Adrian Simmons – Referee #1

Thank you for your very helpful review and support for this work. We have addressed all of your points and feel that the analysis is much improved thanks to your advice. We hope you are happy with the way these points have been dealt with – as described in line below. Changes to text are shown in red. We have changed a few of the figures in addition to comments made by reviewers (e.g., Fig. 5 now shows the ratio of raw to homogenised trends as opposed to difference because this clearly shows gridboxes where the trend is of the same direction vs where the homogenisation has changed the direction of the trend). All amended figures are shown at the end of this report. We have also taken the opportunity to update to the end of 2012 and rectify a problem with the sea level pressure algorithm that became apparent during our revisions (see response to reviewer 2 comment #7). This means that all of the figures have been updated and some of the numbers in the text have changed slightly reflecting the extra year of data, the new SLP algorithm and also some retrospective changes to the ISD source database that we have no control over. The issue with retrospective changes is now noted in the text as it was not something we had thought to be an issue before – see text at the end of this response. It has not changed any of the main conclusions of this work.

Specific comments:
(i) The ERA-Interim values plotted in Figure 10 appear to be incorrect. The agreement between ERA-Interim and HadCRUH/HadCRHext is not as good as published in Simmons et al. (2010). The first panel below is extracted from Fig.10 and shows ERA-Interim to lie below the other curves from 1998 onwards. The second panel is my own update of a figure in Simmons et al. (2010). There's finer-scale detail because I plot 12-month running means, not annual means. But the better agreement is clear. It is essential that this discrepancy be resolved. I am of course willing to assist in this.

Initially we plotted the complete coverage for ERA-Interim for all gridboxes that contained any land. As was done in the Simmons et al. paper, applying a spatial match to HadISDH and weighting ERA-Interim gridboxes by their percentage land cover to reduce the ocean influence gives much closer agreement. Both this time series and the full land average are now shown for the globe, Northern Hemisphere, Tropics and Southern Hemisphere. These show clearly the better agreement in the well-sampled Northern Hemisphere. The Southern Hemisphere shows better agreement than may be expected between the complete ERA-land and HadISDH although the small land coverage here will be a factor. We note that the complete ERA-land will still likely contain some moderating effect of the oceans and the likely low biased SSTs from 2001 onwards. The new text is as follows:

"For the globally averaged annual time series there is very good agreement between all data-products both in long-term changes and inter-annual behaviour. There are sporadic deviations between the HadCRUH family, HadISDH and Dai which may be due to differences in spatial sampling or the homogenisation applied (none has been applied to Dai). The spatially matched ERA-Interim gives closer agreement with HadISDH, as expected, and agreement deteriorates outside of the well-sampled Northern Hemisphere. As noted in Simmons et al. (2010), a change in SST source ingested into ERA-Interim in 2001 led to a cooler period of SSTs henceforth, which almost certainly will have led to slightly lower surface specific humidity over this period, even over the land. This is apparent in Figure 10a-d. While Dai, HadISDH and all varieties of HadCRUH use the same source data, the methods are independent and station selection differs. ERA-Interim does ingest surface humidity data indirectly through its use for soil moisture adjustment, but also has strong constraints from the 4Dvar atmospheric model and many other data products, so it can be considered independent (Simmons et al., 2010). However, it is not impossible that the ERA-Interim reanalysis and the in situ products may be jointly affected by a contiguous region of poor station quality."

(ii) The discussion of results needs to be a little deeper. To be sure specific humidity has increased over land since the 1970s, and the El Nino peaks of 1998 and 2010 are noteworthy, and duly noted in the paper. But inspection of the first panel of Fig. 10, as reproduced on the preceding page, shows that apart from the two El Nino peaks, specific humidity over land has not increased since the late 1990s. This should be discussed, and flagged in the Abstract. Moreover, the discussion of warm years is not the correct one, in that the reference in the second sentence of section 5.4 to the warmest years on record refers to global means, not means over land, and specific humidity is presented here as means over land. The evolution of temperature averaged over land (12-month running means) is presented below, from CRUTEM4 and ERA-40/Interim. Here 2005 does not provide the highest values, and there is recent warming over land that is not accompanied by rising specific humidity. This is consistent with the fall in relative humidity reported by Simmons et al. (2010), which has continued since that paper was published. The discussion given in section 5.4 needs to be revised.

We have revised this section completely and recognise that we should have compared HadISDH with land temperatures in the first place. We now show the good agreement between HadISDH annual anomalies and those from CRUTEM4 land for the globe, hemispheres and tropics in a new figure (see new figures at end of response). We also show HadSST3 for comparison with the thinking that much of the humidity over land originates from evaporation over the oceans. We also note here that the uncertainty for HadISDH is greatest in the tropics, which is not the case for temperature. We suggest that this is due to greater year-to-year variability here in the specific humidity and due to poorer coverage which dominates the uncertainty estimates. We include a comparison of the two moistest years of 1998 and 2010 with the annual averages from HadCRUT4 but for comparison we also show 2007 which stands out as a warm year in the land record from CRUTEM4 but not an especially moist year for HadISDH and make the point that to have a strong moist signal it depends strongly on when and where the warming occurs. The new text is as follows:

Abstract:
"HadISDH is in good agreement with existing land surface humidity products in periods of overlap, and with both land air and sea surface temperature estimates. Widespread moistening is shown over the 1973-2012 period. The largest moistening signals are over the tropics with drying over the subtropics, supporting other evidence of an intensified hydrological cycle over recent years. Moistening is detectable with high (95%) confidence over large-scale averages for the globe, Northern Hemisphere and tropics with trends of 0.089 (0.080 to 0.098) g kg^{-1} per decade, 0.086 (0.075 to 0.097) g kg^{-1} per decade and 0.133 (0.119 to 0.148) g kg^{-1} per decade respectively. These changes are outside the range of uncertainty for the large-scale average which is dominated by the spatial coverage component; station and gridbox sampling uncertainty is essentially negligible on large-scales. A very small moistening (0.013 [-0.005 to 0.031] g kg^{-1} per decade) is found in the Southern Hemisphere but it is not significantly different from zero and uncertainty is large. When globally averaged, 1998 is the moistest year since records began in 1973, closely followed by 2010, two strong El Niño years. The period in between is relatively flat, concurring with previous findings of decreasing relative humidity over land."

"**5.4 Analysis of interannual variability in land surface specific humidity with surface temperature**

The strong El Niño events of 1998 and 2010 are clear in the year-to-year variability of the data, these two years being the moistest since the record began in 1973. These were also two of the three warmest years for the globe (combined land air and sea surface temperature) since 1850, the third being 2005 (Sanchez-Lugo et al., 2012). However, the land air temperature, as shown by CRUTEM4 in Figure 13 shows a number of very warm years in the mid-2000s that were not especially moist years. In fact specific humidity over the 2000s, although mostly above the long-term average demonstrates a period of plateauing more akin to global SSTs. For comparison the global SST record from the median of the HadSST3 ensemble is also shown in Figure 13, with the rationale that specific humidity over land is likely to be related to SSTs given that the majority of evaporation occurs over the ocean. Correlations of the detrended annual time series show relatively strong r values (~0.8) for both land air and sea surface temperatures with the land specific humidity for all regions except the Southern Hemisphere where the land air/specific humidity lowers to r=0.54. The stronger correlation with SSTs is perhaps to be expected here given that the Southern Hemisphere is mostly ocean. The annual average uncertainty estimates are also shown in Figure 13. It is interesting to note that uncertainty is largest in the tropics for specific humidity whereas for land air temperature it is by far the largest in the Southern Hemisphere. This is likely due to the poorer station coverage in the tropics, where year-to-year variability in specific humidity is highest.

CRUTEM4, although presenting a different atmospheric component to HadISDH uses a number of the same stations so is not truly independent. However, HadSST3 uses ship and buoy data and so is an independent record. Overall, these relatively high correlations between HadISDH and both temperature records provides further evidence that HadISDH is a reasonable estimate of large-scale land surface specific humidity. The relatively strong relationship with SST may go some way to explaining the recent plateauing in the land specific humidity record, which concurs with the decreasing RH over land found in Simmons et al. (2010). Assuming that the oceans are the major source of surface specific humidity, even over land, it follows that the slower rate of warming over the ocean cannot support evaporation at a rate sufficient to maintain increases in specific humidity in concert with land surface temperatures. This needs further investigation utilising marine surface specific humidity and marine and land RH (currently unavailable) in addition to assessing rates of change over time. This will be addressed further in future papers.

It is clear from Figure 13 that very warm years do not always lead to very moist years. While we may not expect land specific humidity to follow land air temperatures exactly given that SSTs are also an important factor, the 2000s saw warm years both in the land air and sea surface temperature records that did not constitute especially moist years. Annual anomaly maps of HadISDH and HadCRUT4 for the two warm and moist years of 1998 and 2010 are shown in Figure 14 in comparison to 2007, a very warm year over land but not exceptionally moist. It is clear that the main temperature signal in 2007 originates from the high latitudes whereas in the strong El Niño years it is in the lower latitudes. This matches the spatial distribution of high specific humidity anomalies. Following the Clausius-Clapeyron relation, the warmer lower latitudes can drive a much greater increase in moisture for a given rise in temperature, than the cooler higher latitudes. On further investigation (not shown here), the warmth of 2007 was strongest during the boreal winter and over land whereas during the 1998 and 2010 El Niño years temperature anomalies remained high from the beginning of the year through to boreal summer and featured over both land and ocean. This also helps to explain the enhanced moisture increase in the El Niño years. So, in terms of changes in surface temperature, the ‘where’ and the ‘when’ are important factors governing changes in moisture content, and the surface specific humidity record shows a strong influence from the phase of ENSO. However, the correlation of the detrended monthly HadISDH from the tropics and an optimally lagged (at 4 months) Nino 3.4 index derived from HadSST2 (Rayner et al., 2006; provided by John Kennedy) is only approximately 0.54. This suggests the importance of other factors in explaining individual monthly variability. These could be land-sea temperature differences, changes in atmospheric circulation including subsidence of the dry air in descending regions, the vertical structure of temperature anomalies throughout the atmospheric column, and other modes of variability. ”

(iii) HadCRUH is a dataset that contains values of both specific and relative humidity. HadISDH is not a complete replacement for it as HadISDH currently provides values of only specific humidity. It

is of course pleasing to see it noted on page 5158 that HadISDH paves the way for a relative humidity product, but a bit more explanation could be given as to why one was not produced at the same time as the specific humidity product, as this would have enabled HadCRUH to be superseded. One can understand why this might be, as there are places in the discussion where issues for specific humidity can be put to one side as they are associated with cold regions and low absolute values, which may not be so for relative humidity.

Our initial thinking was to make this paper a 'proof of concept' that the Pairwise Homogenisation test could be used effectively with specific humidity data with a second motivation of trying to get the paper submitted in time for IPCC deadlines of (July 2012). Including RH involved more complex treatment of the measurement uncertainty and also requires some careful thinking in terms of homogenisation to make sure adjustments made to both the RH and specific humidity are physically consistent. We didn't want to rush this. As it turns out, it was more important to take the time to do things carefully than rush for the IPCC deadline so we didn't submit until September but including the RH is still another step which would have taken considerably more time. We have improved the discussion of this as follows:

End of Introduction:
"HadCRUH also included relative humidity. We intend to include relative humidity and other related variables into HadISDH at a later date. This will involve the development of measurement uncertainty estimates specific to each variable and ensuring consistency across all variables after application of homogenisation procedures. Given that both of these are novel ventures it was felt that they could be dealt with more thoroughly in a separate paper. "

(iv) Page 5138, last line: There is a spurious "to" at the beginning of the line.

Now removed – thanks.

(v) Page 5140, line 15: "Instrumentation" could be changed to "measurement", as "instrumental" appears earlier in the line.

Good point – thanks.

(vi) Page 5156, line 5. ERA-Interim is largely independent, though if Td measurements were to be systematically and persistently biased over a significant region by a rather small amount, is there not a risk that this could creep into both ERA-Interim and HadISDH, notwithstanding the sophistication of the latter's various checks on the data?

There is a chance of a spatially consistent bias being present in both HadISDH and ERA-Interim. This paper focuses on the large scale features but we hope that later papers will explore the spatial detail of HadISDH and other products including other variables that are physically related to humidity (e.g., temperature, total column water vapour, etc.). We have amended the text to make this caveat clear as follows:

"For the globally averaged annual time series there is very good agreement between all data-products both in long-term changes and inter-annual behaviour. There are sporadic deviations between the HadCRUH family, HadISDH and Dai which may be due to differences in spatial sampling or the homogenisation applied (none has been applied to Dai). The spatially matched ERA-Interim gives closer agreement with HadISDH, as expected, and agreement deteriorates outside of the well-sampled Northern Hemisphere. As noted in Simmons et al. (2010), a change in SST source ingested into ERA-Interim in 2001 led to a cooler period of SSTs henceforth, which almost certainly will have led to slightly lower surface specific humidity over this period, even over the land. This is apparent in Figure 10a-d. While Dai, HadISDH and all varieties of HadCRUH use the same source data, the methods are independent and station selection differs. ERA-Interim does ingest surface humidity data indirectly through its use for soil moisture adjustment, but also has strong constraints from the 4Dvar atmospheric model and many other data products, so it can be considered independent (Simmons et al., 2010). However, it is not impossible that the ERA-Interim reanalysis and the in situ products may be jointly affected by a contiguous region of poor station quality."

(vii) Page 5159, line 2007. Perhaps some words such as ", and it provides an important [a much-needed][a vital] complement to the reanalysis data that have provided monitoring since then." could be added after "2007".

We agree that this is useful additional text and have added the following:

"and it provides a valuable complement to the reanalysis data that have provided monitoring since then."

Anonymous Referee #2

Thank you for your support on this work. Your comments provide very useful areas for improvement. I hope you are happy with how we have addressed these – as described in line below. Changes to the text are shown in red. We have changed a few of the figures in addition to comments made by reviewers (e.g., Fig. 5 now shows the ratio of raw to homogenised trends as opposed to difference because this clearly shows gridboxes where the trend is of the same direction vs where the homogenisation has changed the direction of the trend). All amended figures are shown at the end of this report. We have also taken the opportunity to update to the end of 2012 and rectify a problem with the sea level pressure algorithm that became apparent during our revisions (see our response to your comment #7 related to the station pressure). This means that all of the figures have been updated and some of the numbers in the text have changed slightly reflecting the extra year of data, the new SLP algorithm and also some retrospective changes to the ISD source database that we have no control over. The issue with retrospective changes is now noted in the text as it was not something we had thought to be an issue before – see text at the end of this response. It has not changed any of the main conclusions of this work.

Specific comments:

1. Fig. 10a shows small differences between this and other similar products, including some without homogenization (e.g., Dai 2006). Does this mean the surface q data are relatively homogeneous and the adjustments made in this analysis have only small impacts? It would be nice if you can compare global and regional time series of q from the cases with and without the homogenization, and make some comments regarding the impact of the adjustments on regional and global q trends.

This is a really good point. We have now added time series and trends for the major regions (Globe, Hemispheres and Tropics) to Figure 5. Please note these trends and uncertainty ranges differ very slightly from those shown in Figure 10 which have been calculated using the monthly series. Figure 5 shows that the homogenisation process has made little difference to the regional averages. Without any benchmarking of the homogenisation it is difficult to assess robustly whether or not the homogenisation process has done a good job. However, the consistency of the story across regions and reasonable comparisons with temperature (shown later) suggest that the homogenised data are reasonable. Hence, as there isn't very much difference between the raw and homogenised data we could conclude that the data are relatively homogeneous. It is clear that inhomogeneities in the raw data are not large enough to introduce any spurious long-term trends when analysed over large scales but individual stations may contain large errors that have been smoothed over by the averaging processes. We have updated the text around Figure 5 and also Figure 10, which now shows regional averages for HadISDH verses the other data products. This shows that agreement is less good away from the Northern Hemisphere. Please note that Figure 5a now shows the ratio of trends in the raw data to those in the homogenised data so the reader can see clearly where the trends are similar/very different and importantly where they are not of the same direction. The text has been adjusted as follows:

Section 3.2:

"Figures 5 a to c show trends in the data before and after homogenisation. There is generally good agreement with 87.2 % of gridboxes being of the same sign (drying or moistening) in both the raw and homogenised data (Figure 5a and b). However, it is clear that the raw data show trends of a slightly greater magnitude (both wetter and dryer) than the homogenised data (Figure 5c). In terms of the large-scale average, homogenisation appears to have very little effect (Figure 5 d-g). The trend for the Northern Hemisphere is very slightly smaller after homogenisation and the trend in the Southern Hemisphere is slightly larger. The largest differences in the time series occur for the tropics and Southern Hemisphere. This is likely an artefact of the low spatial coverage here compared to the extra-tropical and mid-latitude Northern Hemisphere, where averaging over many stations can moderate the effect of changes to a few stations. Furthermore, the tropics include some of the largest magnitude adjustments. The fact that changes are very small on these large scales suggests that seasonal analyses on large scales (not presented here) may be reasonable despite the lack of seasonally varying homogenisation. However, we urge care when analysing over smaller regions, individual gridboxes and stations, where any remaining inhomogeneity or undesirable effect of applying flat adjustments may be larger. A set of individual stations representing some of the largest changes in trends before and after homogenisation are displayed with respect to the surrounding station network in Figure 6 a to c."

2. Fig. 10b-e show that the uncertainty associated with the incomplete area-coverage overwhelmingly dominates over the other uncertainties that are a major focus of this study (sections 4.1 and 4.2). I understand that these uncertainty estimates are useful at grid box levels, but they seems to be less important for regional and global trend analysis? If that's true, you may want to point out that in the Abstract and Summary.

We agree that the station and gridbox sampling uncertainties are small enough to be of little relevance to large scale averages and that this is an important result of the paper. We have taken your advice and added this to the abstract.

Abstract:
"HadISDH is in good agreement with existing land surface humidity products in periods of overlap, and with both land air and sea surface temperature estimates. Widespread moistening is shown over the 1973-2012 period. The largest moistening signals are over the tropics with drying over the subtropics, supporting other evidence of an intensified hydrological cycle over recent years. Moistening is detectable with high (95%) confidence over large-scale averages for the globe, Northern Hemisphere and tropics with trends of 0.089 (0.080 to 0.098) g kg^{-1} per decade, 0.086 (0.075 to 0.097) g kg^{-1} per decade and 0.133 (0.119 to 0.148) g kg^{-1} per decade respectively. These changes are outside the uncertainty range for the large-scale average which is dominated by the spatial coverage component; station and gridbox sampling uncertainty is essentially negligible on large-scales. A very small moistening (0.013 [-0.005 to 0.031] g kg^{-1} per decade) is found in the Southern Hemisphere but it is not significantly different from zero and uncertainty is large. When globally averaged, 1998 is the moistest year since monitoring began in 1973, closely followed by 2010, two strong El Niño years. The period in between is relatively flat, concurring with previous findings of decreasing relative humidity over land.
"

Section 5.3:
"For the globe, Northern Hemisphere and tropics the uncertainty range is smaller than the overall long-term trend (Figure 10 e-h). Hence we can be confident in the long-term moistening signal shown in the data over these regions. The uncertainty is dominated by the spatial coverage, but the station and sampling uncertainty will be more important for any analyses on small scales. The coverage uncertainty at the monthly scale (see Figure 13 for annual uncertainties) is largest for the Southern Hemisphere and tropics, where spatial coverage is poorest. The decadal trend estimates (with 95 % confidence limits in the median of the pairwise slopes) are shown to be 0.089 (0.080 to 0.098) g kg^{-1} per decade for the globe, 0.086 (0.075 to 0.097) g kg^{-1} per decade for the Northern Hemisphere and 0.133 (0.119 to 0.148) g kg^{-1} per decade for the tropics. The narrow ranges of the confidence limits around the trend increases our confidence in these moistening trends. For the Southern Hemisphere, which includes the drying regions of Australia and South America, the overall signal is of very slight moistening but it is not significantly different from a zero trend at 0.013 (-0.005 to 0.031) g kg^{-1}. The variability and uncertainty estimates in the Southern Hemisphere are much larger than elsewhere. This region has few data compared to the Northern Hemisphere,

both because it is mainly ocean and because station density is lower making it harder to identify and adjust for inhomogeneities. Considering these factors, in addition to the known historical changes in the ISD record, we urge caution over Southern Hemisphere trends, which remain unstable with year-to-year updates. ”

Conclusions:
“HadISDH shows widespread and significant moistening across the globe from 1973 to 2012. This is shown to be highly significant and robust to an assessment of uncertainties that for the first time accounts in an explicit and quantified manner for random, systematic and sampling effects on estimates of large-scale specific humidity averages. Moistening is strongest over the tropics. There are a few regions showing a spatially coherent drying signal: southern South America, south-western USA, parts of south-eastern USA, and parts of Australia, all essentially in the subtropics. There is generally lower confidence in these signals given the spread of the trend range. However, this creates a general picture of moistening wet regions and drying dry regions, consistent with the theory of an intensified hydrological cycle resulting from a warming globe. For large-scale averages, uncertainty is dominated by the spatial coverage component; station and gridbox sampling uncertainties are essentially negligible. Large-scale averages exhibit increasing trends that exceed the uncertainty estimate for the globe, Northern Hemisphere and tropics, suggesting that the atmosphere above the global land surface is moister now that it was in the 1970s. The moistest year on record was 1998, followed by 2010, two strong El Niño years and concurrently two of the three warmest years on record. A small moistening trend is discernible for the Southern Hemisphere although it is not statistically significant and variability, both month-to-month and annually, in addition to the estimated uncertainties, are large.”

3. Also, Fig.10b-e seems to suggest that the area-coverage are fairly constant since 1973. Is that true? A few maps showing the number of stations with each grid box with monthly q data for select years (e.g., 1973, 1990, 2011) would be helpful.

Figure 1 now includes a time series of station coverage per month from 1973 to 2012 coloured by region (Southern Hemisphere, Tropics and Northern Hemisphere). This clearly shows that station coverage is fairly steady over time. There is a tail-off from 2006 onwards. Having consulted NCDC we have established that they are implementing improvements to the ISD database which may take some time and result in both the addition and removal of some data. This has been noted in the text. We cannot do anything about this other than make a note of these issues and keep up to date diagnostics such as the number of stations available over time. So, thank you for your advice here as we are now much more aware of some of the issues with the annual updating process. The following text has been added to the paper:

“HadISDH improves coverage over North America, where, for HadCRUH, many records were short and fragmented although they actually referred to the same station. ISD has been improved in this regard since the creation of HadCRUH, and the compositing process has helped further. Central Europe and islands in the Pacific are also areas of better coverage than HadCRUH. However, 878 stations from HadCRUH are no longer in HadISDH. In particular there are now very few data for Madagascar, the Arabian Peninsula, Western Australia and Indonesia. This is mostly because of the lack of up-to-date data from those stations reaching the ISD databank through the GTS. This results in station records being too short to meet the criteria set out above. Hopefully this situation will be improved in future annual updates of HadISDH. In some cases, these stations will have failed to pass the new quality control and homogenisation routines with sufficient data. In a few cases, the compositing process may have resulted in a HadCRUH station having a different identifying number (WMO identifier) in HadISDH. Station coverage, including composites, is shown in Figures 1a and b and a full list is available alongside the data-product at www.metoffice.gov.uk/hadobs/hadisdh. Coverage remains relatively constant over time over both hemispheres and the tropics (Figure 1c). There is a slight tail-off from 2006 onwards for the Northern Hemisphere stations. In part this is due to ongoing updates for known issues with these data so it is expected that 2006+ coverage will improve in the near future (Neal Lott, *pers. comm.*). Users should note that retrospective changes are made to ISD periodically with the addition of new data or removal of old data to and from existing stations. Furthermore, new stations will be added to HadISD, and therefore HadISDH, as they become available. This will be clearly documented.“

4. It appears that the station data won’t be part of the released data set (p. 14, lines 24-28). Thus, the community will have to rely on the gridded data. The gridding used in this analysis appears to be very simple: a simple average of the station data within a 5deg x 5deg box without searching for nearby stations when few or no stations exist inside the box. This may not be the best approach as monthly surface q (like monthly T) has a spatial correlation distance of □*1000km. It would be nice if the station data are included in the released data set, so people can grid the data using other approaches. For may applications, station data (not the gridded data) will be needed.*

Thanks for raising this important point. Actually there is no reason why the station data cannot be made available in addition to the grids so we will state this in the paper and make the station data available, on request, from the hadobs website. We have added this to the text.

5. p. 6, lines 24-30: adjustments did not account for seasonal differences in inhomogeneities, and thus this homogenized monthly data set may be used best for annual trend analysis, may not be suitable for seasonal trend analysis despite of its monthly resolution. People will likely use it for seasonal change analysis. Should the authors warn such applications in advance? Of course, if the authors do not find major impacts of the adjustments from homogenization, then maybe the data are ok for seasonal trend analysis too?

We did test to see whether adjustments could be applied seasonally by scaling by the annual cycle. However, achieving the precise magnitude of the recommended adjustment over the homogeneous sub-period was not straightforward and we remained unconvinced that what we had done to the data was any more defensible than a flat adjustment. While a flat adjustment is not ideal, it is easy to retain and make available the actual size of the adjustment made. To date there are no feasible automated global scale methods for applying seasonal homogenisation to temperature data, for which methods are far more advanced/tried and tested than for humidity. As a first version product we felt that keeping things simple this time around was preferable. So, in response to your comment, we have not yet investigated the suitability of HadISDH for seasonal analyses. Given that the effect of homogenisation on the large scale trends is small it is quite likely that conducting a seasonal analysis for some regions would be sound. Our advice would be to look at the data in as much detail as possible before drawing conclusions. We have updated the text as follows:

“Figures 5 a to c show trends in the data before and after homogenisation. There is generally good agreement with 87.2 % of gridboxes being of the same sign (drying or moistening) in both the raw and homogenised data (Figure 5a and b). However, it is clear that the raw data show trends of a slightly greater magnitude (both wetter and dryer) than the homogenised data (Figure 5c). In terms of the large-scale average, homogenisation appears to have very little effect (Figure 5 d-g). The trend for the Northern Hemisphere is very slightly smaller after homogenisation and the trend in the Southern Hemisphere is slightly larger. The largest differences in the time series occur for the tropics and Southern Hemisphere. This is likely an artefact of the low spatial coverage here compared to the extra-tropical and mid-latitude Northern Hemisphere, where averaging over many stations can moderate the effect of changes to a few stations. Furthermore, the tropics include some of the largest magnitude adjustments. The fact that changes are very small on these large scales suggests that seasonal analyses on large scales (not presented here) may be reasonable despite the lack of seasonally varying homogenisation. However, we urge care when analysing over smaller regions, individual gridboxes and stations, where any remaining inhomogeneity or undesirable effect of applying flat adjustments may be larger. A set of individual stations representing some of the largest changes in trends before and after homogenisation are displayed with respect to the surrounding station network in Figure 6 a to c."

6. p. 3, lines 37-44: Why can't you define e always with respect to water? Isn't e in eq. (3) with respect to water?

It is true that the specific humidity is calculated based on the water vapour pressure. However, although the vapour pressure is the partial pressure exerted by water vapour to define the vapour pressure we assume that a wetbulb thermometer has been used. This requires knowledge of whether the wetbulb was actually an icebulb or not to reduce errors in the conversion process. If we assumed that the wetbulb was always above 0 degrees then we may introduce a small moist bias in the humidity data observed during subzero conditions. The assumption that humidity was always observed using a wetbulb thermometer will not always be the correct one but at present we have no way of identifying sensor type on the global scale.

7. p. 4, top paragr.: Why don't you surface pressure (Ps), which is available at most stations?

Surface pressure is available at the majority, but not all stations. Also, during the QC process we found a number of issues with the station pressure. So to maximise the station coverage for humidity we decided to use derived station pressure using station elevation and standard pressure. But on revisiting this section (during review) we became aware that use of the standard MSLP (1013.25 hPa) was inappropriate and likely to be introducing unnecessary errors given that MSLP can vary +/- 40 hPa both across the globe and seasonally. Furthermore, our simple equation assuming a change of 1 hPa per 10m was inaccurate. We have investigated a more accurate method, while still avoiding using actual station pressure for the reasons given above. This results in a slightly larger annual cycle in the absolute values, especially in the higher elevation stations, and very small changes to anomalies for occasional months of 0.1 hPa. The station trends were altered by a very small amount, reaching a maximum change of +/- 0.04 g/kg per decade. We have added the following text to explain this new method and have highlighted the sections relevant to justifying not using station P in red:

"A climatological monthly mean station pressure component is used for calculating q. The ideal would be to use the simultaneous station pressure from HadISD. However, this is not always available, or of suitable quality and so we give preference to maximising station coverage with a trade off of very small potential errors. Climatological monthly mean sea level pressure (P_{msl}) is obtained from the 20th Century Reanalysis V2 (20CR [Compo et al., 2010]; data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, http://www.esrl.noaa.gov/psd/). This is available for 2° by 2° grids and has been averaged over the 1976 to 2005 climatological period to match that used for the humidity data. For each station the closest gridbox is converted to climatological monthly mean station level pressure (P_{mst}), using station elevation (Z in metres) and station climatological monthly mean temperature T (in kelvin), by an equation based on the Smithsonian Meteorological Tables (List, 1963):

$$P_{mst} = P_{msl}\left(\frac{T}{T+0.0065Z}\right)^{5.625} \qquad \text{Eq. 6}$$

Using a non-varying station pressure introduces small errors at the hourly level. These will be largest for high elevation stations. For stations at 2000m and temperature differences (from climatology) of ± 20 °C an error in q of up to 2.3 % could be introduced. However, the majority of stations are below 1000m where potential error for ± 20 °C reduces to ~1 % and then 0.5 % for 500m. We assume that during a month the station pressure will vary above and below the estimated P_{mst} and so essentially cancel out. Using a non-varying station pressure (year-to-year) ensures that any trends in q originate entirely from the humidity component as opposed to changes in T introduced into station pressure indirectly through conversion from mean sea level pressure. Hence, for studying long-term trends in q anomalies this method is sufficient. However, users of actual monthly mean q should be aware of the small potential errors here."

8. p. 12, lines 21-23: Could you provide more details on how you derive the certainties for the regionally averaged q values from the uncertainties at each grid box?

We have improved the description of derivation of regional uncertainties from the gridbox scale and also added a description of how we compute annual regional uncertainties as these are now shown in the new Figure 13 comparing temperature with HadISDH in response to a comment from Reviewer 1. We have added the following text:

"To explore the uncertainty over large-area averages (Sections 5.3 and 5.4) the spatial coverage uncertainty is estimated and combined with the station and sampling uncertainties, after Brohan et al. (2006) for the globe, Northern Hemisphere, tropics and Southern Hemisphere. As the spatial coverage of the gridded data is not globally complete and varies from month to month, this uncertainty needs to be accounted for when creating a regional average time series. To estimate the uncertainties of these large-area averages, which are based on incomplete coverage, we use the ERA-Interim reanalysis product due to its good agreement with the *in situ* surface humidity (Simmons et al. 2010). For each month in the HadISDH q anomalies, the ERA-Interim q anomalies from all matching calendar months are selected (i.e., for a January in HadISDH, all Januaries in ERA-Interim are selected). The ERA-Interim fields are then masked by the spatial coverage in HadISDH for that particular month and a cosine-weighted regional average is calculated. The residuals between these masked averages and the full regional average are then calculated. From the distribution of these residuals the standard deviation is extracted and used as the spatial coverage uncertainty for that HadISDH month in the regional time series. The sampling and station uncertainties are estimated from the individual sampling and station uncertainties for each grid box, and then combined with the overall coverage uncertainty for the region in question. On a month-by-month basis, the sampling and station uncertainties from each gridbox are treated as independent errors, and so the regional sampling and station uncertainty is the square-root of the sum of the normalised cosine-weighted squares of the individual gridbox uncertainties. Individual components (station, gridbox sampling and spatial coverage) are also treated as independent, and so root-sum-squared as appropriate to obtain the final 2σ uncertainty on the area average time series.

To obtain the annual uncertainties, the autocorrelation of the different uncertainty components needs to be accounted for as well as possible. The sampling uncertainty is treated as uncorrelated between months in Brohan et al. (2006), and so each of the uncertainties is independent, and the annual sampling uncertainty is the root-sum-square of the monthly uncertainties, normalised by 12 to account for the number of months. The station uncertainty, however is treated as completely autocorrelated, and so the annual station uncertainty is the mean of all 12 monthly uncertainties. For the annual coverage uncertainty, the comparison between ERA-Interim and HadISDH q fields is repeated for annual averages (as for monthly). The three individual components are then combined as described above. We note that the treatment of the station uncertainty as completely autocorrelated, and the sampling uncertainty completely uncorrelated is an approximation, as these uncertainty components are themselves combinations of separate estimates of the uncertainty from different sources. The climatology component (Eq. 7-10) for example, although uncorrelated between months, is correlated across years (i.e., January to February is uncorrelated, but January in year 1 to January in year 2 is correlated)."

9. It is unclear whether gridboxes without observations for individual months will have values (of climatological data), like in CRU TS data sets. It is best to assign a missing data code to these boxes because may users will take the climatological values as real observations.

There has been no interpolation over gridboxes that do not contain stations so HadISDH climatologies remain gappy because we have only used climatologies from the stations included in the HadISDH dataset. We have now noted in the text that the presented climatologies are simple averages of the stations present in the gridbox rather than trying to represent the 'true' climatology of the region within the gridbox. For example, where a gridbox contains complex topography the gridbox climatology will be biased to the sampled stations only rather than attempting to interpolate to the mean elevation of that gridbox. The following text has been added:

"Gridbox estimates (for all quantities) use only stations within the gridbox, all weighted equally: there is no interpolation of information from surrounding gridboxes or accounting for any elevation sampling bias (Brohan et al., 2006; Jones et al. 2012). Both the absolute values and the anomalies relative to the 1976-2005 reference period are gridded in addition to the monthly climatologies calculated over the reference period. The standard deviation of all contributing stations is also given for each gridbox month, providing an estimate of gridbox variability. Where only one station contributes, an arbitrarily large standard deviation of 100 is given so that these can be easily identified. Station numbers for each gridbox month are also recorded."

New Figures:

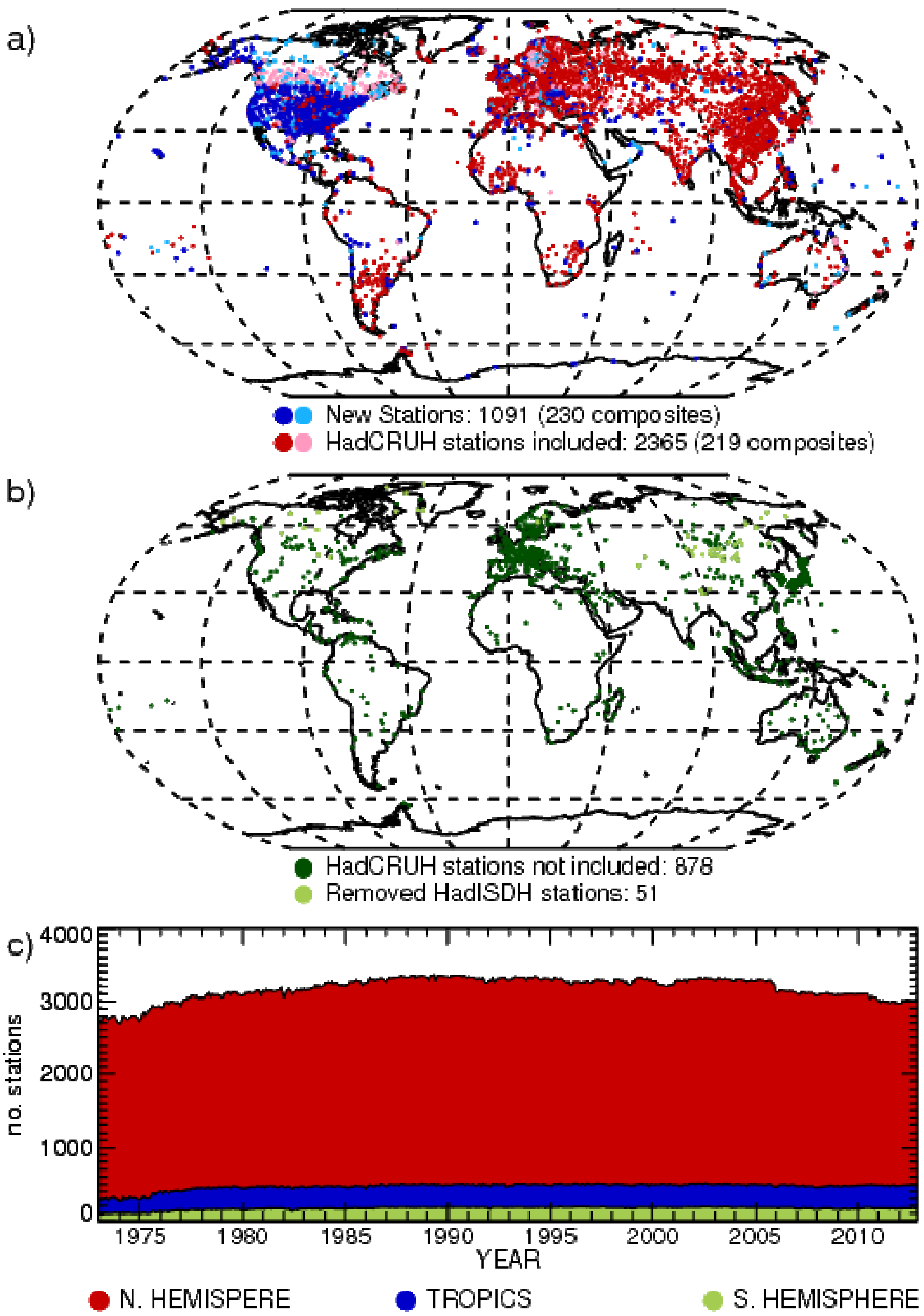

Figure 1. Station coverage comparison between HadCRUH and HadISDH. a) Station coverage in HadISDH. Stations in red/pink were also in HadCRUH. Stations in blue/turquoise are new. Pink and turquoise stations are stations that are composites of more than one original source station. b) Stations from HadCRUH that are no longer in HadISDH (dark green) and HadISDH stations with subzero specific humidity issues after homogenisation that are not included in any further analyses (light green). c) Station coverage by month for HadISDH, coloured by region (N. Hemisphere = 20°N-90°N, Tropics = 20°S-20°N, S. Hemisphere = 20°S-90°S). The tail-off from 2006 onwards is likely due to ongoing improvements to the ISD historical archive. Station coverage should improve over this period with future updates of HadISDH.

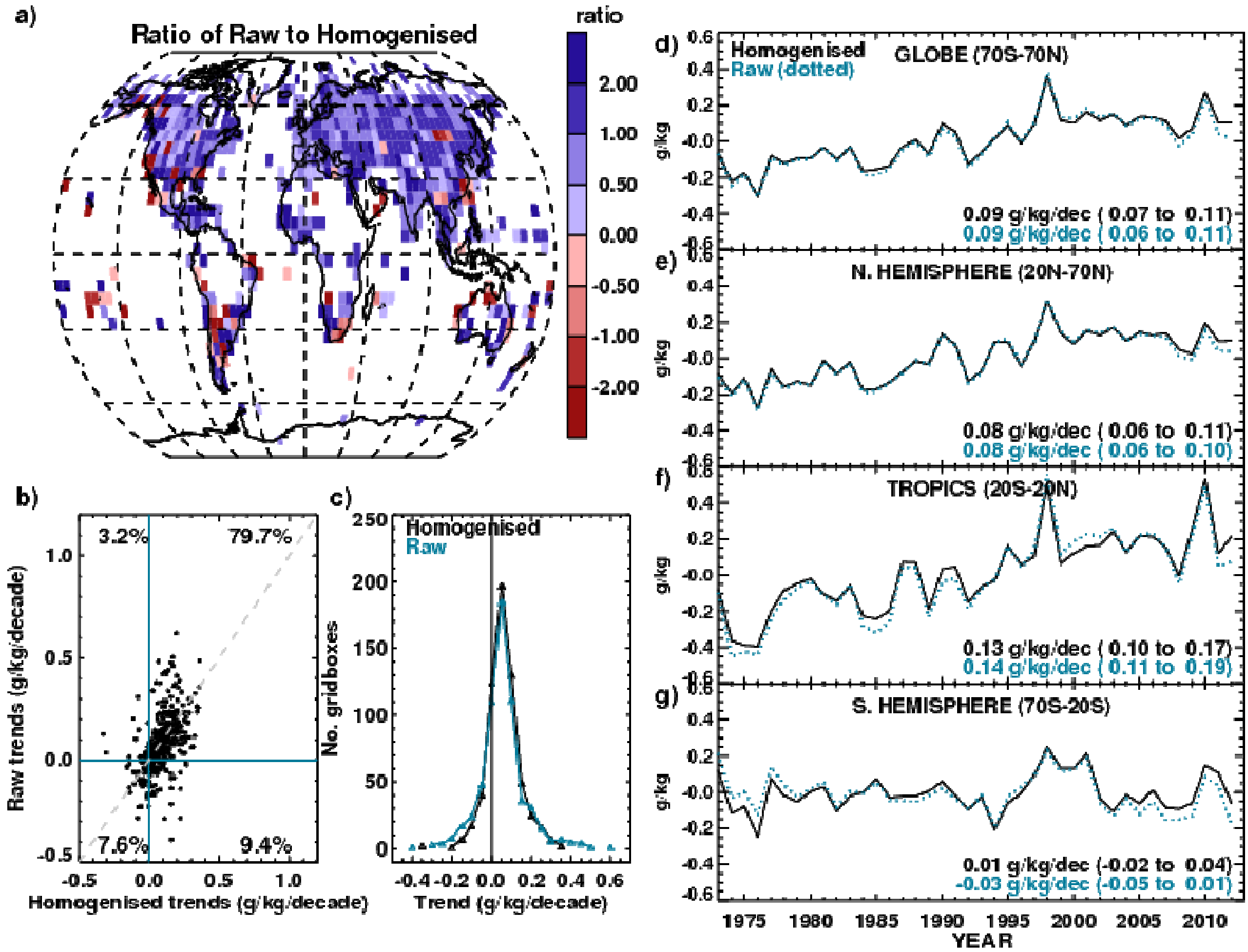

Figure 5: Difference between trends in HadISDH before and after the pairwise homogenisation process. a) Ratio of decadal trends from the raw HadISDH compared to homogenised HadISDH (trend methodology is described in Figure 9). Note non-linear colourbars. b) Scatter relationship between homogenised and raw decadal trends for HadISDH. The percentage of gridboxes present in each quadrant is shown. c) Distribution of gridbox trends for the homogenised and raw data. d-g) Large scale area average annual anomaly time series and trends for homogenised HadISDH and the raw data relative to the 1976-2005 climatology period.

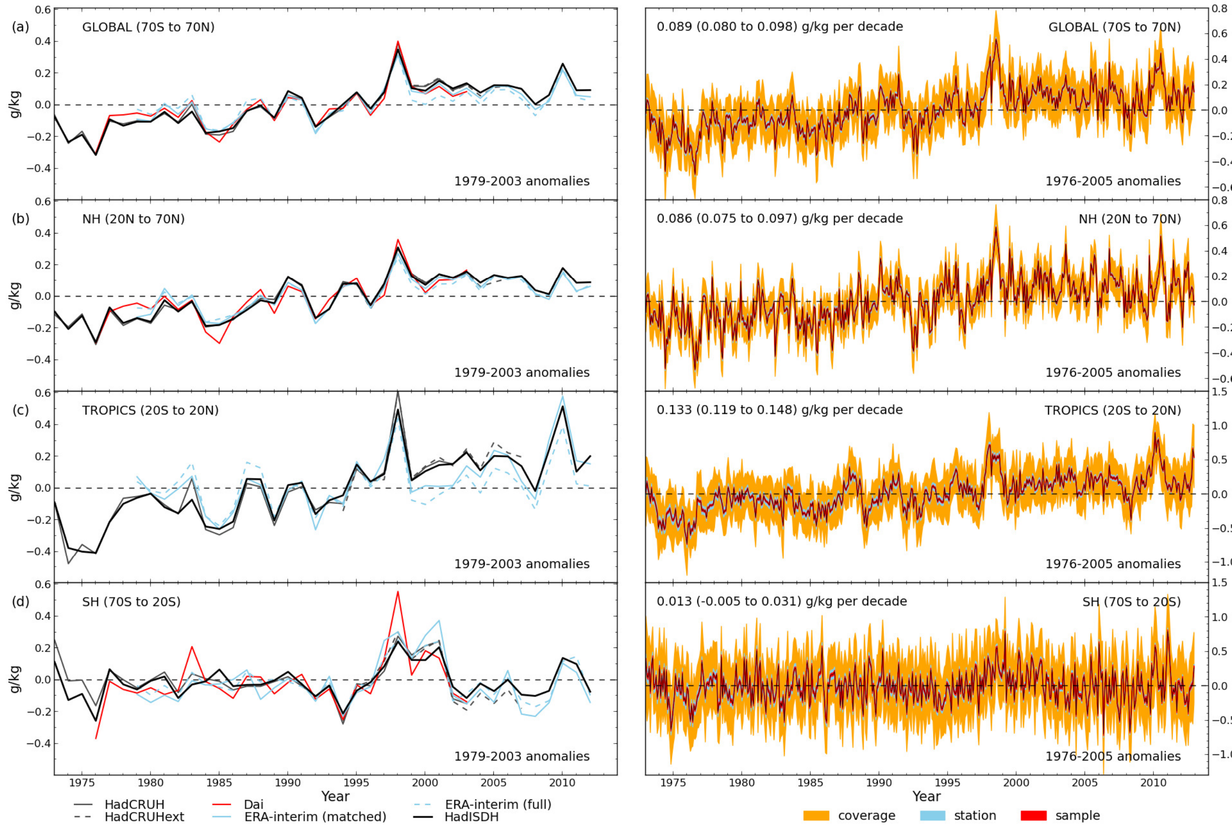

Figure 10: Time series of large-scale average specific humidity over land for HadISDH and existing data-products. a-d) Annual time series from all other global surface humidity products given a zero mean over the common period of 1979-2003. Dai covers 60°S to 70°N. ERA-Interim has been weighted by % land coverage in each gridbox and is shown both spatially matched to HadISDH and with complete coverage. e-h) Monthly time series (relative to the 1976-2005 climatology period) for HadISDH with 2σ uncertainty estimates. The black line is the area average (using weightings from the cosine of the latitude). The red, blue and orange lines show the +/- combined uncertainty estimates from the gridbox sampling uncertainty, the station uncertainty and the spatial coverage uncertainty respectively. Trends are shown for each region for the period 1973-2012. These have been fitted using the median of pairwise slopes as described in Figure 9 with the 95 % confidence intervals shown. Where these are both of the same sign (i.e., the globe, Northern Hemisphere and tropics) there is high confidence that trends are significantly different from zero.

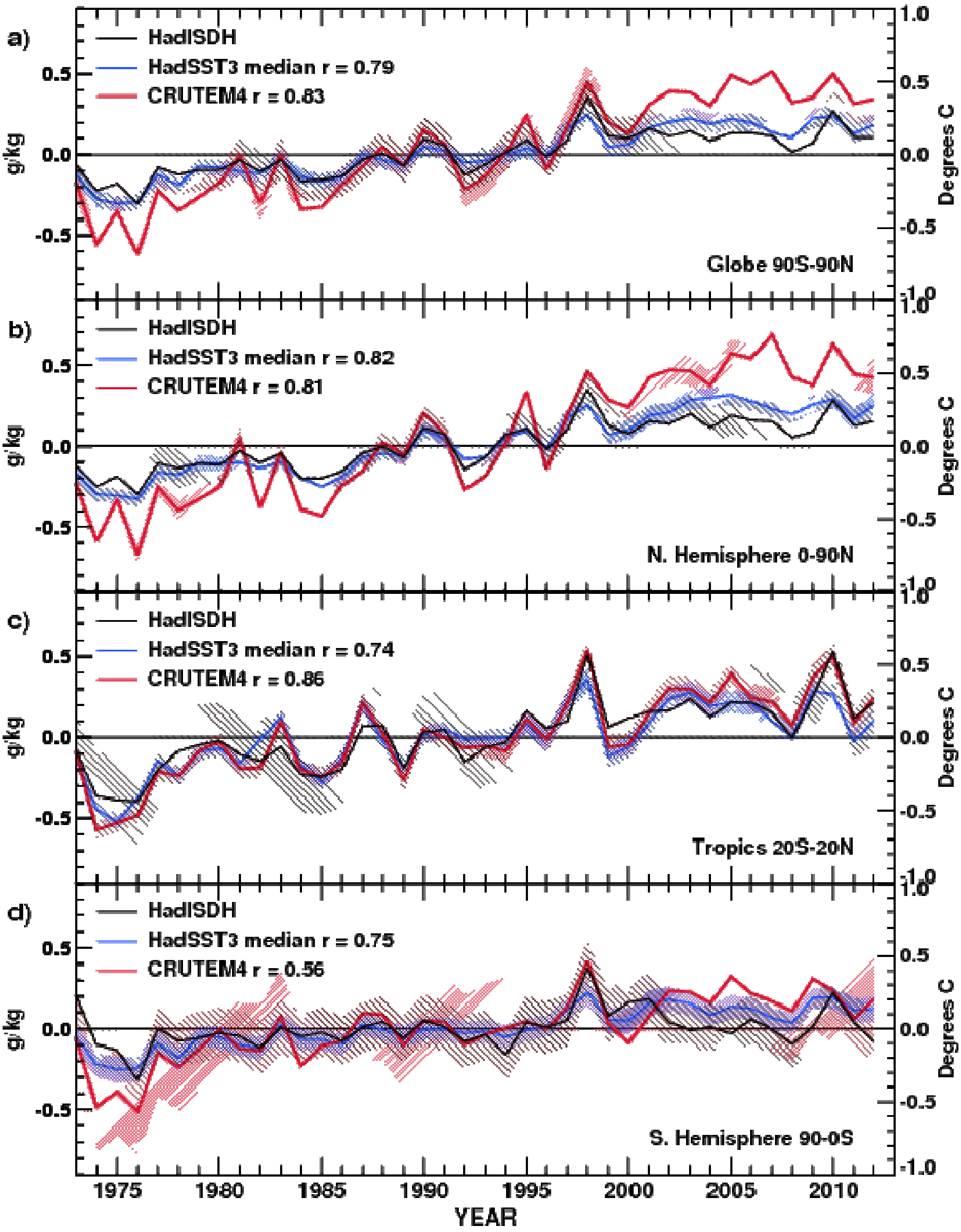

Figure 13. Comparison of large scale annual average time series from HadISDH land specific humidity with land surface air temperature from CRUTEM4 (Jones et al. 2012) and sea surface temperature from HadSST3 (Kennedy et al. 2011a, b) including uncertainty ranges. Temperature data have been adjusted to have a zero-mean over the 1976-2005 climatology period of HadISDH. Correlations between the land air temperature and SST and land surface humidity have been performed on the detrended time series.

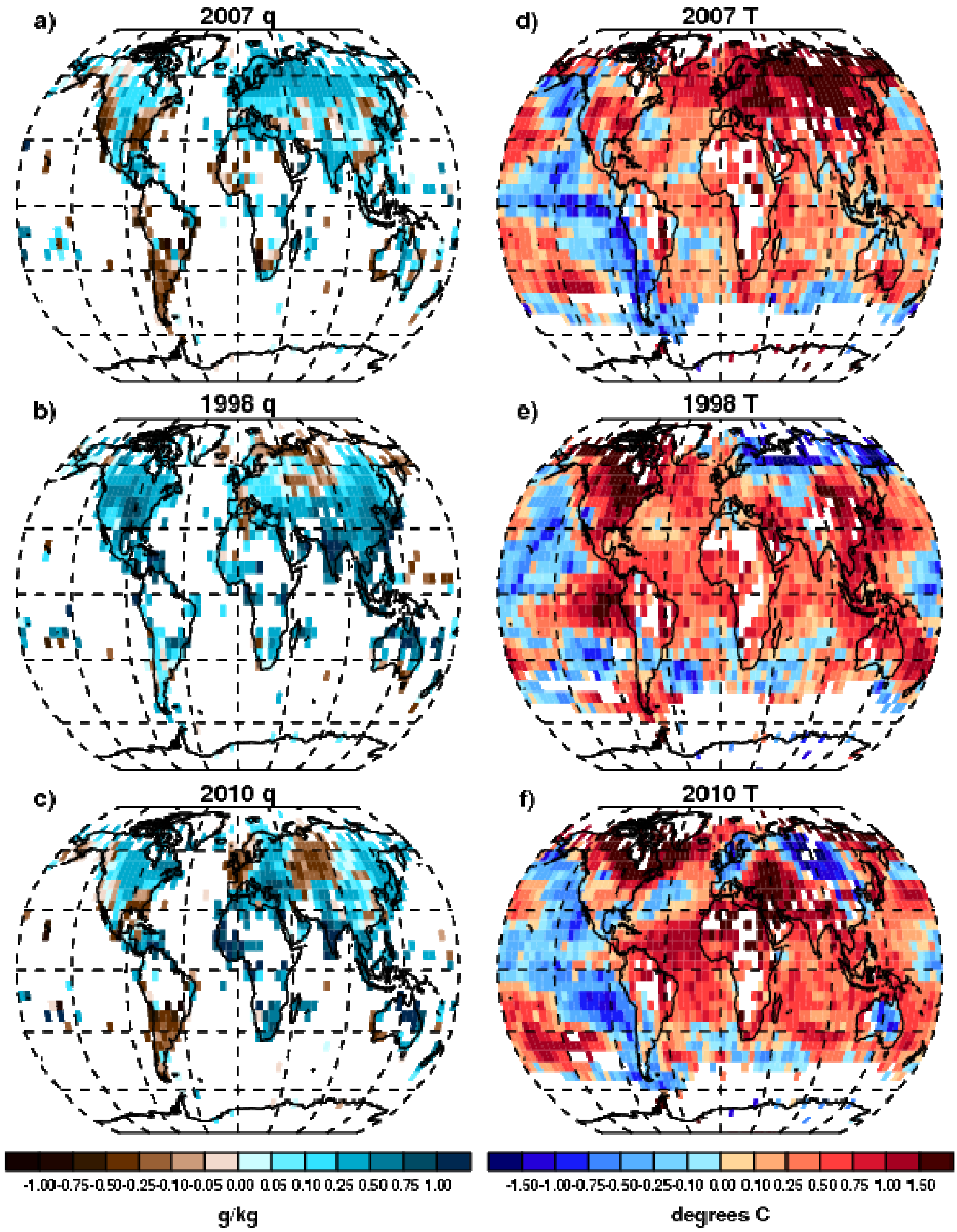

Figure 14 (was figure 13). Annual average anomalies (from the 1976-2005 climatology period) for HadISDH specific humidity and HadCRUT4 (Morice et al. 2012) temperature, for the two moistest years within the HadISDH record (1998 and 2010) which were also among the warmest years since records began in 1850, and one of the warmest years in the land record from CRUTEM4 (2007:see Figure 13) that was not simultaneously very moist. Note non-linear colour bars.

Other changes made to the paper:

Updating HadISDH to include 2012:
We wished to update HadISDH to include 2012 and as the paper had not yet been published we preferred to include an as up to date product as possible. Also, when preparing the time series of station coverage in response to reviewer 2's comments we noticed that there had been a failure in the download of a number of northwest North American stations during 2007. All figures and quoted statistics now cover the period 1973 to 2012 unless otherwise stated. This has resulted in small changes both because of the extra year (and also the improved SLP algorithm) but also because the ISD source database is undergoing continual updates to its historical record which can lead to both the addition of station years where they were previously missing and also the removal of station years where a problem has been noted and a solution is being applied. We have gained a few stations and lost a few others because of the station completeness thresholds. At a station level this has lead to quite a few changes in trends, in some cases a change of trend direction. This also follows through to the gridbox level in a few regions. The most impacted region is North America where a number of stations are undergoing problem fixing. For the large-scale regions the story is still the same – significant moistening trends for the globe, Northern Hemisphere and tropics and a very small positive but not significant trend for the Southern Hemisphere.

We could wait for these ISD updates to be finalised but this is likely to take the best part of this year and we cannot say that there will not be further changes down the line. We have added some text to make the reader aware of these historical changes and noted that when updating HadISDH comparisons will be made and documented on our website:

Section 5.1:
"Users should note that annual updates to HadISDH will likely also involve some changes to the historical record as the ISD source database is undergoing continual improvements to its historical archives. This can result in the addition of some stations into HadISDH that will then have sufficiently long data series. It may also result in the loss of some stations where ISD updates have resulted in their removal or merges with another record. There may be loss or addition of years of data for stations that remain in HadISDH. In some cases this may change the underlying station trends. While using gridbox average anomalies mitigates the effects of this instability somewhat, some notable differences could persist through to the gridbox level. Changes are unlikely to affect the large-scale features of the data. In updating from 2011 to 2012 HadISDH trends large-scale average changed minimally (±0.01 g/kg per decade). Comparisons will be made after each update and documented at www.metoffice.gov.uk/hadobs/hadisdh/."

Section 6:
"Furthermore, there is some instability resulting from continual ISD updates and improvements to the historical data, as noted in Section 5.1. For each update an assessment will be made of any resulting differences in HadISDH. This will be documented on the website. "

Given this update the values in the text and abstract have been edited as follows:

Abstract:
"Moistening is detectable with high (95%) confidence over large-scale averages for the globe, Northern Hemisphere and tropics with trends of 0.089 (0.080 to 0.098) g kg^{-1} per decade, 0.086 (0.075 to 0.097) g kg^{-1} per decade and 0.133 (0.119 to 0.148) g kg^{-1} per decade respectively. These changes are outside the uncertainty range for the large-scale average which is dominated by the spatial coverage component; station and gridbox sampling uncertainty is essentially negligible on large-scales. A very small moistening (0.013 [-0.005 to 0.031] g kg^{-1} per decade) is found in the Southern Hemisphere but it is not significantly different from zero and uncertainty is large."

Section 5.3:
"The decadal trend estimates (with 95 % confidence limits in the median of the pairwise slopes) are shown to be 0.089 (0.080 to 0.098) g kg^{-1} per decade for the globe, 0.086 (0.075 to 0.097) g kg^{-1} per decade for the Northern Hemisphere and 0.133 (0.119 to 0.148) g kg^{-1} per decade for the tropics. The narrow ranges of the confidence limits around the trend increases our confidence in these moistening trends. For the Southern Hemisphere, which includes the drying regions of Australia and South America, the overall signal is of very slight moistening but it is not significantly different from a zero trend at 0.013 (-0.005 to 0.031) g kg^{-1} per decade. The variability and uncertainty estimates in the Southern Hemisphere are much larger than elsewhere. This region has few data compared to the Northern Hemisphere, both because it is mainly ocean and because station density is lower making it harder to identify and adjust for inhomogeneities. Considering these factors, in addition to the known historical changes in the ISD record, we urge caution over Southern Hemisphere trends, which remain unstable with year-to-year updates."

Conclusions:
"A small moistening trend is discernible for the Southern Hemisphere although it is not statistically significant and variability, both month-to-month and annually, in addition to the estimated uncertainties, are large."

Other small changes:
We have made what we feel to be general improvements to the text. These are highlighted in red throughout. We hope you agree that they improve the text and the work in general.