We thank Referee 2 for the constructive and thoughtful comments. The comments are helpful to improve the paper. We appreciate the positive comments about the value of the correlation analysis. We have made major revisions following the suggestions. Responses to the specific comments are detailed below (the referee's comments are cited in italics; unless stated otherwise, we will refer to the original manuscript for changes made).

1) I am missing a clear definition of the term UTWV. It should be made clear that UTWV is not an upper tropospheric humidity product with units of relative humidity but is the brightness temperature observed by an infrared water vapour channel.

A clear definition of the term UTWV, which is defined as the brightness temperature observed by HIRS channel 12, is now included in the beginning of the revised manuscript.

2) The paper is lacking a description and discussion of important product specifications:
a) As far as I know the UTWV product is affected by the surface under dry conditions as they e.g. occur at high latitudes. A specification/discussion of this aspect is needed. Then a discussion on how this affects the climate index analysis north of a certain latitude is needed. It might be adequat to restrict the figures to a certain latitude band and change figures and analysis accordingly.

Discussions are added to the revised paper on the effect of surface under dry conditions. Though HIRS channel-12 data in higher latitudes are affected by temperature, the data continue to be valuable in monitoring climate changes in terms of tracking atmospheric pattern changes. Therefore correlation analyses for higher latitudes are included.

b) To me it makes more sense to carry out radiative transfer simulations to characterize the sensitivity of the water vapour channel to clouds. This way, a product (here: UTWV) specific brightness temperature threshold which might be a function of latitude can be identified and applied. The discussion at the beginning of section 2.1 is dealing with the identification of deep convective clouds/MCS/squall lines and not with the sensitivity of the UTWV to clouds and can be shortened. I wonder if the change in T threshold affects the climate index analysis.

We did simulations from a radiative transfer model (RTTOV) during the development of cloud screening procedure. The simulations are now included in the revised manuscript to show the sensitivity of channels 12 and 8 to clouds. The discussion at the beginning of section 2.1 is largely shortened and re-written in a more logical flow. We tested changes of cloud-clearing thresholds on the correlation analyses with climate indices, and found that the effect is qualitatively insignificant.

3) An improved UTWV product is introduced. It is clear that the cloud handling is in general an advantage. However, the quality of the new product has not been evaluated. An evaluation should be included in the paper. To me the comparisons to the previous product and to the AMSU-B product are not adequate in this context because the new cloud screening can introduce uncertainties which can not be identified in the above comparisons.

As mentioned above, radiative transfer simulations are now included in the revised manuscript to show the sensitivity of cloud screening. The simulations show that clouds in the upper troposphere have distinctively different radiative properties from those of lower clouds. The differences of channels 8 and 12 for clouds at 400 hPa and above are sharply smaller than those for lower clouds. The channel 8 brightness temperatures for these upper clouds are also clearly lower. The thresholds used in the study are well positioned to identify clouds above 400 hPa. The thresholds may leave a small amount of clouds with tops between 450 and 500 hPa in the dataset, which can introduce uncertainty if the brightness temperatures are to be converted to humidity units (%). The cloud screening algorithm can remove some lower-level clouds, on the other hand. These are discussed in the revised manuscript.

4) The UTWV product is not an all-sky product. Please remove all such statements from the paper, in particular in the abstract and in the conclusions.

Removed.

5) The introduction would benefit from an overview of all available, satellite-based UTH/FTH/UTWV data records and studies related to climate index analysis and maybe also climate model evaluation. One aspect of using climate indices is to identify predictive skill on seasonal scales and beyond in climate model forecasts/hindcasts. This can also be recalled in the introduction.

An overview of available satellite-based UTH/FTH/UTWV data records is added in the introduction. More references and discussions are included on climate model evaluation studies and climate index analyses that lead to improved predictive skills on long-range scales.

6) 33416, 28 and subsequent discussion: My understanding is that the clear sky bias (e.g., Soden and Lanzante, 1996, as stated in this paper) is the bias between cloudy and clear sky conditions, that is, there is more water vapour within clouds than outside clouds, at least when global averages are considered. But is this also true relative to the above cloud case? I doubt it because in Fig. 2 minima are observed off the coast of South America and Africa. In these areas stratus clouds are present (O'Dell et al., 2008). Maybe an "above cloud bias" can be observed in certain areas/situations? Do you have a reference?

Clarifications are added to the revised manuscript that the studies by Soden and Lanzante (1996) and Lanzante and Gahrs (2000) are on the differences between all-sky and column-clear-sky conditions, and our paper shows the differences between above-cloud-sky and column-clear-sky. The areas of minima in the bottom panel of Fig. 2 correspond nicely with the areas where total cloud frequencies are relatively small, as shown in Wylie et al. (2005). The total cloud frequencies off the coasts of South America and southern Africa are only 20-40%. Even though there are stratus clouds in these areas, the relatively small cloud frequencies result in relatively small above-cloud differences compared to other areas with larger cloud frequencies. The Fig. 2 difference minima in the bottom panel also correspond to brightness temperature maxima in the top panel, indicating that these areas are very dry climatologically. The above-cloud difference minima are the combined result from both small cloud frequencies and dry conditions in these locations where persistent descending branches of the general circulations are found. Unfortunately we are unable to find a reference for "above cloud bias".

Response to technical comments:

Section 2: A short description of the previous product can be added.

A brief description of the previous product is added.

Section 2.2: This section would benefit from reformulations and can be shortened.

We think that it is important to show the detailed differences between the previous and new versions of the datasets, therefore the details are kept.

33416, 10: Change to "to better retain".

Changed the phrase.

33412, 23: Add "operated by EUMETSAT"

Added "operated by EUMETSAT".

33412, 24: Add wavelength.

The wavelength is discussed at the beginning of section 2, therefore it is not added here.

Section 3.2: I am missing a clear motivation for this comparison. It is maybe out of scope but it might be more appropriate to consider homogeneity and stability in this inter-comparison.

The upper tropospheric water vapor measurements have been routinely made from several satellite sensors, including HIRS and AMSU-B/MHS from polar orbiting satellites and MVIRI/ SEVIRI and GOES images from geostationary satellites. We consider that it is valuable to show to the general scientific community what differences and limitations are due to different sensor characteristics. The brightness temperatures (in K) are the foundation database leading to the derivation of upper tropospheric humidity (in %) datasets. Understanding the difference in brightness temperature measurements helps the understanding of differences in derived upper tropospheric humidity. Comparison between HIRS and AMSU-B is one of such activities. Inter-comparisons among more sensors are to be extended to. We agree that homogeneity and stability of datasets are very important issues. When resources allow, these will be examined in the next phase of studies along with other projects we participate in.

33427, 7: "the processing of: : :"->"water vapour measurements in the upper troposphere using HIRS."?

Edited the sentence as suggested.

33427, 9: Change to "to better retain".

Changed the phase.

334247, 20-21: This should be stated in section 4 first, together with a few comments on the differences.

The similarities and differences of correlation analyses between the previous and current versions are discussed in Section 4 of the revised manuscript.

33428, 7-9: Please remove because it was not discussed in this paper.

The sentence is removed.

Fig. 1: Use same scale/range in both plots.

The figure is re-plotted using the same scale/range in both plots.

Fig. 4: Either use same scale/range in both plots or explicitly mention a 10 K difference in the figure caption.

The 10 K difference in the color scales between HIRS and microwave plots are mentioned in the revised figure caption.

Figures are hard to read.

Most of the figures are re-plotted and larger fonts are used. Figures 5-10 in the original manuscript (Figures 7-12 in the revised version) are re-done to show the cold seasons only, and the grids having insignificant correlation (based on null hypothesis test) are stippled. These should improve the figure presentation.