HadISDH: An updateable land surface specific humidity product for climate monitoring

Kate M. Willett[1], Claude N. Williams Jr.[2], Robert J. H. Dunn[1], Peter W. Thorne[3], Stephanie Bell[4], Michael de Podesta[4], Phil D. Jones[5,6] and David E. Parker[1]

[1]{Met Office Hadley Centre, FitzRoy Road, Exeter, UK}
[2]{NOAA's National Climatic Data Center, Asheville, NC, USA}
[3]{Cooperative Institute for Climate and Satellites North Carolina, NCSU and NOAA's National Climatic Data Center, Asheville, NC, USA}
[4]{National Physical Laboratory, Teddington, UK}
[5]{Climatic Research Unit, University of East Anglia, Norwich, UK}
[6]{Center of Excellence for Climate Change Research / Dept of Meteorology, Faculty of Meteorology, Environment and Arid Land Agriculture, King Abdulaziz University, P. O. Box 80234, Jeddah 21589, Saudi Arabia}

Correspondence to: K. M. Willett (kate.willett@metoffice.gov.uk)

To be submitted to Climate of the Past

Abstract

HadISDH is a near-global land-surface specific humidity monitoring product providing monthly means from 1973 onwards over large-scale grids. Presented herein to 2012, annual updates are anticipated. HadISDH is an update to the land component of HadCRUH, utilising the global high resolution land surface station product HadISD as a basis. HadISD, in turn, uses an updated version of NOAA's Integrated Surface Database. Intensive automated quality control has been undertaken at the individual observation level, as part of HadISD processing. The data have been subsequently run through the pairwise homogenisation algorithm developed for NCDC's US Historical Climatology Network monthly temperature product. For the first time, uncertainty estimates are provided at the gridbox spatial scale and monthly time scale.

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.

1 Introduction

Specific humidity at the surface is, on a physical basis, expected to increase commensurate with rising surface temperatures, where the presence of liquid water is not a limiting factor (Held and Soden, 2000). This has been observed over recent decades (Dai, 2006; Willett et al., 2008; Berry and Kent, 2009) with specific humidity increases in excess of 7 % per kelvin (as expected from the Clausius-Clapeyron relation) for some regions over 1973-1999 (Willett et al., 2010). Surface water vapour drives a positive feedback effect, supplying the upper atmosphere with additional water vapour through vertical mixing processes. Here, it acts as a greenhouse gas, modifying the radiation budget and augmenting climate change. Water vapour is also an important component of the Earth's atmosphere for a number of additional reasons beyond determining climate sensitivity. The amount of water vapour in the atmosphere, quantified here as specific humidity, is a crucial element within the hydrological cycle: it governs heavy rainfall amounts where a large fraction of the water is often rained out (Trenberth, 1999). A number of variables are now showing what appears to be an intensified hydrological cycle (e.g., precipitation – Zhou et al., 2011 ; ocean salinity – Durack et al., 2012; evaporation – Brutsaert and Parlange, 1998), which is consistent with large-scale increasing water vapour concentration. Through latent heat, water vapour stores and releases energy, which can then be transported around the globe. Increasing water vapour also has implications for regulation of thermal comfort, increasing the risk of heat stress or heat related health problems in humans (Taylor, 2006) and impacting milk yields in cattle (e.g., Segnalini et al., 2011; Vujanac et al., 2012) amongst other physiological impacts on ecosystems more generally.

Since early in the 21st Century however, humidity increases over land have abated somewhat as global land temperatures have continued to rise. This has been observed as a decrease in the relative humidity and a plateauing in the specific humidity (Simmons et al., 2010; Willett et al., 2012). Simmons et al. suggest a link to the observed greater warming over the land than over the oceans in recent years. Potential mechanisms for such warming asymmetry have been discussed in the literature (e.g., Brutsaert and Parlange, 1998; Joshi et al., 2008; Rowell and Jones, 2006). Much of the moisture over the land comes from evaporation over the oceans, so if the air over the ocean surface warms more slowly than that of the land, then the saturated vapour pressure (water-holding capacity) will also increase more slowly over the ocean. Therefore, evaporation over the oceans is unlikely to increase at a rate high enough to sustain constant relative humidity (and hence proportionally increasing specific humidity) over the warmer land mass. Large-scale changes in the atmospheric circulation may also play a part and reduced moisture availability over land may lead to increased partitioning of incoming energy into sensible heating as opposed to evaporation (latent heating). This further escalates the warming over land and may diminish specific humidity increases. Whatever the drivers or processes, the crucial issue is how well we can characterise the true changes in surface humidity. Without a robust estimate of the observed behaviour the potential for false conclusions or inferences is substantial.

Previously, HadCRUH, a quality-controlled and homogenised global surface humidity product, has been widely used to look at these changes. However, it was last updated in 2007 and an improved version, extending spatial coverage and with capacity for operational annual updates, is required for near-real-time monitoring activities. Here we describe the creation of the land-surface specific humidity component of an envisaged next generation HadCRUH product: HadISDH (Met Office Hadley Centre [in collaboration with the National Oceanic and Atmospheric Administration's National Climate Data Center, the National Physical Laboratory and the Climatic Research Unit] ISD humidity product). This builds upon the new hourly land-surface dataset HadISD (Dunn et al., 2012) which is a quality-controlled database of global synoptic data since 1973. HadISDH will be the first operational *in situ* land-surface specific humidity product, and also the first to provide an estimate of uncertainties in the data. This product is designed for assessing year-to-year changes over large scales. While the data are intended primarily for scientific research, they are freely available to all.

Section two describes the source data and processing. Section three describes the building process including the pertinent aspects of the HadISD quality control suite and the applied homogenisation procedure. Section 4 describes the development of the uncertainty model for both the station data and the gridded product. Methods for exploring these uncertainties in following analyses are also documented. An analysis of recent changes is given in Section 5 followed by the logistics for using the product in Section 6. Conclusions are drawn in Section 7.

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.

2 Data Source and Processing

HadISDH uses the global high-resolution quality-controlled land-surface database HadISD as its source. HadISD was designed for studying extreme events and provides hourly to six-hourly temperature (T), dewpoint temperature (T_d), sea level pressure (SLP) and wind speed for 6103 stations. To date, HadISD has not been homogenised. Therefore, care must be taken when looking at any long-term changes. It is described fully in Dunn et al. (2012). Elements of this processing relevant to the creation of a specific humidity dataset will be discussed here in Section 3.1. We apply additional processing to make HadISDH suitable for assessing long-term trends over large scales (Section 3.2).

The station data for HadISDH are essentially the same as for HadCRUH: NOAA's National Climate Data Center's Integrated Surface Database (ISD) (Smith et al., 2011). This is available online from http://www.ncdc.noaa.gov/oa/climate/isd/index.php. For HadISDH, 3456 stations are found that have sufficient length of data after passing through the quality control and homogenisation procedures (there are 51 additional stations that are sufficiently long but not included due to homogenisation issues – Section 3.2). In order to be able to calculate a reliable climatology, each station must have at least 15 years of data within the 1976-2005 climatology period for each month of the year, where each month must contain at least 15 days. To prevent biasing towards night or day, or biases arising from systemically changing observation times aliasing into the record, there must be at least four observations per day, with at least one in each eight hour tercile (0-8hr, 8-16hr, 16-24hr) of the day. HadISDH includes 1091 stations that were not in the specific humidity land component of HadCRUH. Furthermore, a total of 449 stations are the result of compositing multiple stations where they appeared to be the same station. For example, certain countries changed their WMO identifier code leading to changes in station reporting ID over the Global Telecommunication System (GTS), which is the basis for ISD. Without such compositing many Canadian, Scandinavian and Eastern European stations would be truncated or treated as two stations artificially. Unfortunately, the compositing does not manage to resolve the WMO identifier change over eastern Germany. Compositing was done during the HadISD processing and is fully documented therein (Dunn et al., (2012)).

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 Figs. 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 (Fig. 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.

The quality controlled (see Section 3.1) HadISD T_d are converted to specific humidity (q) using the same equations as for HadCRUH (Table 1 – Eqs. (1) to (5)). First, T_d are converted to vapour pressure (e) using Eq. (1). The wet-bulb temperatures (T_w) are then calculated using Eq. (5). Where T_w values are below zero, values of e are recalculated with respect to ice (Eq. [2]). This assumes that the wet-bulb was indeed an ice-bulb at that time and that the measurement was taken with a wet-bulb thermometer as opposed to a resistance or capacitance sensor. This assumption will be incorrect in some cases, especially in the later record where more automated sensors are in use. This potentially introduces a dry bias in q where resistance or capacitance sensors are used when the ambient temperature is near or below 0 °C because e calculated with respect to ice is lower than that with respect to water at the same temperature (*A Guide to the Measurement of Humidity*: NPL/IMC, 1996). Given the increasing propensity in the record for such measurements, unless the effects are detected and accounted for in the homogenisation this would tend to yield a spurious drying signal in locations and seasons where sub-freezing temperatures are frequent. However, absolute values of specific humidity are small under such conditions so absolute errors will be small even if they are large in percentage terms. They will not affect records in seasons with temperatures above freezing. Without metadata for all 3456 stations it is impossible to correct for this and so it remains an uncertainty in the data, but it should bear little influence on the large-scale assessments for which this product is intended. From e, Eq. (3) is used to calculate q.

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):

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

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.

3 Building the data-product

3.1 Quality control

Synoptic data contain random and systematic errors which must be removed as far as is possible to ensure robust climate analyses. The random errors can be caused by instrument error, observer error or transmission error. As part of the HadISD processing a suite of quality control tests were designed for use with hourly synoptic data. These tests have been optimised with the aim of removing random errors while retaining the 'true' extremes. The quality control suite included tests particular to humidity and also neighbour intercomparisons. It is an automated procedure, necessitated by the large number of stations and observations. It is fully documented in Dunn et al. (2012), and the HadISDH input stations are freely available for research purposes as part of the HadISD dataset at www.metoffice.gov.uk/hadobs/hadisd.

The HadISD QC comprises 14 tests which looked at 6103 stations selected from the ISD database after compositing. These tests are more sophisticated than those conducted for HadCRUH as they have been designed iteratively by validation with stations where specific problems were known or record values documented, and then further tuned to optimise test performance. Like HadCRUH, a set of three logical checks are included to test for humidity measurement failures. The first tests for supersaturation: where T_d exceeds T, the T_d observations are removed. If this occurs for more than 20 % of the observations within a month, the whole month is removed. The second is for occurrences of the wet-bulb wick drying out, either through reservoir drying or freezing, which again assumes the majority of humidity measurements were taken using psychrometers. This test uses dewpoint depression: if there are ≥ 4 consecutive observations spanning 24 hours or more where the dewpoint depression is <0.25 °C, T_d is flagged, unless simultaneous observations of precipitation or fog are present, which may indicate a true high humidity event. The leeway of 0.25 °C is added to account for instrumental error in either the T or T_d measurement. Finally, a dewpoint cut-off check is done, following the discovery in Willett et al. (2008) that T_d observations can be systematically absent when T exceeds apparent threshold values in hot and cold extremes. Similar behaviour has been documented for radiosondes (e.g., McCarthy et al., 2008). Most quality control tests are variable specific such that a flagged value does not lead to removal of observations for other parameters at the same time step. However, there are a number where flags for T and T_d are linked. When checking for overly frequent values, T_d observations coincident with flagged T observations are also flagged, and where T observations exceed WMO record values for that region, the T_d values are also removed. There is also a neighbour comparison where suspect values can be removed, but also, flagged values can be recovered should they agree with unflagged neighbouring values. No such comparison was made in HadCRUH.

For HadISDH stations data removal is highest in the regions of greatest data density (North America and North-western Europe) as shown in Fig. 2a, b. This is similar for both T and T_d but with a higher percentage of T_d data removed, especially around the tropics. This is likely an artefact of higher observation density (fewer missing data and higher temporal frequency and reporting resolution) within a station giving the internal station QC tests greater power, and higher station density giving the neighbour QC test greater power. Greater data density will increase the sensitivity to outliers, thus improving the signal-to-noise ratio. Unfortunately, this means that there is a greater chance of poor data remaining in regions where station and data density are low. This underlines the importance of improving both current station coverage and historical data rescue and thus support for these efforts through initiatives (e.g., ACRE http://www.met-acre.org/, Allan et al., 2011). For T_d, in total 78.1 % of stations have ≤ 1 % of hourly data removed and 98.0 % of stations have ≤ 5 % of hourly data removed. For T, 89.9 % of stations have ≤ 1 % of hourly data removed and 99.2 % of stations have ≤ 5 % of hourly data removed.

3.2 Homogenisation

The monthly mean values are likely to contain systematic errors due to changes in instruments, station moves, incorrect station merges, changes in observing practices or changes to local land-usage. For this reason, the monthly mean *q* data are reprocessed to detect and adjust for undocumented changepoints. There are now a number of available homogenisation algorithms that have been developed and benchmarked for temperature and precipitation as part of the COST HOME project (Venema et al., 2012; www.homogenisation.org). However, very few are suitable to be run on large global networks, which require an automated process. The pairwise homogenisation algorithm designed for NCDC's US Historical Climatology Network monthly surface temperature record (Menne and Williams, 2009; Menne et al. 2009), and later applied to their Global Historical Climatology Network (GHCN) monthly temperature data set (Lawrimore et al. 2011), has been chosen here. This has been shown to be one of the more conservative algorithms, giving a very low rate of changepoint detection where none are actually present (false alarm rate) (Venema et al., 2012). Also, the pairwise method enables attribution of a changepoint to a station or stations in a more robust manner than a simple candidate verses composite reference series approach. In the candidate-composite reference series approach, network wide changes may be missed or wrongly attributed to a single station. Furthermore, the pairwise homogenisation algorithm has been through a substantive benchmarking assessment for the US temperature network (Williams et al., 2012). This showed that in all benchmark cases, the pairwise algorithm reduced the errors in the data. Importantly, it did not over-adjust or make the data any worse. This is the first time that the pairwise algorithm has been used on surface humidity data or indeed any data outside of station temperature records. This is also the first time that a fully automated (and reproducible) homogenisation process has been applied to global land surface humidity.

The pairwise algorithm (Menne and Williams, 2009; Williams et al., 2012) undertakes a number of sequential steps to find and adjust for suspected changepoints in the series:

1. For a candidate station a set of neighbours are selected based upon geographic proximity and monthly mean time series correlation, the latter being the dominant factor.
2. The difference series between each station and every neighbour are assessed iteratively using the standard normal homogenisation test (SNHT; Alexandersson, 1986) to locate undocumented change points. At this point both the candidate and master are tagged as potential breaks.
3. The large array of potential breakpoint locations is resolved iteratively as shown by the following overly simple illustration. A station might have 20 potential breaks assigned in close proximity and its 20 neighbours only one each. In this case it is clear that this station contains the true breakpoint. All of the changepoints are assessed together to determine the date of the changepoint. The break count for all the remaining stations is reduced by one so all then would be treated as homogeneous.
4. The changepoint is then assessed to define whether it is indeed a step-change or actually part of a local trend. Where the magnitude of a changepoint can be reliably estimated, and reasonable confidence can be assigned that it is non-zero based upon the spread of pairwise adjustment estimates arising from apparently homogeneous neighbour segments resulting from step #3, a flat adjustment is made to the mean of the homogeneous sub-period, using the most recent period as a reference. Where the magnitude of the changepoint cannot be reliably estimated, that period of data is removed. The spread of estimated changepoint magnitudes across the network also provides a 2σ estimate of uncertainty for the applied adjustment. This is fed through to the station uncertainty (Section 3.3).

Overall, the pairwise homogenisation results in an adjustment rate of approximately two per station. The adjustments applied to HadISDH are reasonably symmetrical about zero with a median of -0.07 g kg^{-1} and 90 % of the adjustments lying between –0.86 and 0.81 g kg^{-1} (Fig. 3a). The historical spread is also relatively even (Fig. 3b). The first two years and last two years are artificially free from changepoints because it is more difficult to detect changepoints close to the end of record. As all adjustments made are seasonally invariant and additive rather than proportional it is highly likely that adjustments may be biased low in some seasons and biased high in others. This is a particular problem for very dry months where adjustments to already very low specific humidity results in unphysical negative values. This occurs in 51 stations (identified in Fig. 1b) although only 13 of these have more than 2 % of their data affected. For this version of HadISDH these stations will not be included in any further analyses. There may also be issues at the saturation end where positive adjustments bring the specific humidity above the saturation level imposed by the original unhomogenised temperature data. However, it is likely in many cases that any inhomogeneities appearing in specific humidity co-occur in the dry-bulb temperature, which would change the saturation level. This suggests that seasonally varying and proportional adjustments may be a better approach for specific humidity, such that the humidity can never go below 0 g kg^{-1}. Homogenisation of specific humidity is still a relatively new endeavour. Exactly how the different types of inhomogeneity affect the specific humidity across the seasonal cycle, or even in wet verses dry years, is not well understood. On further investigation (Fig. 4a, b), there is no obvious relationship between adjustment magnitude and climatological mean specific humidity, as shown by looking at adjustment magnitude by latitude. Adjustment direction is relatively evenly spread across 0 g kg^{-1} at all latitudes. Should the relationship between adjustment magnitude and specific humidity be strong, we would expect to see the largest adjustments made in the more humid tropics. In fact, the largest adjustments occur in the extratropics. Station coverage is poorer in the tropics (Figs. 1a, b) and so the ability to detect inhomogeneities in the first place is decreased, like the effectiveness of quality control (Section 3.1 and Fig. 2). Indeed, Fig. 4a also shows that it is easier to detect smaller changepoints in well sampled regions, as shown in Menne et al. (2009) for the USA. There is little geographical coherence in adjustment magnitude or direction, as shown by Fig. 4b.

Given the complexity of seasonal adjustment magnitude, we have chosen to start with the simple approach of seasonally-invariant flat adjustments, where the transforms to the data are easily traceable, rather than making more complicated assumptions. In terms of detecting long-term trends in the anomalies over large spatial scales, this approach should differ very little from a seasonally-varying and proportional adjustment approach over each homogeneous subperiod. The absolute values, however, especially on gridbox spatial scales and sub-annual temporal scales, should be used with caution.

Figs. 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 (Fig, 5a, b). However, it is clear that the raw data show trends of a slightly greater magnitude (both wetter and dryer) than the homogenised data (Fig. 5c). In terms of the large-scale average, homogenisation appears to have very little effect (Fig. 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 Fig. 6 a to c.

There will always be changepoints (both step-changes and local trends) which remain undetected because they are either too small to detect or too close to other changepoints. It is very difficult to estimate the uncertainty remaining in the data due to missed detections and adjustments without a rigorous benchmarking exercise as has been undertaken for temperature over the USA (Williams et al., 2012). Benchmarking is a relatively new concept and so has not yet been attempted for humidity and as such, is beyond the scope of this paper.

4 Estimating an uncertainty model for specific humidity

4.1 Station uncertainty

Our estimate of the monthly mean anomaly q_{anom} is given, following Brohan et al. (2006), by:

$$q_{anom} = q_{ob} - q_{clim} + q_{adj} \tag{7}$$

where q_{ob} is the observed monthly mean; q_{clim} is the climatological monthly mean over the 1976 to 2005 reference period and q_{adj} are the adjustments applied to improve the long-term homogeneity. In fact, there is an error term, ε, inherent in each of these terms such that the true monthly mean anomaly can be described as:

$$q_{anom} = q_{ob} - q_{clim} + q_{adj} + \varepsilon_{ob} + \varepsilon_{clim} + \varepsilon_{adj} \tag{8}$$

Unfortunately, these errors cannot be quantified explicitly, and so the uncertainty, u, in each monthly mean anomaly value needs to be estimated. To determine the significance of q_{anom} we estimate the uncertainty u_{anom} that captures the likely error from each of the error terms in Eq. (8):

$$u_{anom} = \sqrt{u_{c\lim}^2 + u_{adj}^2 + u_{ob}^2} \tag{9}$$

where u_{clim} is the uncertainty in the calculation of the climatological monthly mean due to missing data (temporal sampling uncertainty); u_{adj} is the uncertainty in the adjustments applied for homogeneity; and u_{ob} is the measurement uncertainty of meteorological measurements. We now consider each of these in turn.

The standard uncertainty in the climatological monthly mean due to missing data is given by:

$$u_{c\lim} = \frac{\sigma_{c\lim}}{\sqrt{N_M}} \tag{10}$$

where σ_{clim} is the standard deviation of the N_M months making up the climatological mean of the 30-year period from 1976 to 2005.

The standard uncertainty in the applied homogeneity adjustments, u_{adj} is estimated as the quadrature sum of two terms:

$$u_{adj} = \sqrt{u_{applied}^2 + u_{missed}^2} \tag{11}$$

The first term $u_{applied}$ arises from the adjustments which *have* been applied to the data, and the second term u_{missed} arises from the adjustments which *have not* been applied to the data, but which should have been.

We estimate $u_{applied}$ from the 5th to 95th percentile spread of all possible adjustments magnitudes given by the network of pairwise evaluations, as described in Section 3.1, adjusting by a factor 1.65 to obtain a standard uncertainty (1σ, coverage factor of $k = 1$).

We estimate u_{missed}, the uncertainty arising from missed changepoints, using methods described in Brohan et al. (2006). We assume that large changepoints, shown in the tails of the distribution in Fig. 3a (black line), are well captured because they are easy to detect given the high signal-to-noise ratio. However, the small adjustments, close to 0 g kg^{-1}, are not well captured, as shown by the 'missing middle' of the distribution. The central part of the distribution can be approximated by a Gaussian distribution. A best-fit curve is then derived by merging a fitted Gaussian curve (near the centre) with those points of the actual adjustment distribution that are larger (in the wings). The standard deviation of the difference between this 'best-fit' and the actual distribution (blue dotted line) is 0.135 g kg^{-1} and provides an estimate of u_{missed}.

The uncertainty u_{ob} relates to the uncertainty of measurement of the instrument at the point of observation. The BIPM *Guide to the Expression of Uncertainty in Measurement* (BIPM, 2008) describes uncertainties as belonging to one of two categories. Type A uncertainties are those which can be estimated from analysis of repeated observations. Type B uncertainties are those which cannot be estimated by repeated observations, and so must be estimated from *a priori* knowledge of the measurement apparatus and the measuring conditions. Type B uncertainties may have randomly varying components, u_{rand}, and components which cause 'systematic' errors, u_{sys}.

In a meteorological context it is not possible to derive Type A estimates of uncertainty because the measurand – the weather – is intrinsically variable, and so the variability due to the instruments themselves cannot be isolated. Since Type A uncertainties are likely to be random and uncorrelated, they should reduce with temporal and spatial averaging to a large extent, and so be attenuated by averaging both over a month and over a gridbox. Since the station metadata do not reliably record the instrumentation used, we have derived estimates of the Type B uncertainty of an individual measurement u_i based on knowledge of hygrometers in use in the field. Until the 1980s, psychrometers were probably the most common type of hygrometer, but since then there has been a move towards electronic devices (typically capacitance sensors) and dewcels which can be more readily automated. Typically electronic devices have a lower uncertainty than psychrometers and so we can conservatively estimate u_i (Table 2) assuming that all humidity measurements were taken using aspirated psychrometers.

Psychrometer errors arise either from the use of an incorrect psychrometer coefficient, or from temperature errors in measurements of either the wet-bulb or dry-bulb (MOHMI, 1981). In general, these errors are not random or symmetrically-distributed and they may be correlated with other meteorological variables, such as wind speed. However, we expect that within any one month, the uncertainty of measurement for psychrometers will contain some random component, u_{rand}, whose effect can be reduced by averaging, and a systematic component, u_{sys}, whose effect will be unaffected by averaging.

We estimate u_i with a standard uncertainty of 0.15 °C in the wet-bulb depression above 0 °C. The resulting standard uncertainty in %rh varies from 1 %rh to 3 %rh, decreasing with increasing T and increasing with decreasing RH (*A Guide to the Measurement of Humidity*: NPL/IMC, 1996). Although not ideal, this is done at the monthly mean resolution because the homogeneity adjustments have already been applied at the monthly mean level and so it is not possible to go back to the hourly values at this stage. The concomitant uncertainty in q is estimated from the uncertainty in %rh by calculating the change in vapour pressure, e (Eq. [3]) caused by changes of ±1 standard uncertainty in %rh. Combining this with an estimate of the saturation vapour pressure calculated from simultaneous monthly mean T (Eqs. [1] and [2]), under the necessary (and in many cases incorrect) assumption that the T data are homogeneous, the resulting change in q can be estimated. The reading uncertainties of the wet-bulb and dry-bulb temperatures are unlikely to be biased and so we assume that the resulting uncertainty is randomly distributed. We thus estimate the random component of the uncertainty in the monthly mean as:

$$u_{rand} = \frac{u_i}{\sqrt{N_O}} \qquad (12)$$

In order to pass ISD quality control there must be at least 15 days of data for a monthly mean and at least four observations per day, implying $N_O \geq 60$. Hence, we conservatively use $N_O = 60$ in our calculation of u_{rand}.

There are a number of weaknesses in this approach. Firstly, these non-linear conversions (Eqs. [1] to [5]) are imperfect for monthly mean data. Secondly, when the monthly value is already close to 100 %rh the addition of uncertainty in RH can then result in estimates >100 %rh. Although it is physically possible for RH to exceed 100 %rh it is not common, nor reliably measured in operational circumstances (Makkonen and Laakso, 2005). Thirdly, errors will also be introduced because the simultaneous monthly mean *T* data have not been homogenised. This is due to the issue of maintaining physical continuity when homogenising across simultaneously observed variables which will be addressed in future work. False wet-bulb depressions may occur at 100 %rh, but the low-resolution conversion between humidity variables makes accurate detection of such cases impossible. However limiting the new RH (%rh + derived uncertainty in %rh) to 100 %rh can imply an unrealistically small variability. To counter this, we have set a minimum threshold for u_{rand} of two standard deviations below the mean by examining the u_{rand} estimates for each month for the station. All values below this threshold are assumed to be unrealistically low and are substituted with the mean value of u_{rand} for that station.

Despite the difficulties in estimating u_{rand} one clear feature emerges (Table 2). Although the fractional uncertainties are largest at low temperatures, the absolute values of specific humidity are low in this range – saturation vapour pressure varies by a factor 20 from 0 °C to 50 °C – and so contributions to the uncertainty in the specific humidity of a station will be dominated by the uncertainty during periods of high temperature and high relative humidity.

In addition to randomly-varying components, u_i the uncertainty of measurement of each station will also have contributions which do not reduce on averaging, u_{sys}. Such uncertainties arise from the limitations of the calibration, and the shortcomings of psychrometers. We have not included an explicit assessment of u_{sys} because we consider that their effect on our estimate of q_{anom} is likely to be small. The origin of this insensitivity can be seen by considering the case in which a changepoint is identified as occurring in the record from a particular station, and also the case in which no changepoints are identified. For example, where instruments or observing practices change or stations move, u_{sys} will change, and so we expect that some fraction of u_{sys} should be found during homogenisation and so should be partially accounted for in terms of u_{adj}. Additionally, when a changepoint is found by comparison with neighbouring stations, the algorithm adjusts the target station’s older data to match its newer data on the assumption that more modern measurements are likely to have lower uncertainty. Where instruments or observing practices do *not* change, then we can assume that u_{sys} will be substantially unchanged. So we expect that a substantial fraction of u_{sys} will be common to q_{anom} and q_{clim}. Thus when calculating q_{anom} we can expect this fraction of the uncertainty to cancel (Eq. [8]). However, we note that care must be taken if the final gridded data are used to estimate *absolute* values of specific humidity. For such cases, the full value of u_{sys} should be evaluated to fully capture the uncertainty. The uncertainty estimates provided alongside HadISDH will therefore be underestimates with respect to the *absolute* values.

The uncertainties are calculated as standard uncertainties (1σ), and then a coverage factor $k = 2$ is applied such that there is ~95 % confidence (2σ) that these uncertainties capture the true error. As an example, the individual uncertainty components and the combined station uncertainty are shown in Fig. 7 for station 486650 (Malacca, Malaysia, 2.267°N, 102.250°E, 9.0 m). Climatological uncertainty is constant year-to-year but has an annual cycle and is greatest during the season of greatest natural variability in *q*. Measurement uncertainty (not including u_{sys}) is usually the smallest component. It changes throughout but has a clear annual cycle due to the temperature dependence. The adjustment uncertainty is usually the largest component, reducing towards 0 g kg^{-1} because the most recent period is treated as the reference period. This is the first attempt at quantifying uncertainty in specific humidity and is a basis which will benefit from future improvements in the model design and application as a greater understanding of this issue accrues.

4.2 Gridding methodology and sampling uncertainty

HadISDH is intended for the purpose of studying change on large temporal and spatial scales, so gridding is essential. It reduces the effect of individual outliers and remaining random errors in the data. Given that station density is rather sparse over large parts of the globe, there is little value in gridding at finer than 5° by 5° resolution. For the station-rich regions, specific high-resolution grids could be produced but will not be presented here. Using 5° by 5° grids also allows comparison with other products such as HadCRUH and CRUTEM4.

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.

Station uncertainty estimates, as defined in Section 3.3, are also brought through to the gridbox level by assuming independence of, and combining in quadrature, all constituent station uncertainties and then multiplying by $1/\sqrt{(N_S)}$ where N_S is the number of stations in the gridbox at that time. Fig. 8a shows an example field of gridded station uncertainty for June 1980 in g kg^{-1}. Station uncertainty is largest around the tropics, whereas, for the CRUTEM3 temperature product in Brohan et al. (2006) it is largest at the poles. The largest component is by far the adjustment uncertainty, until the most recent years of the record where it tends towards zero as a result of choosing the most recent period as the reference period. The measurement uncertainty is comparable to the climatological uncertainty when averaged over the gridbox scale. All are generally largest in the tropics, where station density is generally least. These uncertainties are also gridded individually.

Given that there are relatively small numbers of stations within each gridbox the gridbox value is unlikely to be the true gridbox average. Some estimate of the sampling uncertainty is necessary. Following Brohan et al. (2006), the sampling uncertainty is estimated using the method laid out in Jones et al. (1997). For gridboxes with data we first estimate the mean variance of individual stations in the gridbox, $\bar{s}_i^2$, using:

$$\bar{s}_i^2 = \frac{\hat{S}^2 N_{SC}}{(1+(N_{SC}-1)\bar{r})} \qquad (13)$$

where $\hat{S}^2$ is the variance of the gridbox anomalies calculated over the 1976-2005 climatology period and N_{SC} is the mean number of stations contributing to the grid-box mean. The last term, $\bar{r}$, is the average inter-site correlation and is estimated using:

$$\bar{r} = \frac{x_O}{X}\left(1-\exp\left(-\frac{x_O}{X}\right)\right) \qquad (14)$$

where X is the diagonal distance across the gridbox and x_0 is the correlation decay length between gridbox averages. Gridbox sampling uncertainty, SE^2, is then estimated by:

$$SE^2 = \frac{(\bar{s}_i^2\bar{r}(1-\bar{r}))}{(1+(N_S-1)\bar{r})} \qquad (15)$$

However, here N_S is the actual number of stations contributing to the gridbox in each month, giving a time varying SE^2. The number of stations contributing to the gridbox mean makes a large difference to SE^2 with a 10-fold increase in stations making SE^2 an order of magnitude smaller. Sampling uncertainties, in g kg^{-1}, are shown in Fig. 8b as 2σ uncertainties. The main driver of the sampling uncertainty is the standard deviation of gridbox monthly specific humidity anomalies.

The sampling uncertainty and station uncertainty estimates are assumed to be independent and are combined in quadrature to provide a combined uncertainty statistic, shown for June 1980 in Fig. 8c as a percentage of June climatology. Station uncertainty is the largest component and dominates the combined uncertainty fields where there are data. The magnitude of the combined uncertainty relative to climatology is generally less than 5 % (for 69.3 % of gridboxes) but exceeds 10 % of June climatology in 8.0 % of gridboxes which are mostly located in parts of the subtropics. This reflects the large uncertainty in adjustments made to the data. For there to be confidence in any changes apparent in the data, these changes must be larger than the combined spread of uncertainty.

Brohan et al. (2006) also provide a bias estimate. However, for humidity over land, no such broad scale estimates have been assessed to date. While there are likely biases locally for urbanisation and land-use changes such as increased irrigation, it is assumed here that their effect at the large 5° by 5° gridbox scale is small. A recent study by Asokan et al. (2010) found changes in evapotranspiration flux resulting from irrigation over the Mahanadi River Basin in India suggesting that local water use could be important in regional climate change. Further work is needed to quantify this impact for the global scale.

Another way to explore the uncertainty would be to produce plausible ensemble estimates of HadISDH, as was done for e.g., HadCRUT4 (Morice et al., 2012) or Remote Sensing Systems' Microwave Sounding Unit product (Mears et al., 2011). This is the first time that a global humidity estimate has been given any measure of uncertainty. Creating a meaningful ensemble product that enables the uncertainty model developed here and its interdependencies through the HadISDH processing chain to be more fully explored is a future aspiration.

4.3 Using the uncertainty model to explore uncertainty in long-term trends

To explore the uncertainty in individual gridbox trends (Section 5.2) a simple 100 member ensemble of HadISDH is created, randomly sampling across the spread of the 2σ uncertainty for each individual uncertainty component (climatology, measurement, adjustment and sampling uncertainty). This is distinct from that described at the end of Section 4.2 which would be a far more rigorous exploration of the uncertainty fields. The ensemble members created here while available to users, are purely for exploring the spread of uncertainty and not to be used singularly as a plausible estimate of land surface humidity.

For each station, ten versions of anomaly time series are created by adding the actual anomaly values to random values of climatology, measurement and adjustment errors as follows:

- *Climatology error time series*:
  Ten values are randomly selected from a Gaussian distribution (μ=0, 2σ=1) for each station. The Gaussian is forced to have 2σ=1 because this then provides a ~95 % chance that the randomly selected values lie between -1 and 1. These values are then used as a scaling factor on the 2σ climatology uncertainty which has an annual cycle but is constant year-to-year.
- *Measurement error time series:*
  Ten time series are created by randomly selecting values from the Gaussian distribution for each station and each month. These are then used as scaling factors on the 2σ measurement uncertainty, such that the error randomly varies over time.
- *Adjustment error time series:*
  Ten time series are created by randomly selecting values from the Gaussian distribution for each station and each homogeneous subperiod (the period between two adjustments). These are then used as scaling factors on the adjustment uncertainty for each homogeneous subperiod.

For each station, the actual anomalies are then added to the first climatology error time series, the first measurement error time series and the first adjustment error time series to give the first station realisation. The second to tenth station realisations are created similarly. These ten realisations of each station are then gridded in the same manner as the actual HadISDH to give ten gridded realisations. For sampling error, ten values are randomly selected from a Gaussian distribution and used as scaling factors on the sampling uncertainty. The scaling factor is consistent across all gridboxes and months within each of the ten sampling error realisations. These are then combined with each of the ten gridded station realisations to give the 100 member ensemble of the gridded HadISDH.

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 (Eqs. [7] to [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).

5 Recent trends in land surface specific humidity

5.1 Validation against other land surface humidity products

It is first necessary to assess the likely quality of HadISDH before we can use it with any confidence. This has been done by comparing gridbox decadal trends with the older HadCRUH product (Fig. 9) and large-scale area average time series with all other existing products: HadCRUH (Willett et al., 2008); HadCRUHext (Simmons et al., 2010); Dai (Dai, 2006); and the reanalysis product ERA-Interim (from 1979 onwards) regridded to 5º by 5º and weighted by percentage of land within each gridbox (Dee et al., 2011; Willett et al. 2011, 2012) (Fig. 10a to d). The ERA-Interim time series are shown both spatially matched to HadISDH and with complete land coverage.

In all cases trends have been estimated using the median of pairwise slopes (Sen, 1968; Lanzante, 1996). Confidence in the trend is assigned using the 95 % confidence range in the median value. Where the intervals defined by the confidence limits are either both above zero or both below zero there is high confidence that the trend is significantly different from a zero trend. The spread of these intervals gives an estimate of confidence in the magnitude of the trends.

For the gridbox trends over the common period 1973-2003, HadISDH shows generally good agreement with HadCRUH – the key feature of widespread moistening is common to both with drying apparent in parts of South Africa, southern South America, southern Australia and New Zealand. HadCRUH shows more moistening in the tropics (e.g., Brazil, West Africa and northern India). HadISDH shows more moistening over the USA and southeast Asia. There is very little overall difference with 92.3 % of HadISDH gridboxes showing moistening verses 89.5 % in HadCRUH. It is likely that HadISDH contains fewer outlying/poor quality data issues due to the improved quality control and homogenisation methods used. There are large differences over the USA owing to the improvements in coverage and station compositing described in Section 2; see also Smith et al. (2011). In HadISDH, moistening is now far more widespread over the USA.

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, although 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 Fig. 10a to 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.

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/.

5.2 Spatial patterns in long-term changes in land surface specific humidity

The spatial pattern of long-term trends over the full 1973-2012 period (Fig. 11) shows a very similar picture (widespread moistening) to that shown for 1973-2003 in Fig. 9. There are some notable differences, with more extensive drying in parts of South America and southwestern and southeastern USA, and Mexico; and South Africa showing more moistening. Overall, there is widespread moistening which is strongest across the tropics. The subtropics over the USA, South America and Australia show drying. This is consistent with the now well-observed and documented intensification of the hydrological cycle over recent decades (Allan et al., 2010).

The gridbox trends range from approximately -0.1 g kg^{-1} to 0.3 g kg^{-1} per decade. This is comparable to the uncertainty ranges shown in Fig. 8. To explore the uncertainty in these trends, an ensemble of HadISDH is created with 100 members as described in Section 4.3. Trends are fitted to each ensemble member at the gridbox scale. The 5th percentile, median and 95th percentile trends for each gridbox (assessed individually) are shown in Fig. 12 a to c respectively. Moistening remains the main feature of all three maps so the conclusion of widespread moistening appears to be robust to the quantified uncertainties, especially across the tropics, Eurasia and northeastern North America. Drying over the southwestern USA also appears to be significant relative to uncertainty but the extratropical drying regions show relatively large uncertainty.

5.3 Long-term changes in large-scale area-average land surface specific humidity

For the globe, Northern Hemisphere and tropics the uncertainty range is smaller than the overall long-term trend (Fig. 10e to 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 Fig. 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} 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.

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 Fig. 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 Fig. 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 Fig. 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 Fig. 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 Fig. 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.

Despite the moistening (in absolute terms) shown here, other research shows that the land surface atmosphere became less saturated over recent years, as shown by the decreasing relative humidity in Simmons et al. (2010). The decrease is too recent to be defined as a long-term trend. HadISDH paves the way for a later development of a relative humidity product in addition to other humidity variables which will allow this aspect to be fully explored. In absolute terms, the globe contains more moisture over the land surface now than in the 1970s. In relative terms, this depends on the simultaneous temperature changes and whether enough water has been evaporated to sustain the relative humidity. Following Clausius-Clapeyron this would need to be 7 % for every 1 kelvin rise in temperature.

6 Data availability and logistics

The gridded product of HadISDH used here is HadISDH.landq.1.0.0.2012p. It is freely available for research purposes from www.metoffice.gov.uk/hadobs/hadisdh along with supporting material, diagnostics and also some of the source code used in development. Individual stations are also available on request. Version control will follow the HadISD format (Dunn et al., 2012) with

HadISD updates being fed through to HadISDH. HadISDH version control and format is fully described on the download webpage. The version of the pairwise algorithm used is that associated with the GHCN v3.2 release and can be downloaded from (http://www.ncdc.noaa.gov/oa/climate/research/ushcn/#phas). While great effort has been made to ensure high quality and long-term homogeneity of the data, all users are advised to use the uncertainty estimates and station numbers contributing to each gridbox mean where possible. 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. Feedback is very much appreciated and future versions/annual updates will endeavour to address any issues found. Table 3 documents the fields available.

7 Conclusions

We have presented a new improved and updatable surface specific humidity product over land, HadISDH, for the purpose of assessing long-term changes. It benefits from improved station coverage and compositing, more in-depth quality control, and more thorough and objective homogenisation. It also has uncertainties parameterised through a formal error model. HadISDH has been compared against all existing global products over their respective overlaps and shown to be in very good agreement. It is the only purely observationally based estimate that exists after 2007, and it provides a valuable complement to the reanalysis data that have provided monitoring since then. This is the first time that the pairwise homogenisation algorithm has been used for surface humidity. The close agreement with existing products suggests that the pairwise algorithm is an effective tool for homogenising the surface humidity data. Further work is necessary to thoroughly assess the strengths and weaknesses of this important process using humidity benchmark data in addition to exploring seasonally varying and proportionally applied adjustments. The uncertainty model could also be refined.

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.

It is intended for HadISDH to be updated annually so that it can be used to monitor year-to year changes in specific humidity. Future work will deliver similar products for relative humidity, vapour pressure, wet-bulb temperature and dewpoint temperature and also the simultaneously observed temperatures. Such a suite of simultaneously derived temperature and humidity products will be a valuable addition to further our understanding of the water cycle under climate change.

Acknowledgements

Kate Willett, Robert Dunn and David Parker were supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101). Phil D. Jones has been supported by the Office of Science (BER), U.S. Dept. of Energy, Grant No. DE-SC0005689. Stephanie Bell and Michael de Podesta were supported by the UK National Measurement System Programme for Engineering and Flow Metrology, and by the MeteoMet Project of the European Metrology Research Programme. We thank John Kennedy for providing an ENSO index time series and Adrian Simmons for a thorough review and help thereafter working with the ERA-Interim data. We also thank the internal reviewers at NCDC, NPL and the Met Office for their thorough reviews and advice. Support for the Twentieth Century Reanalysis Project dataset is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office.

References

Allan, R. J., Brohan, P., Compo, G. P., Stone, R., Luterbacher J., and Brönnimann, S.: The International Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative, Bull. Amer. Meteor. Soc., 92, 1421-1425, doi: http://dx.doi.org/10.1175/2011BAMS3218.1, 2011.

Allan, R. P., Soden, B. J., John, V. O., Ingram W., and Good, P.: Current changes in tropical precipitation, Environ. Res. Lett., 5, 025205 doi:10.1088/1748-9326/5/2/025205, 2010.

Alexandersson, H.: A homogeneity test applied to precipitation data, J. Climatol., 6, 661–675. doi: 10.1002/joc.3370060607, 1986.

Asokan, S. M., Jarsjö J., and Destouni, G.: Vapor flux by evapotranspiration: effects of changes in climate, land use and water use, J. Geophys. Res., 115, D24102, doi:10.1029/2010JD014417, 2010.

Berry, D. I and Kent, E. C.: Air–Sea fluxes from ICOADS: the construction of a new gridded dataset with uncertainty estimates, Int. J. Climatol., 31, 987–1001, doi: 10.1002/joc.2059, 2011.

BIPM (Bureau International des Poids et Mesures): Guide to the Expression of Uncertainty in Measurements. Joint Committee for Guides in Metrology JCGM 100, pp120, 2008. http://www.bipm.org/en/publications/guides/gum.html [accessed August 2012]

Brohan, P., Kennedy, J. J., Harris, I., Tett S. F. B., and Jones, P. D. Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850, J. Geophys. Res., 111, D12106, doi: 10.1029/2005JD006548, 2006.

Brutsaert, W. and Parlange, M. B.: Hydrologic cycle explains the evaporation paradox, Nature, 396, 30-30, doi: 10.1038/23845, 1998.

Buck, A. L., 1981: New equations for computing vapor pressure and enhancement factor, J. Appl. Meteorol., 20, 1527–1532, 1981.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D, Matsui, N., Allan, R. J., Yin, X., Gleason, Jr B. E., Vose, R. S., Rutledge, G,. Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, Q. J. R. Meteorol. Soc. 137: 1-28. doi:10.1002/qj.776, 2011.

Dai, A.: Recent climatology, variability, and trends in global surface humidity, J. Climate., 19, 3589–3606, 2006.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G.., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L. J., Bidlot, L., , Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137 (656), 553-597, doi: 10.1002/qj.828, 2011.

Dunn, R. J. H., Willett, K. M., Thorne, P. W., Woolley, E. V., Durre, I., Dai, A., Parker, D. E., and Vose, R. S.: HadISD: a quality-controlled global synoptic report database for selected variables at long-term stations from 1973–2011, Clim. Past, 8, 1649-1679, doi:10.5194/cp-8-1649-2012, 2012.

Durack, P. J., Wijffels, S. E., and Matear, R. J.: Ocean salinities reveal strong global water cycle intensification during 1950 to 2000, Science, 336 (6080), 455-458, doi: 10.1126/science.1212222, 2012.

Held, I. M. and Soden, B. J.: Water vapour feedback and global warming. Annu. Rev. Energ. Env., 25, 441–75, 2000.

Jensen, M. E., Burman, R. D., and Allen, R. G., Eds.: Evapotranspiration and Irrigation Water Requirements: ASCE Manuals and Reports on Engineering Practices No. 70, American Society of Civil Engineers, New York, 360 pp, 1990.

Jones, P. D., Osborn T. J., and Briffa, K. R.: Estimating sampling errors in large-scale temperature averages, J. Climate, 10, 2548–68, 1997.

Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M., and Morice, C. P.: Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010, J. Geophys. Res., 117, D05127, doi:10.1029/2011JD017139, 2012.

Joshi, M. M., Gregory, J. M., Webb, M. J., Sexton D. M. H., and Johns, T. C.: Mechanisms for the land/sea warming contrast exhibited by simulations of climate change, Clim. Dyn., 30, 455–465, doi 10.1007/s00382-007-0306-1, 2008.

Kennedy J. J., Rayner, N. A., Smith, R. O., Saunby M., and Parker, D. E.: Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: measurement and sampling errors, J. Geophys. Res., 116, D14103, doi:10.1029/2010JD015218, 2011a.

Kennedy J. J., Rayner, N. A., Smith, R. O., Saunby M., and Parker, D. E.: Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 2: biases and homogenisation. J. Geophys. Res., 116, D14104, doi:10.1029/2010JD015220, 2011b.

Lanzante, J. R.: Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data, Int. J. Climatol., 16, 1197–1226, 1996.

Lawrimore, J. H., Menne, M. J., Gleason, B. E., Williams, C. N., Wuertz, D. B., Vose, R. S. and Rennie, J.: An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3, J. Geophys. Res., 116, D19121, doi:10.1029/2011JD016187, 2011.

List, R. J.: Smithsonian Meteorological Tables. Vol 114. 6th Ed. Smithsonian Institution, Washington DC. pp268, 1963.

Makkonen, L. and Laakso, T.: Humidity Measurements in Cold and Humid Environments, Bound.-Lay. Meteorol., 116, 131–147,, doi: 10.1007/s10546-004-7955-y, 2005.

McCarthy, M. P., Titchner, H. A., Thorne, P. W., Tett, S. F. B., Haimberger, L., and Parker, D. E.: Assessing bias and uncertainty in the HadAT adjusted radiosonde climate record, J. Climate, 21, 817-832, doi: 10.1175/2007JCLI1733.1, 2008.

Mears, C. A., Wentz, F. J., Thorne P., and Bernie, D., Assessing uncertainty in estimates of atmospheric temperature changes from MSU and AMSU using a Monte-Carlo estimation technique, J. Geophys. Res. - Atmos., 116, D08112, doi:10.1029/2010JD014954, 2011.

Menne, M. J. and Williams Jr, C. N., 2009: Homogenization of temperature series via pairwise comparisons. J. Climate, 22, 1700-1717, doi: http://dx.doi.org/10.1175/2008JCLI2263.1.

Menne, M., Williams Jr. C. N., and Vose, R. S.: The U.S. historical climatology network monthly temperature data, Version 2, Bull. Amer. Meteor. Soc., 90 (7), 993:1007, doi: http://dx.doi.org/10.1175/2008BAMS2613.1, 2009.

MOHMI (Meteorological Office): Handbook of Meteorological Instruments – Volume 3, Measurement of Humidity (second edition), Her Majesty's Stationery Office, London, pp10-12, 1981.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset, J. Geophys. Res., 117, D08101, doi:10.1029/2011JD017187, 2012.

NPL/IMC (National Physics Laboratory and The Institute of Measurement and Control): A Guide to the Measurement of Humidity, The Institute of Measurement and Control, London, 68pp, 1996.

Peixoto, J. P. and Oort, A. H.: The climatology of relative humidity in the atmosphere, J. Climate, 9, 3443–3463, 1996.

Rayner, N.A., Brohan, P., Parker, D. E., Folland, C. K., Kennedy, J. J., Vanicek, M., Ansell T., and Tett S. F. B.: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: the HadSST2 data set, J. Climate, 19(3), 446-469, doi: http://dx.doi.org/10.1175/JCLI3637.1, 2006.

Rowell, D. P. and Jones, R. G..: Causes and uncertainty of future summer drying over Europe. Clim. Dyn.., 27, 281-299, doi 10.1007/s00382-006-0125-9, 2006.

Sanchez-Lugo, A., Kennedy J. J., and Berrisford, P.: Surface Temperature, in "State of the Climate in 2011". (Eds. Blunden, J. and Arndt, D. S.) Bull. Amer. Meteor. Soc., 93 (7), S14-S15, doi: http://dx.doi.org/10.1175/2012BAMSStateoftheClimate.1, 2012.

Segnalini, M., Nardone, A., Bernabucci, U., Vitali, A., Ronchi B., and Lacetera, N.: Dynamics of the temperature-humidity index in the Mediterranean basin, Int. J. Biomet., 55(2), 253-263, doi: 10.1007/s00484-010-0331-3, 2011.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63, 1379–1389, 1968.

Simmons, A., Willett, K. M., Jones, P. D., Thorne P. W., and Dee, D.: Low-frequency variations in surface atmospheric humidity, temperature and precipitation: inferences from reanalyses and monthly gridded observational datasets, J. Geophys. Res., 115, D01110, doi:10.1029/2009JD012442, 2010.

Smith, A., Lott, N., and Vose, R.: The Integrated Surface Database: Recent Developments and Partnerships, B. Am. Meteorol. Soc., 92, 704–708, doi:10.1175/2011BAMS3015.1, 2011.

Taylor, N. A. S.: Challenges to Temperature Regulation When Working in Hot Environments, Industrial Health, 44, 331–344, 2006.

Trenberth, K. E.: Conceptual framework for changes of extremes of the hydrological cycle with climate change, Climatic Change, 42, 327-339, 1999.

Venema, V. K. C., Mestre, O., Aguilar, E., Auer, I., Guijarro, J. A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C. N., Menne, M. J., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Prohom Duran, M., Likso, T., Esteban, P., and Brandsma, T.: Benchmarking homogenization algorithms for monthly data, Clim. Past, 8, 89-115, doi:10.5194/cp-8-89-2012, 2012.

Vujanac, I., Kirovski, D., Samanc, H., Prodanovic, R., Lakic, N., Adamovic M., and Valcic, O.: Milk production in high-yielding dairy cows under different environment temperatures, Large Animal Review, 18(1), 31-36, 2012.

Willett, K. M., Jones, P. D., Gillett, N. P., and Thorne, P. W.: Recent changes in surface humidity: Development of the HadCRUH dataset, J. Climate, 21, 5364–5383, doi: http://dx.doi.org/10.1175/2008JCLI2274.1, 2008.

Willett, K. M., Jones, P. D., Thorne, P. W., and Gillett, N. P.: A comparison of large scale changes in surface humidity over land in observations and CMIP3 GCMs, Environ. Res. Lett., 5, 025210 doi: 10.1088/1748-9326/5/2/025210, 2010.

Willett, K. M., Dai, A., and Berry, D.: Global climate, in “State of the Climate in 2010” (Eds. Blunden, J., Arndt, D. S., and Baringer, M. O.), Bull. Amer. Meteor. Soc., 92(6), S40-S41, doi: http://dx.doi.org/10.1175/1520-0477-92.6.S1, 2011.

Willett, K. M., Berry D., and Simmons, A.: Global climate, in “State of the Climate in 2011” (Eds. Blunden, J. and Arndt, D. S.), Bull. Amer. Meteor. Soc., 93(7), S23-S24 doi: http://dx.doi.org/10.1175/2012BAMSStateoftheClimate.1, 2012.

Williams, C. N., Menne, M. J., and Thorne P. W. : Benchmarking the performance of pairwise homogenization of surface temperatures in the United States, J. Geophys. Res., 117, D05116, doi:10.1029/2011JD016761, 2012.

Zhou, Y. P., Xu, K-M., Sud Y. C., and Betts, A. K.: Recent trends of the tropical hydrological cycle inferred from Global Precipitation Climatology Project and International Satellite Cloud Climatology Project data, J. Geophys. Res., 116, D09101, doi:10.1029/2010JD015197, 2011.

Table 1: Equations (1) to (5) used to derive humidity variables from dry-bulb temperature and

dewpoint temperature.

Variable	Equation	Source	Notes
Vapour Pressure calculated with respect to water (e) (when T_w > 0ºC)	e = 61121*f_w*EXP(((18.729 – (T_d/227.3))* T_d) / (257.87 + T_d)) $f_w = 1 + 7x10^{-4} + (3.46x10^{-6} *P)$	Buck (1981)	(1), substitute T for T_d to give saturated vapour pressure (e_s)
Vapour Pressure calculated with respect to ice (e_{ice}) (when T_w < 0ºC)	e = 61115*f_i*EXP(((23.036 – (T_d /333.7))* T_d) / (279.82 + T_d)) $f_w = 1 + 3x10^{-4} + (4.18x10^{-6} *P)$	Buck (1981)	(2), as above for e_s
Specific Humidity (q)	$q = 1000((0.622*e)/(P – ((1 – 0.622)*e)))$	Peixoto and Oort (1996)	(3)
Relative Humidity (RH)	$RH = (e/e_s)*100$	---	(4)
Wet-bulb Temperature (T_w)	$T_w = ((a*T) + (b\ T_d))/(a + b)$ $a = 6.6x10^{-5}*P$ $b = (409.8*e)/(T_d + 237.3)^2$	Jensen et al., (1990)	(5)

Table 2: Estimates of standard uncertainty in humidity measurements calculated in terms of equivalent psychrometer uncertainty to represent a ‘worst case scenario’. At lower temperatures the measurement uncertainty becomes large, but the low absolute specific humidity values make only a small contribution to global estimates of specific humidity. Calculations of specific humidity used Eqs. (1) to (5).

Dry-bulb Temperature (°C)	**Uncertainty in %rh given by a 0.15 °C uncertainty in wet-bulb depression**	**Specific Humidity ($g\ kg^{-1}$) at saturation**	**u_i ($g\ kg^{-1}$)**	**u_{rand} ($g\ kg^{-1}$)**
-50 and below	15	0.02	0.003	0.001
-40	15	0.08	0.012	0.002
-30	15	0.23	0.035	0.005
-20	10	0.64	0.064	0.008
-10	5	1.60	0.080	0.010
0	2.75	3.78	0.104	0.013
10	1.8	7.60	0.137	0.018
20	1.35	14.54	0.196	0.025
30	1.1	26.60	0.293	0.038
40	0.95	46.82	0.445	0.057
50+	0.8	79.85	0.639	0.082

Table 3: Description of data contained in the HadISDH CF-compliant netCDF file.

Field	Description	Dimensions	Maximum and Minimum Values
qhum_abs	Monthly mean specific humidity	72 by 36 by 468	0.000 to 23.570 g kg^{-1}
qhum_anoms	Monthly mean anomaly specific humidity from the 1976-2005 climatology period	72 by 36 by 468	-6.308 to 5.867 g kg^{-1}
qhum_std	Standard deviation of all station monthly mean anomalies within the gridbox for each month	72 by 36 by 468	0.0 to 100.0 g kg^{-1}
qhum_combinederr	Station uncertainty and sampling uncertainty combined in quadrature to give a 2σ uncertainty	72 by 36 by 468	0.031 to 2.272 g kg^{-1}
qhum_samplingerr	2σ Sampling uncertainty	72 by 36 by 468	0.002 to 0.778 g kg^{-1}
qhum_rbar	Average inter-site correlation	72 by 36	0.100 to 0.891 g kg^{-1}
qhum_sbarSQ	Estimate the mean variance of individual stations in the gridbox	72 by 36	0.030 to 10.000 g kg^{-1}
qhum_stationerr	Climatological, measurement and adjustment uncertainty combined in quadrature to give a 2σ station uncertainty	72 by 36 by 468	0.015 to 2.656 g kg^{-1}
qhum_adjerr	1.65σ adjustment uncertainty	72 by 36 by 468	0.015 to 2.186 g kg^{-1}
qhum_obserr	2σ measurement uncertainty	72 by 36 by 468	0.001 to 0.131 g kg^{-1}
qhum_climerr	1.65σ climatological uncertainty	72 by 36 by 468	0.003 to 1.144 g kg^{-1}
qhum_clims	Monthly climatologies over the 1976-2005 period	72 by 36 by 12	0.041 to 22.364 g kg^{-1}
mean_n_stations	Total number of stations within the gridbox over entire record	72 by 36	1 to 42
actual_n_stations	Actual number of stations within the gridbox for each time step	72 by 36 by 468	0 to 41
lat	Latitude in 5°	72 by 36	-87.5°S to 87.5°N
lon	Longitude in 5°	72 by 36	-177.5°W to 177.5°E
times	Months since January 1973	468	1= January 1973, 480= December 2012
months	1-12	12	1 to 12

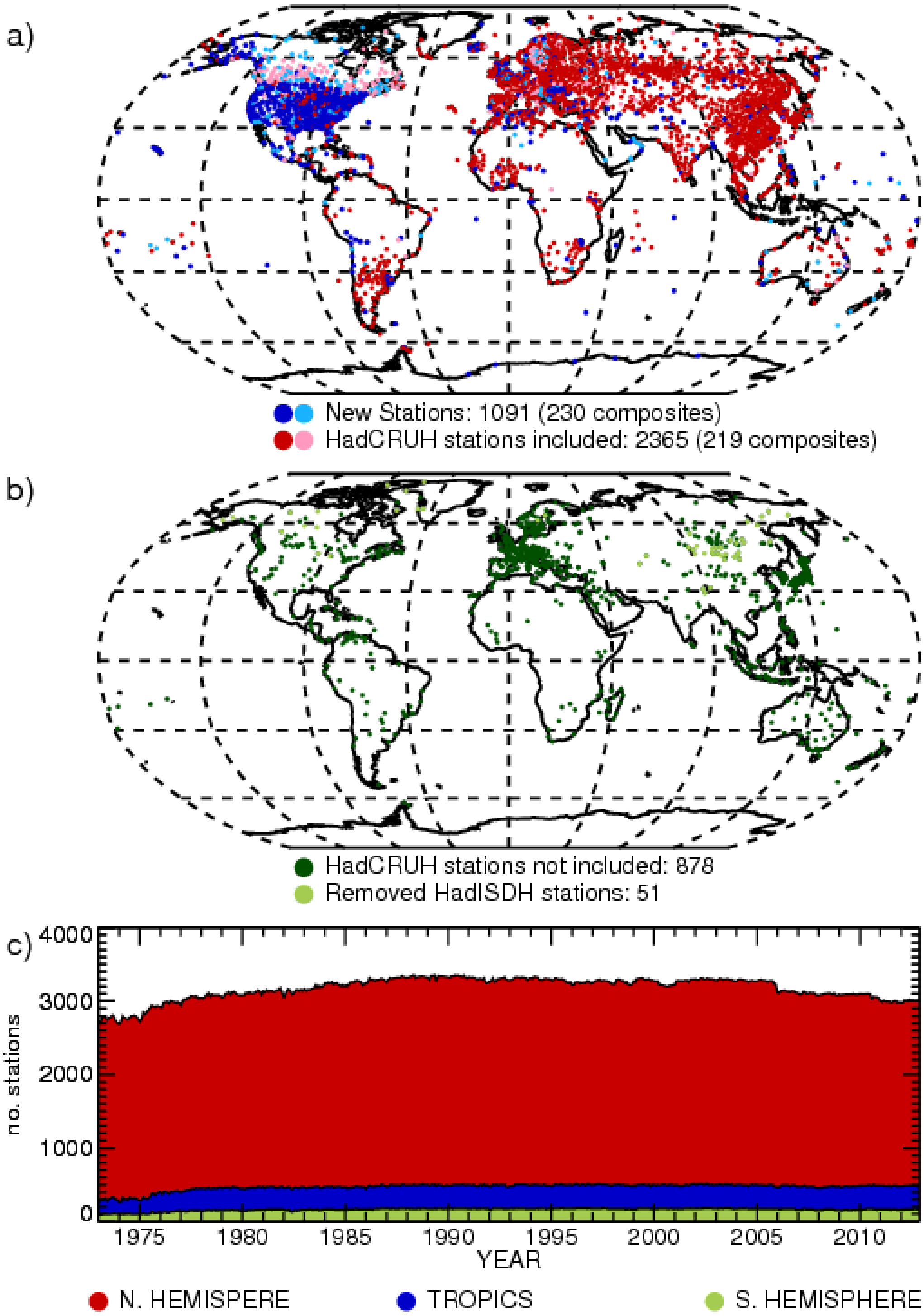

Figure 1. Station coverage comparison between HadCRUH and HadISDH. a) Station coverage in HadISDH. Stations in red/pink were also in HadCRUH. Stations in blue/turquoise are new. Pink and turquoise stations are stations that are composites of more than one original source station. b) Stations from HadCRUH 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.

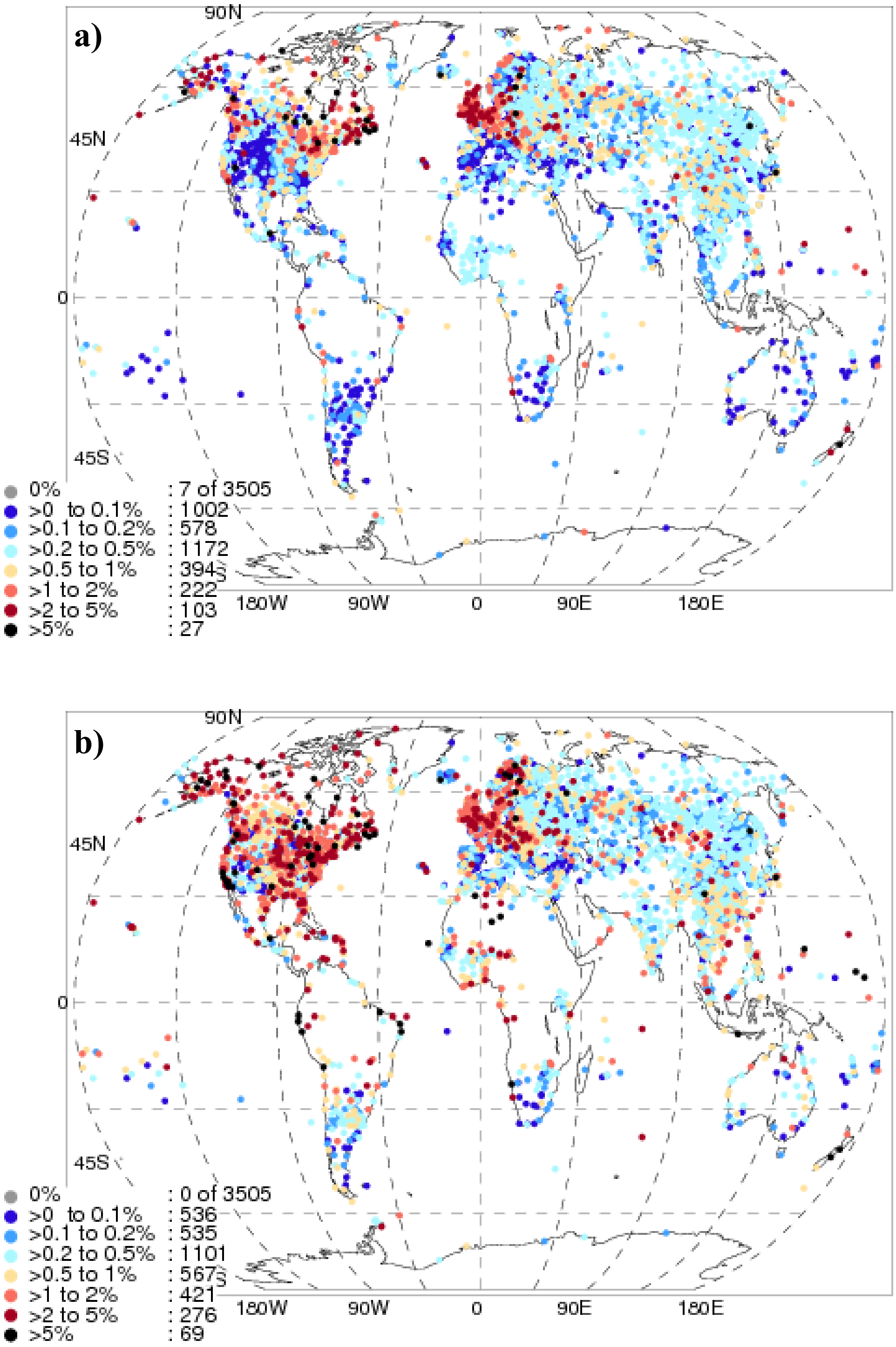

Figure 2. Percentage of hourly observations removed for each HadISDH station during the HadISD quality control procedure for a) temperature and b) dewpoint temperature.

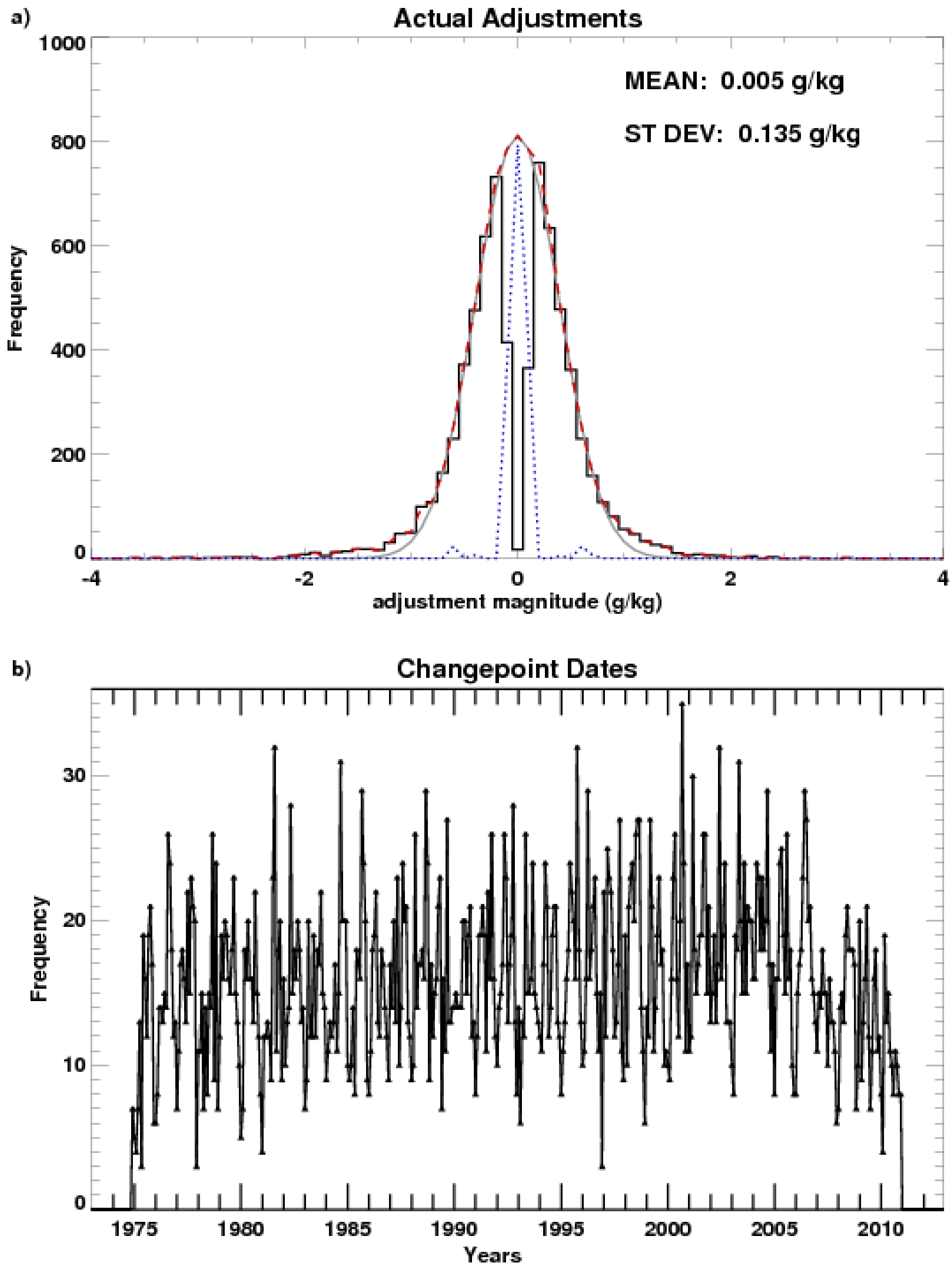

Figure 3. Summary of adjustments applied to HadISDH during the pairwise homogenisation process. Figure a) shows the actual adjustments in black (stepped). The best-fit Gaussian is shown in grey. The merged Gaussian plus larger actual distribution points ‘best-fit’ is shown in dashed red. The difference between the merged ‘best-fit’ and the actual adjustments is shown in dotted blue with the mean and standard deviation of the difference.

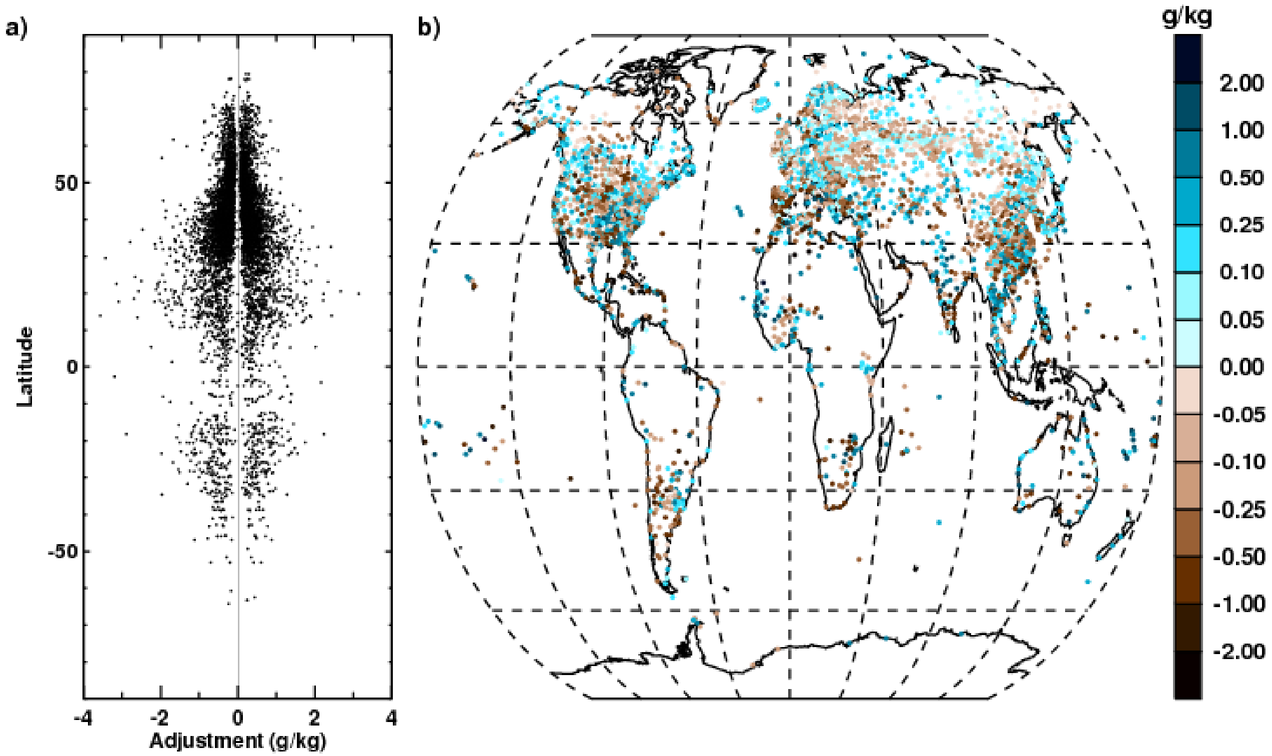

Figure 4. Distribution of adjustments made and their magnitude during the pairwise homogenisation process: a) adjustments by latitude; b) largest adjustments for each station. Note non-linear colour bars.

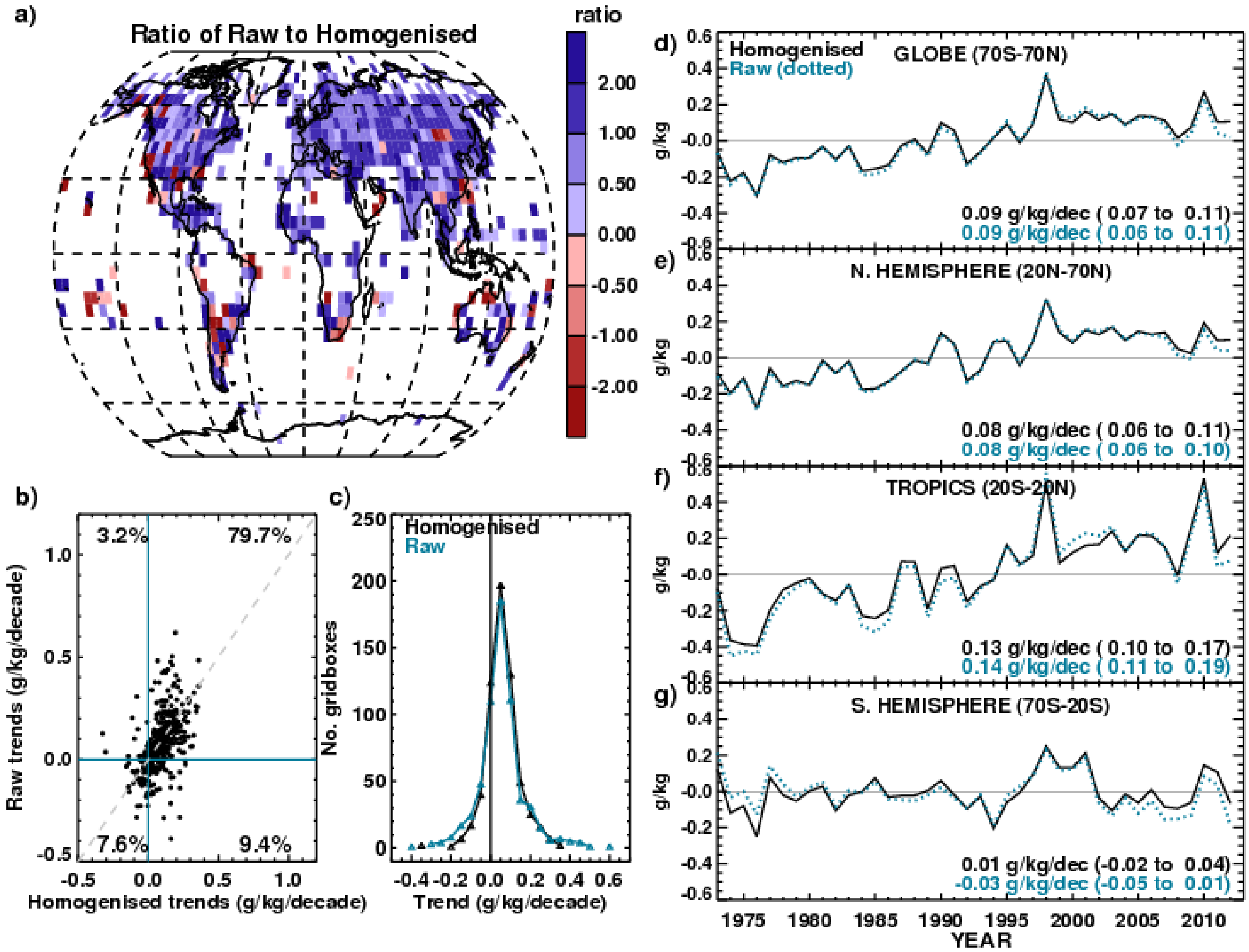

Figure 5: Difference between trends (1973-2012) in HadISDH before and after the pairwise homogenisation process. a) Ratio of decadal trends from the raw HadISDH compared to homogenised HadISDH (trend methodology is described in Figure 9). Note non-linear colour bars. b) Scatter relationship between homogenised and raw decadal trends for HadISDH. The percentage of gridboxes present in each quadrant is shown. c) Distribution of grid-box 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.

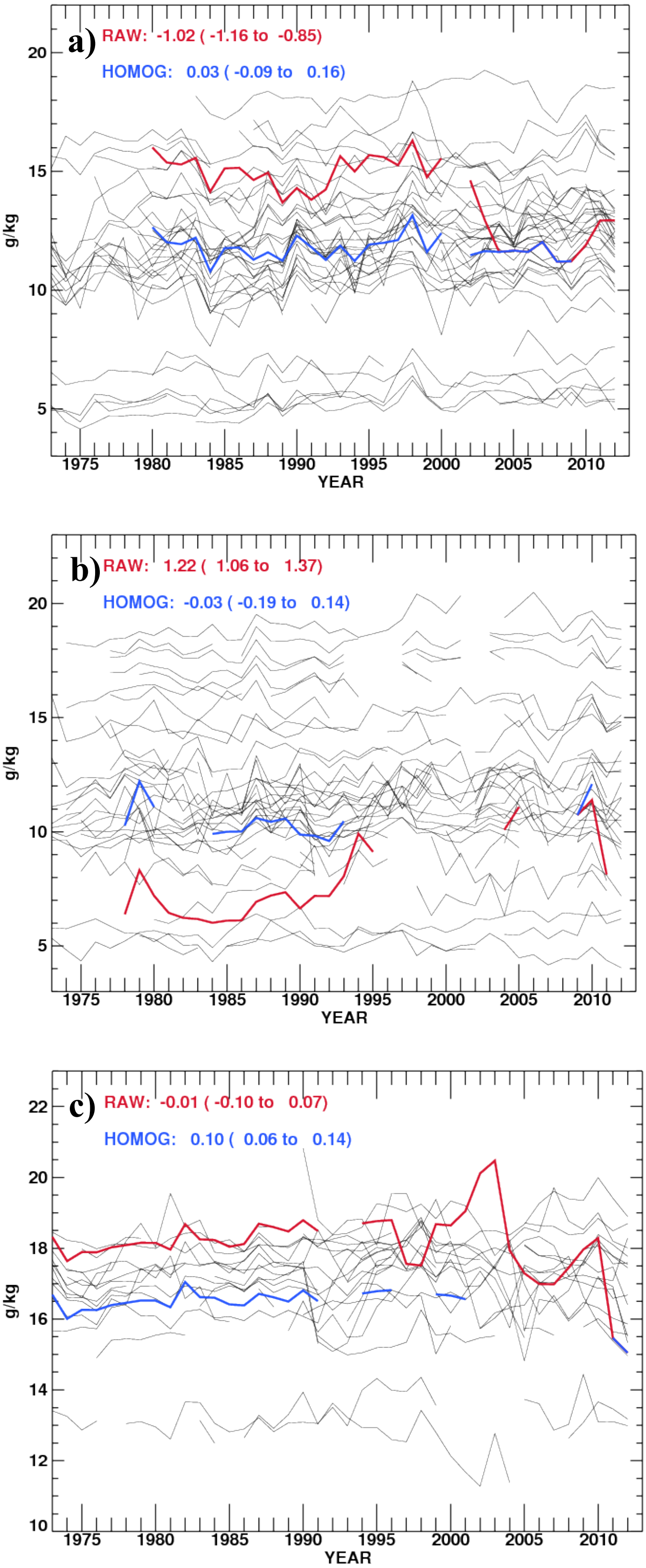

Figure 6. Results for three stations as examples of some of the largest changes of the pairwise homogenisation algorithm. Red lines represent the original station time series. Blue lines represent the adjusted time series. Black lines show the original time series for all stations within the designated network. a) Sur, Oman, WMO ID: 412680, 22.533°N, 59.467°E, 14.0m. b) Atar, Mauritania, WMO ID: 614210, 20.5170°N, 13.0670°W, 224.0m. c) Sao Luiz, Brazil, WMO ID: 822810, 2.6000°S, 44.2330°W, 53.0m.

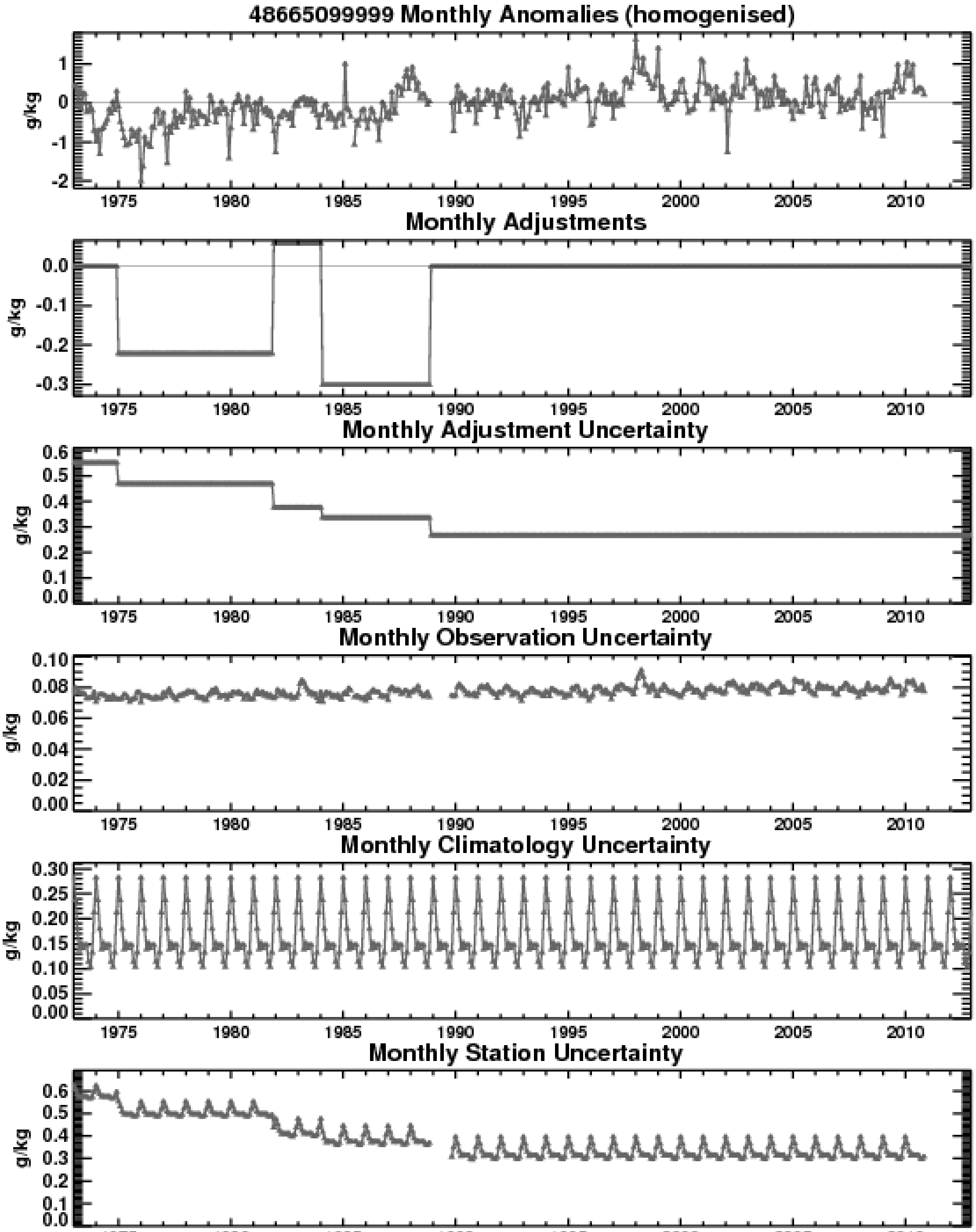

Figure 7. The components of station uncertainty estimates for station 486650 (Malacca, Malaysia, 2.267°N, 102.250°E, 9.0m). All uncertainties represent 2σ (approximately 95 % confidence intervals).

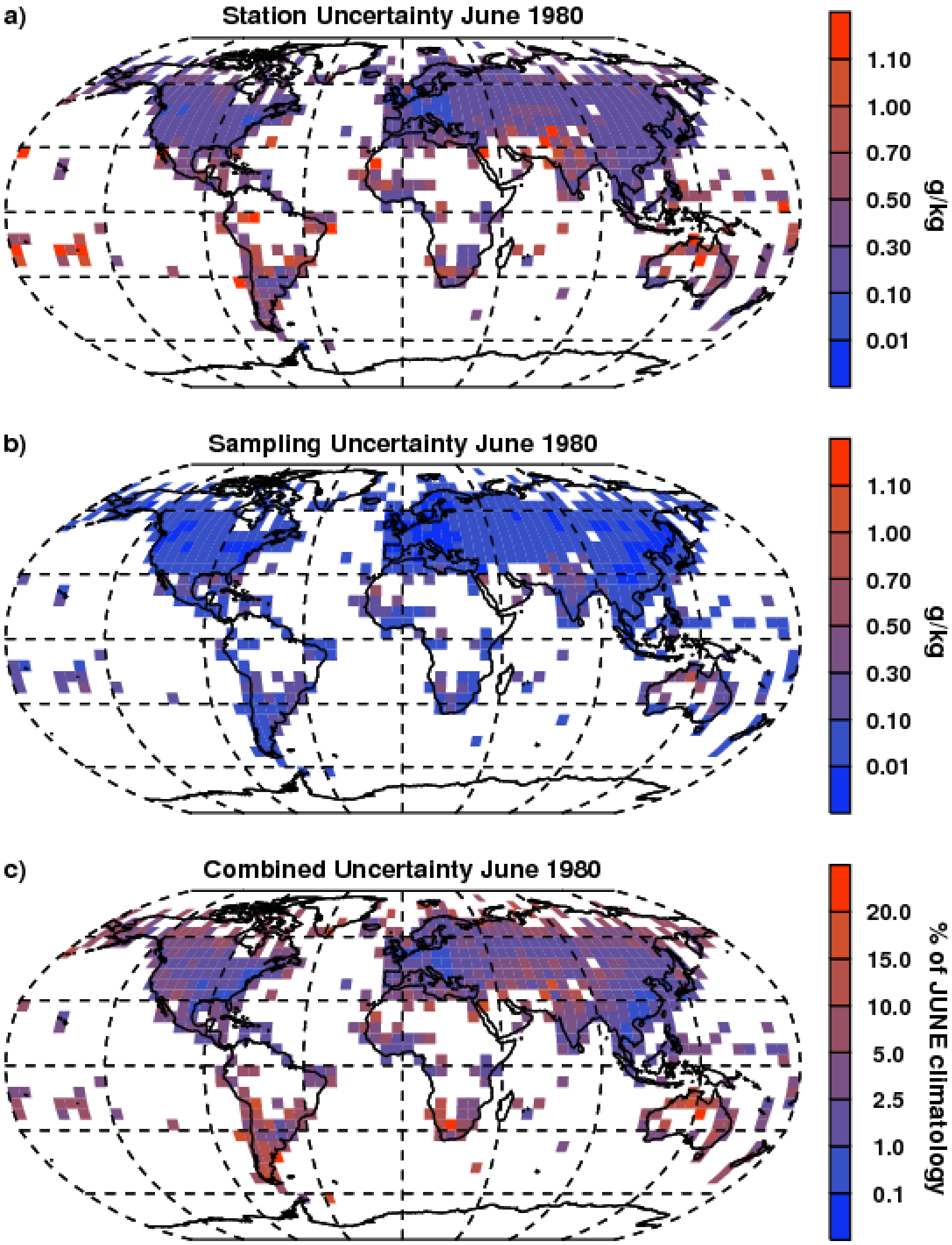

Figure 8: Gridded 2σ uncertainty fields for HadISDH. a) June 1980 gridded station uncertainty, b) sampling uncertainty, c) combined uncertainty for June 1980 as a percentage of the grid-box climatological (1976-2005) value for June. Note non-linear colour bars.

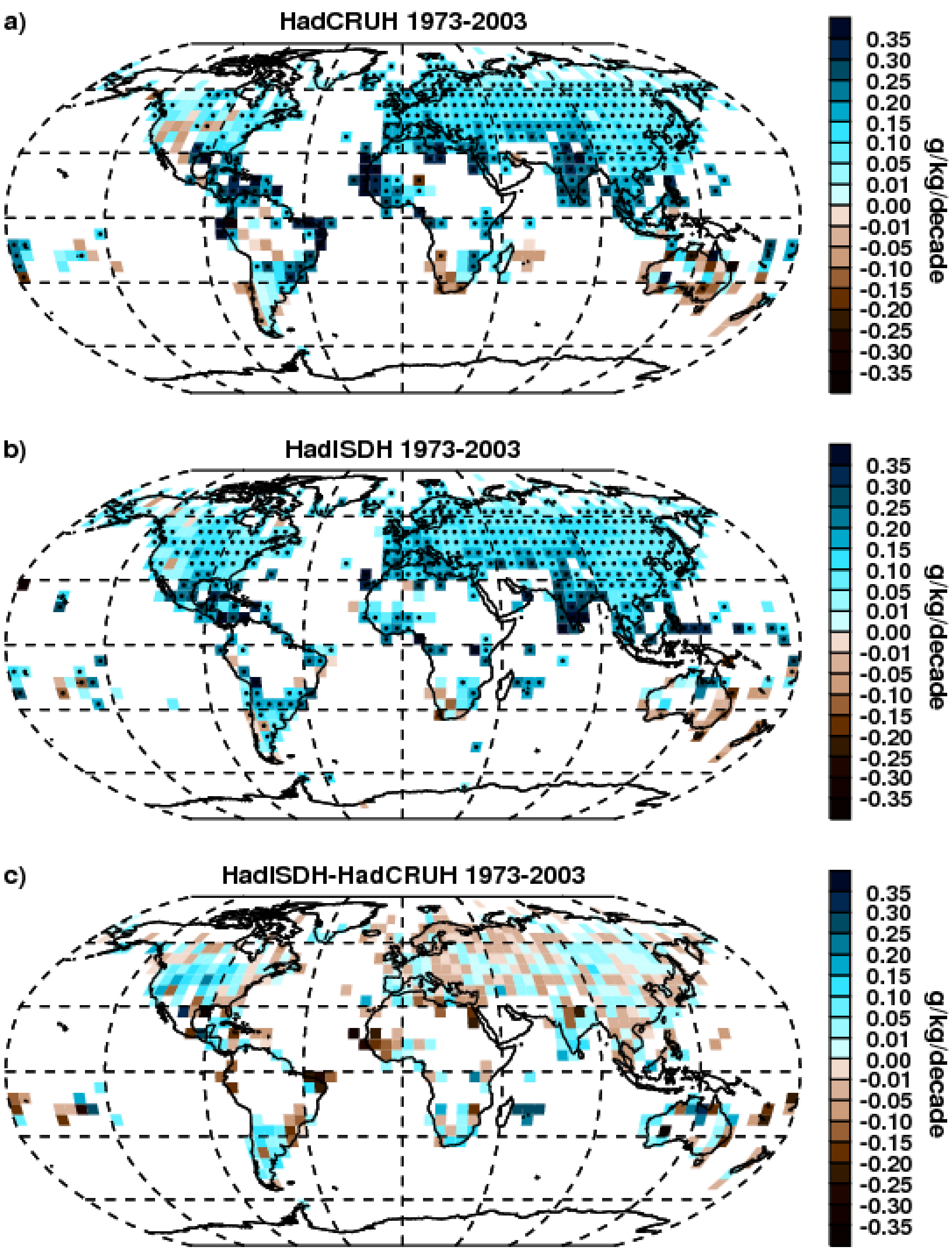

Figure 9: Decadal trends in specific humidity for HadCRUH verses HadISDH over the 1973-2003 period of record. Trends have been estimated using the median of pairwise slopes (Sen 1968; Lanzante 1996) method. Where intervals defined by the 95 % confidence limits on the median of the slopes are both of the same sign as the median trend presented in the gridboxes the trend is presumed to be significantly different from a zero trend. This is indicated by a black dot within the gridbox. This means that there is higher confidence in the direction of the trend, but not necessarily the magnitude. The spread of the confidence interval provides the confidence in the magnitude, these values are available online at www.metoffice.gov.uk/hadobs/hadisdh. Note non-linear colour bars.

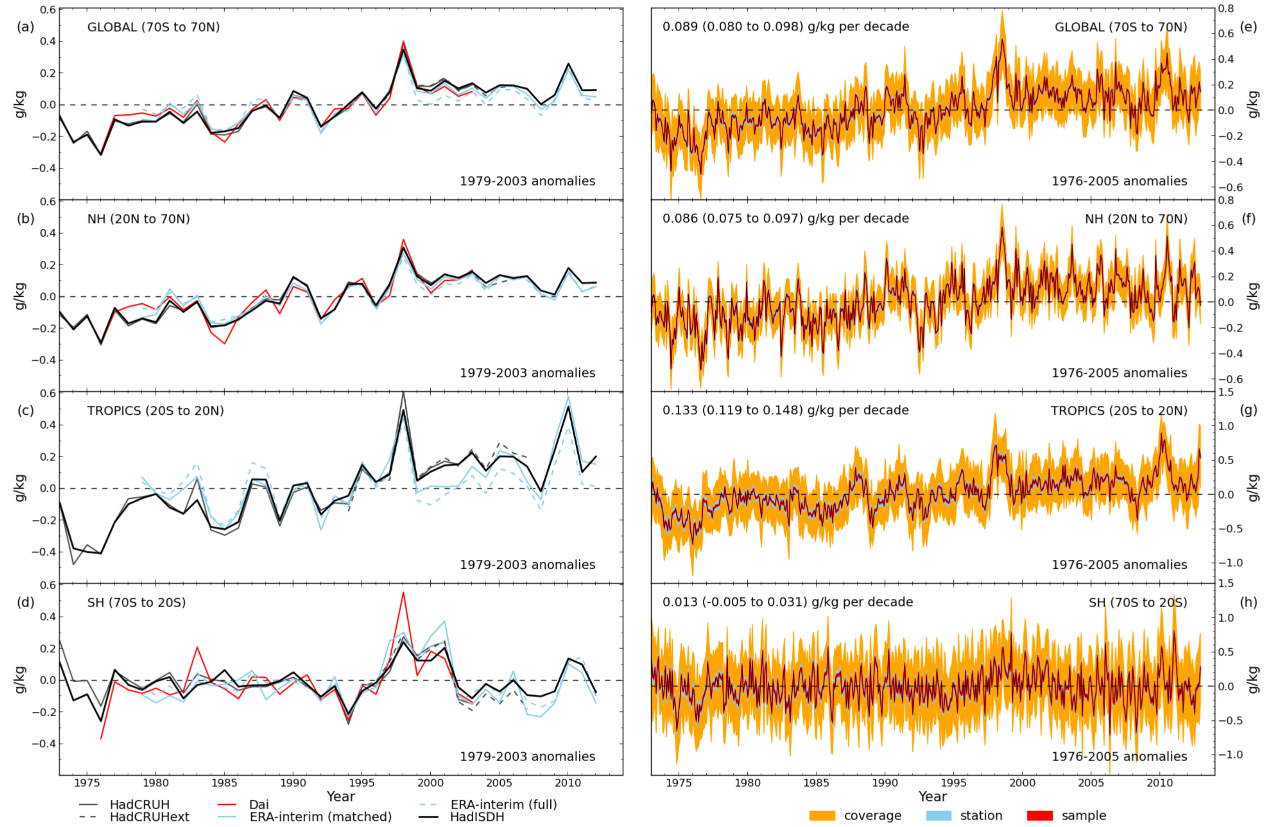

Figure 10: Time series of large-scale average specific humidity over land for HadISDH and existing data-products. a-d) Annual time series from all other global surface humidity products given a zero mean over the common period of 1979-2003. Dai covers 60°S to 70°N. ERA-Interim has been weighted by % 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 grid-box 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.

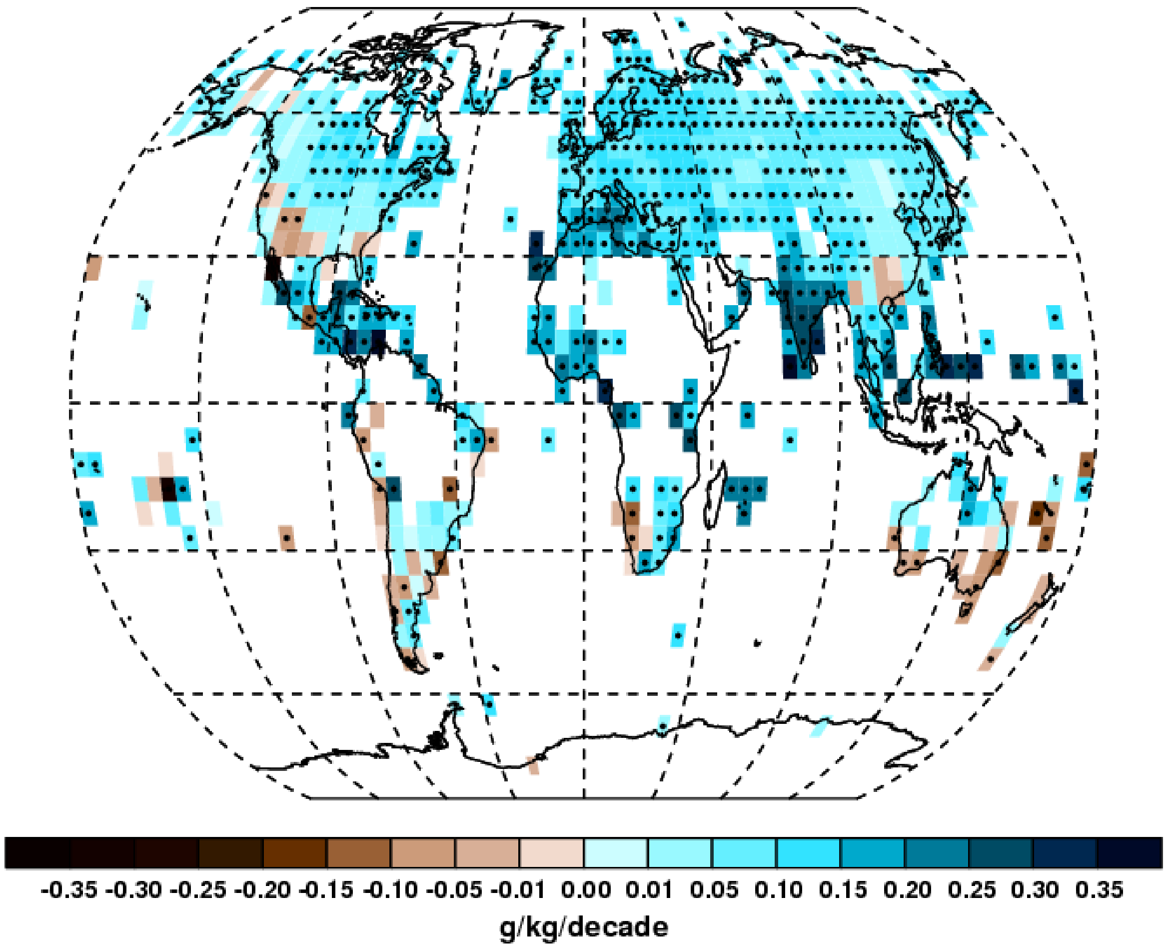

Figure 11: Decadal trends in specific humidity for HadISDH over 1973-2012. Trends are fitted and confidence assigned as described in Figure 9. Note non-linear colour bars.

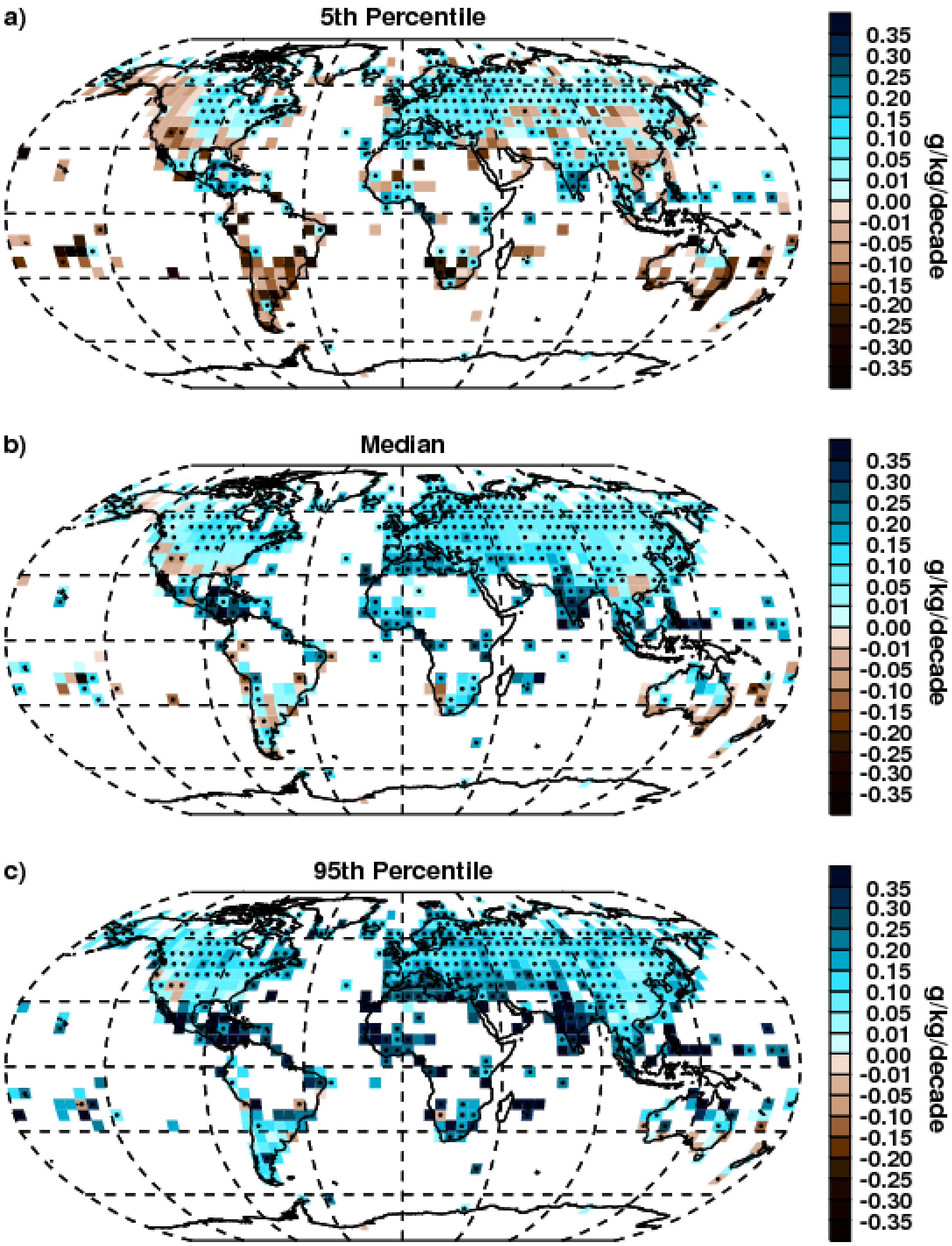

Figure 12. Exploration of the uncertainty in decadal trends using 100 realisations of HadISDH spread across the 2σ uncertainty estimates. Median pairwise trends were fitted over the period for each realisation, with higher confidence assigned by a black dot as described in Figure 9. For each gridbox, the 5th percentile (a), median (b) and 95th percentile (c) trends are shown. If the uncertainty was large enough to obscure the long-term trends then it would be expected that the 5th and 95th percentiles would starkly disagree with each other. In fact, there is very little difference as shown by a, b and c above. Note non-linear colour-bars.

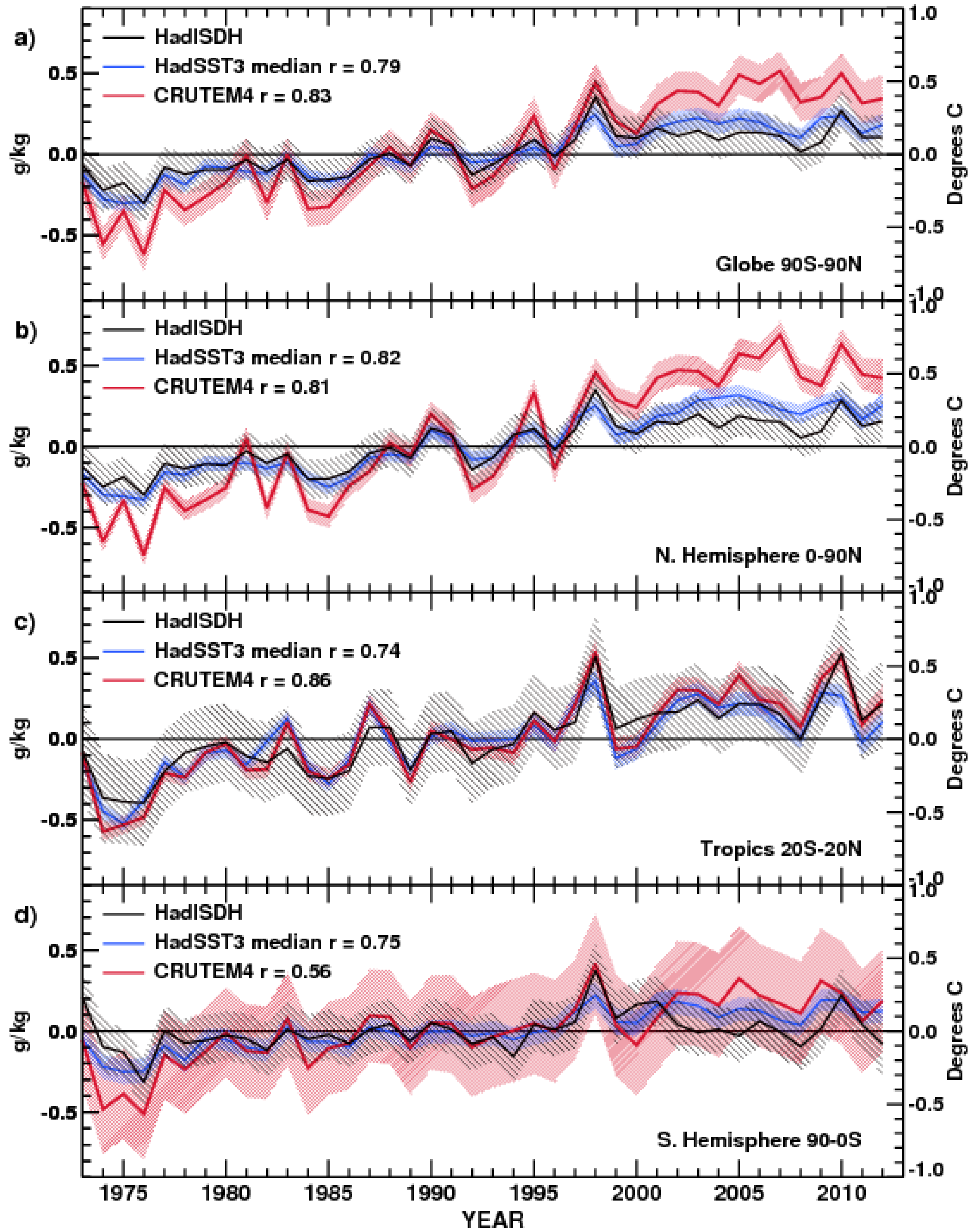

Figure 13. Comparison of large scale annual average time series from HadISDH land specific humidity with land surface air temperature from CRUTEM4 (Jones et al. 2012) and sea surface temperature from HadSST3 (Kennedy et al. 2011a, b) including uncertainty ranges. Temperature data have been adjusted to 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.

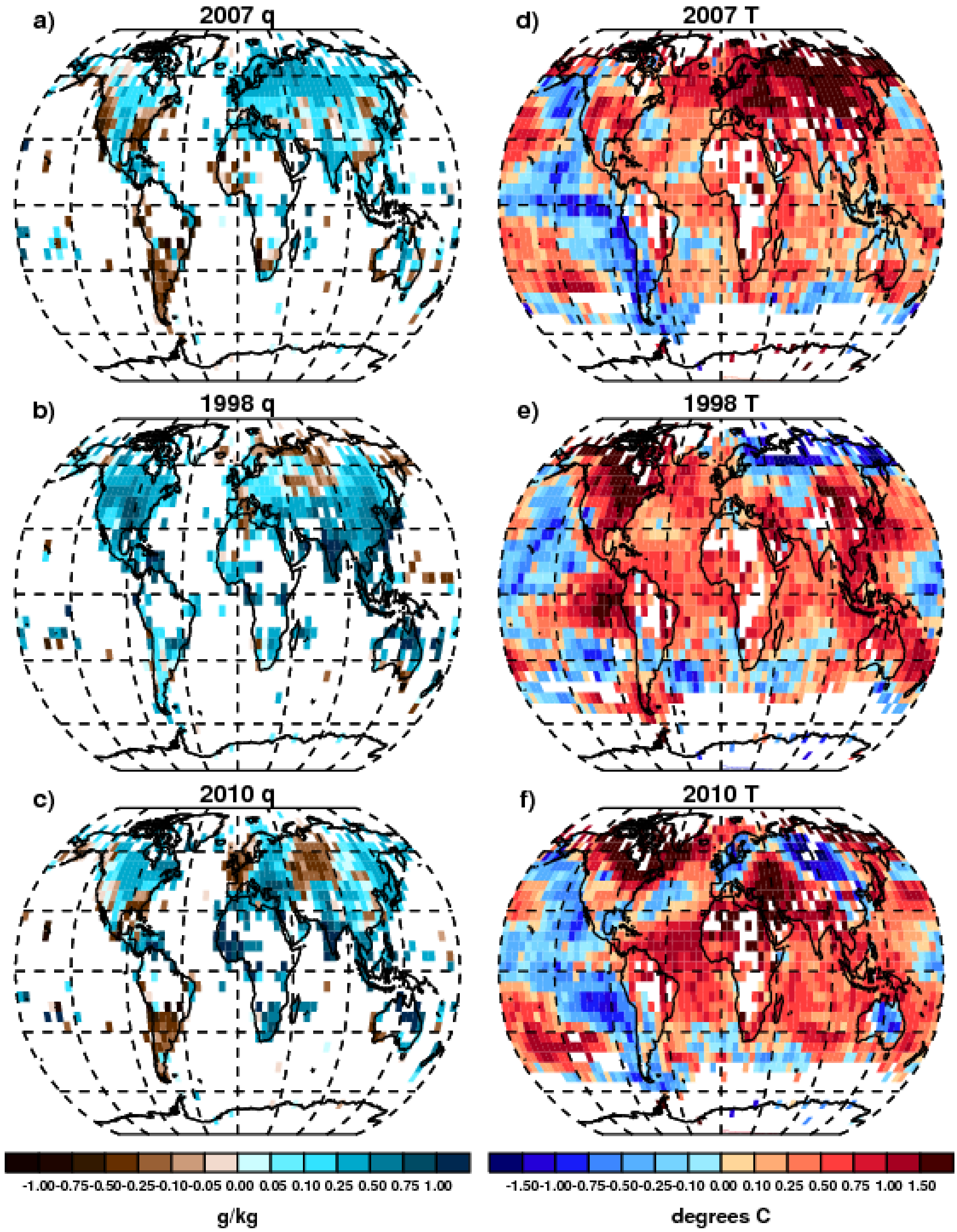

Figure 14. Annual average anomalies (from the 1976-2005 climatology period) for HadISDH specific humidity and HadCRUT4 (Morice et al. 2012) temperature, for the two moistest years within the HadISDH record (1998 and 2010) which were also among the warmest years since records began in 1850, and one of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