figure 6b showed that offsets between the UTH time series from consecutive satellite missions in the CMSAF record tend to be positive over time. When all the satellites are merged into one time series this may lead to a positive trend."

Over the Indian Ocean, decreased UTH centered over the equatorial central Indian Ocean is surrounded with increased UTH in most datasets, except that the center of decreasing rates is confined to a smaller area around 15°S for the CMSAF UTH.