We thank Referee 1 for his/her effort in reviewing the manuscript. The comments are constructive and helpful to improve the paper. We appreciate the Referee's positive comments about the importance of the work. We have made major revisions following the suggestions. Responses to the specific comments are detailed below (referee's comments are cited in italics; unless stated otherwise, we will refer to the original manuscript for changes made).

1) The cloud clearing values seem rather arbitrary. I suspect this is the nature of cloud clearing but I think the authors need to do two things: a) Motivate the choice of values. How does it relate to the channel 12 weighting function? b) Explore the sensitivity of their results to changes in the critical values.

a) A radiative transfer model (RTTOV) was used in the development of the cloud screening procedure that accounted for the channel 12 weighting function. In determining the threshold values for cloud clearing we followed the following thoughts, aided by RTTOV simulations. First, we used threshold values that are published in peer-reviewed articles and that have long been used. The majority of cloudy pixels are identified by the criterion "$T_{ch8} - T_{ch12} \leq 25$ K". The threshold value of 25 K is taken from Soden (2000). It takes into account the spectral properties and weighting functions of the upper tropospheric water vapor 6.7 μm channel and the window 11 μm channel. Second, when past publications show a range of values, we used a threshold value to ensure clear-sky pixels, but it can be in the expense of identifying some lower-cloud pixels as upper-tropospheric-cloud pixels. In the criterion "$T_{ch8} \leq 235$ K" the threshold is at the warmer side of values shown as upper tropospheric cloud-top temperatures in the past studies, which range between 208K and 245K. It is found that in cold seasons over land surfaces, this criterion can identify too many pixels as upper tropospheric clouds. And third, we compared the identified cloud amount with other independent studies. The upper tropospheric cloud amount identified in this study is found consistent with the studies of Wylie et al. (1994) and Wylie et al. (2005). More detailed discussions on the choice of values are now included in the revised manuscript.

b) During the algorithm development phase we had tested the sensitivity of the primary criterion "$T_{ch8} - T_{ch12} \leq 25$ K" by altering its threshold value. We found that by changing the "$T_{ch8} - T_{ch12}$" by 5K, it changes the identified cloudy pixels by approximately 10% (absolute value). A new figure (new Figure 1) is now included in the revision to show the RTTOV simulated relationships of the two criteria with cloud-top pressures. The figure shows that both criteria clearly separate clouds at and above 400 hPa from lower clouds.

2) I was worried that the comparison between AMSU 183 GHz and HIRS channel 12 had such a large bias. The authors did not seem concerned by this. I think the paper needs some quantitative calculations to show that this bias is not due to cloud contamination. Perhaps this could be done using ERA-Interim for the two months + RTTOV (or other radiation transfer codes) to compute the expected difference for clear sky cases.

RTTOV clear-sky simulations for the two months are now included in the revised manuscript and shown in the new Figure 6. For dry cases the simulated clear-sky differences are close to those in Figure 4g-h (Figure 5g-h in the revision). For wet cases the simulated clear-sky differences are smaller. The difference between the two figures may come from two different effects. For wet cases convective low-to-middle clouds are often present, and the water vapor contents above these clouds are larger. With increased water vapor above clouds, HIRS channel 12 detects lower brightness temperature values compared to column-cloud-free simulations.

Another effect may come from clouds with tops in the lower portion of the upper troposphere. As indicated by RTTOV cloudy profile simulations (shown in Figure 1 of the revised manuscript), a small amount of clouds with tops between 450 and 500 hPa may not have been removed by the cloud screening criteria. As HIRS channel 12 senses a thick layer of the atmosphere, these clouds in the lower portion of the troposphere can be sensed by the channel and contribute to the reduction of HIRS brightness temperature values. These discussions are added to the revised manuscript.

3) When using the entire 33 year record some issues of data inhomogeneity need discussion and possible analysis either to correct for or to demonstrate that the issues are small enough not to be important.
a) There is a big shift in the spectral response function for channel 12 when going from HIRS/2 to HIRS/3. I suspect this has been corrected for (as Shi has done so in the past) but some discussion is needed.
b) All the NOAA platforms have orbits that drift and so will alias in the climatological diurnal cycle into the record unless corrected for. Lindfors et al, 2011 and McKenzie et al, 2012 explored this. Shi is a co-author on Lindfors et al so should be aware of the issue.

a) As described in the beginning of section 2, the study uses intersatellite calibrated brightness temperatures. More discussion on channel 12 going from HIRS/2 to HIRS/3 is added in the revised manuscript.
b) The intersatellite biases were corrected using values derived from overlapping satellites' monthly zonal averages. The method corrected the combined biases from both instrument spectral function difference and sampling in different times of a day. Lindfors et al. (2011) and McKenzie et al. (2012) showed that the diurnal variation of channel 12 brightness temperatures over low-mid latitudes are generally within 0.5 K. When both passes in a day from a satellite is combined, as in the case of the present study, the intersatellite biases due to diurnal variation is generally less than 0.2 K; these are corrected as part of overall combined biases between two overlapping satellites. More discussions are added in the revised paper.

4) The correlation analysis also needs to consider statistical significance and make clear what the null hypothesis for the analysis is. I suspect for most of the indices a white noise model is adequate.

A statistical significance analysis is incorporated in the revised manuscript. Linear correlation coefficients are tested for null hypothesis. In the new correlation maps the grids that are insignificant at 90% level are marked.

5) I think the title is somewhat misleading. Channel 12 is not a pure water vapour channel but is also sensitive to temperatures. The analysis presented in this paper only considers brightness temperatures. The authors should come up with a better title that reflects this.

Following the suggestion, we change the title to "HIRS channel-12 brightness temperature data with improved clear-sky coverage in the upper troposphere".

6) I think the section on indices would benefit by referring to review papers on the various indices and also by suggesting mechanisms by which the correlations happen. For example for the NAO I'd think that the primary mechanism was transport of water vapour into the upper-troposphere by mid-latitude storms. I think this section could be shorted by dropping the AO (or NAO) as they are much the same. I think the authors would help the paper by more intelligent use of seasons – I don't think for every index we need to

see JJA, DJF and the annual mean. For the NAO are they using the same index for JJA and DJF? I think that Folland et al, 2008 suggested a different index for the summer NAO though given the weakness of the summer NAO I'm not convinced that the discussion adds much to the paper.

Following the suggestions, more review papers and more discussions on mechanisms are included in the revised manuscript. The correlation figures are re-plotted for cold seasons only. Discussions are more focused. The NAO for JJA is removed, but both NAO and AO for DJF are kept as NAO shows much more significant signals over the Atlantic.

7) There are places where the language is hard to follow and a bit staccato. I suggest the authors try to improve the language.

Efforts are made to improve the writing in the revised paper.

8) The figures are small and hard to read.

We increased the font sizes for most figures. As correlation maps in the revised manuscript now show only cold seasons, the images as well as font sizes are enlarged for better display. In correlation maps the grids having insignificant correlation based on null hypothesis test are stippled to improve the figure presentation.