Chapter 4

The Marine Data

SUMMARY

The marine data component of HadCRUH comes from a well known and much used data source and for which a tried and tested system is already in place to convert raw data to a gridded product. The system required adaptation to work with humidity data. This Chapter discusses the data source, the implementation of humidity-specific quality control tests, homogeneity checking of the data and analysis of the gridded product including some issues with data quality pre-1982.

4.1 THE MARINE DATA SOURCE

The marine observations used for 1973 to 1997 are taken from ICOADS (International Comprehensive Ocean-Atmosphere Data Set) release 2.1 (Worley *et al.*, 2005). From 1998 onwards, the observations used are provided through the Global Telecommunications System (GTS), and made available by the NOAA National Centers for Environmental Prediction (NCEP) (http://icoads.noaa.gov/ncep_obs). This combination of data sources is chosen: for comprehensiveness; because of regular updates; and because it has been used for numerous Hadley Centre marine datasets (e.g. interpolated global sea ice and *SST* dataset – HadISST), uninterpolated global *SST* dataset - HadSST2, uninterpolated global night *MAT* dataset - HadMAT1 (Rayner *et al.*, 2003, 2006; www.hadobs.org)).

Marine observations come from a range of platform types (Fig. 4.1). Historically ships have been the primary source, however, from the mid-1990s ship observations have been in decline while buoy observations have increased. Furthermore, ship height, instrument type and exposure and observing practices also vary widely through time and over space. These differences have been found to cause biases and uncertainty in several meteorological variables (*SST* – Rayner *et al.*, 2006; *MAT* – Rayner *et al.*, 2003; Berry & Kent, 2005). Humidity measurements will not be immune from such problems and previous attempts have been made to apply adjustments (Table 4.1). Unlike other variables, errors in q were found to be smaller at high latitudes and largest in the Tropics (Kent *et al.*, 1999). However, in most cases, adjustments are applied to specific ships (height corrections - Kent *et al.*, 1999) or instrument types (Screen as opposed to Psychrometer measured humidities - Kent *et al.*, 1993). Although digitised metadata are available from 1973 to undertake such an analysis within this project is considered too great a task. Furthermore, it is of interest to create an un-adjusted dataset as adjustments may introduce errors, referred to as structural uncertainty (Thorne *et al.*, 2005a).

In both ICOADS and the GTS data, humidity observations are reported as T_{dw} and sometimes accompanied by T_w. Observations are hourly or less frequent. The accuracy is to one decimal place although sometimes rounded to the nearest half- or whole-degree. These variables along with T are converted to e, RH and q as described in Chapter 2. Unlike *MAT* datasets (Rayner *et al.*, 2003) there is no need to exclude daytime observations because the solar heating bias has been found not to significantly affect the humidity measurement (Kent & Taylor, 1996) (Table 4.1).

For consistency with the land data, the period of record is from 1973 to 2003 with a 1974 to 2003 climatology (section 3.1.2). Sampling density remains fairly constant at ~1.5 x 10^6 observations yr^{-1} globally (Fig. 4.1). Spatially, the marine component incorporates all oceans of the world, the Great Lakes and the Caspian Sea. Coverage is good over the Northern Hemisphere mid latitudes and common ship tracks (Fig. 4.2). The high latitudes and Southern Hemisphere mid-latitudes are a particular problem with very low reporting densities. In total only 47 % of the ICOADS/GTS database contains humidity data (where both T and T_{dw} are present).

4.2 CREATING THE MARINE DATASET

4.2.1 The Marine Data System for Humidity

The existing Hadley Centre Marine Data System (MDS) is a set of programs to undertake: raw data extraction; quality control; homogeneity checks; and gridding. It has been well used and documented for: *SST* (HadISST – Rayner *et al.*, 2003; HadSST2 – Rayner *et al.*, 2006), *MAT* (MOHMAT43N - Parker *et al.*, 1995; Rayner *et al.*, 2003) and sea level pressure (*SLP*) (HadSLP2 – Allan & Ansell, 2006). The system, tuned specifically to humidity, is described below and summarised in Fig. 4.3. It is separated into two stages.

4.2.2 First Stage Quality Control

All tests included in the first stage quality control are described in Table 4.2. Collectively, the tests remove 15.7 % of the data. The Base QC is part of the original MDS. It ensures the correct positioning of each observation in space and time. In total, 3.6 % of humidity data are removed with higher percentage removals in the high northern latitudes, the north east coast of Asia and inland Seas (Fig. 4.4).

The Dewpoint Temperature QC leads to a small data removal (0.3 %) with a small cluster of higher percentage removals around the west coast of the Antarctic Peninsula (Fig. 4.4). Although in reality T_{dw} can be slightly greater than the simultaneously measured *T*, especially in very cold humid environments, instruments used for weather observation are considered highly unlikely to be of the standard to measure this accurately (Makkonen & Laakso, 2005).

The Relative Humidity QC is based on the premise that the ocean is an infinite moisture source for the near-surface atmosphere. Hence, as a general check on quality it is assumed that *RH* should never really drop below 60 %. An offshore wind could bring drier air across the sensor thus humidity in coastal regions could be lower but it was assumed that this should not be less than 40 %. This test removes 0.7 % of the data, proportionally more around coastlines and inland lakes/seas (Fig. 4.4). The Caspian Sea and north Chilean coastline are two key examples. This implies that such low humidities are plausible in these areas and that the test needs to be modified in future versions to allow for lower humidities around coastlines and inland water bodies.

The Temperature QC is included because *RH* values are derived directly from *T*. A ‘bad’ *T* value is also considered indicative of poor quality of the simultaneously measured variables (related or not). Thus observations are removed if the simultaneous *T* value: is missing; differs by more than 10 ^{o}C from the climatology (Outlier QC –

section 4.2.4) (the test uses a climatology from HadMAT1 – Rayner *et al.*, 2003), or fails the Neighbour Check QC (section 4.2.5). This test removes 11.3 % (more than any other test) of the data. High latitudes and the inland Seas / Great Lakes experience the highest percentage removals (Fig. 4.4).

The daytime heating bias in *MAT* has been well documented (Parker *et al.*, 1995; Rayner *et al.*, 2003), although it is not thought to affect humidity conversions (Kent & Taylor, 1996). Proportionally more daytime than nighttime data are removed by the first stage quality control (Fig. 4.4) suggesting that daytime data are more problematic than nighttime. Unlike the land data, no measures have been taken to ensure balanced sampling across the diurnal cycle. Previous humidity studies generally concur on trends in *q* and *RH* being larger during the nighttime (Dai, 2006; Gaffen & Ross, 1999; Wang & Gaffen, 2001) although others find no difference (*RH* – Van Wijngaarden & Vincent, 2005). As such, it is possible that uneven sampling across the diurnal cycle may introduce a slight bias in the marine data relative to the land. However, the final proportions of nighttime and daytime data have not been calculated and so no inferences can be given for this first version of HadCRUH.

4.2.3 Gridding the Marine Data

The marine data are gridded by first creating 1 ^{o} by 1 ^{o} grid-box means at pentad (Box 3.1) resolution. Pentad mean anomalies are then created by subtracting the climatology (1974 to 2003 pentad mean for each 1 ^{o} by 1 ^{o} grid-box). These are then averaged by winsorising to produce monthly mean anomalies at the 5 ^{o} by 5 ^{o} grid-box resolution. Winsorising, a technique common to Hadley Centre marine datasets, reduces the effects of extreme values and is described in Rayner *et al*. (2006). Three gridded fields are output: anomalies; number of observations; and standard deviations.

As climatological fields based on the observations cannot be created until the data have been gridded, NCEP Reanalyses are used to create first version climatologies. For further iterations within the dataset building process and the end product it is desirable to have a climatology created from the marine data itself for a number of reasons. Firstly, the NCEP Reanalyses are only available in *q* and *RH*. Although conversions are made to create *e* fields, for accuracy, humidity conversions should be done for each observation rather than monthly mean gridded fields (McCarthy & Willett, 2006).

Secondly, the spatial (grid-box) and temporal (monthly mean) resolution of the Reanalyses are not sufficient to reflect the standard deviation actually found in an hourly dataset at point of source. Finally, the Reanalyses climatology fields, although mostly within 1 g kg^{-1} of the observational climatologies, are consistently dry biased over the Tropics and mid-latitude Northern Hemisphere and consistently moist biased over the Southern Hemisphere (below the Tropics) and in the high latitude Northern Hemisphere (Fig. 4.5). Generally, these biases are stronger in the respective hemispheric Summer. Regions such as the Persian Gulf and the Red Sea are strongly underestimated (>5 g kg^{-1}) relative to the observations. This bias is consistently dry throughout the seasonal cycle but stronger from April to September.

New monthly climatology and standard deviation fields are created each time the observations are gridded. This is done by adding the latest climatology fields (re-gridded to monthly 5 ^{o} by 5 ^{o} resolution) back to the now gridded anomaly fields to give absolute values. Notably, this means that by using Reanalyses for the initial climatologies they are in some small way part of the HadCRUH dataset. However, Reanalyses data are at no point incorporated into HadCRUH as actual data and are very unlikely to have any real effect on the final dataset timeseries and trends for the following reasons. Firstly, as climatological values, Reanalyses will have a consistent effect on the anomalies over time thus not affect trends. Secondly, any persistent bias in the Reanalyses climatologies relative to the observational data should be lost when the gridded anomalies are initially added back to the Reanalyses climatology to create new absolute values. Thirdly, the quality control, gridding and recreation of climatologies are iterated several times further reducing any influence of the Reanalyses.

For some data sparse grid-boxes it was not possible to calculate climatologies or standard deviations from the actual data. To preserve the little data that remains in these boxes, gaps in the observation based climatologies are filled in with Reanalyses climatologies. Gaps in the standard deviation field are filled in with an arbitrary large value of 10 000. Both fields are then re-gridded to 1 ^{o} by 1 ^{o} at pentad resolution (Box 3.2) as this is necessary for quality control and gridding. For the climatology, pentads are interpolated from a non-linear fit of the annual cycle as for HadSST2 (Rayner et al., 2006). For the standard deviation fields, each month is divided into the appropriate number of pentads with values identical throughout the month.

4.2.4 Second Stage Quality Control and Subsequent Iterations

The derived climatology and standard deviation fields facilitate two further checks on the data: an outlier check (Outlier QC) and a homogeneity check (Neighbour Check QC). Following application of these, the remaining data are then re-gridded and a new climatology and standard deviation field created. This process, including creating new climatologies and standard deviation fields can then be repeated several times. At least two iterations are deemed necessary in order to have a climatology created without the first sweep of failed Outlier and Neighbour Check QC values. Following this, some subjective criterion is required to avoid an infinite loop and removal of all data. A judgement is made to end iterations when total data removal falls below an additional 0.5 % as this is a small amount of data and hence unlikely to have a large effect on the end product. There is also a large computational overhead needed to run further iterations.

The Outlier QC is part of the original MDS but has been modified to use four standard deviations rather than an absolute value (as with *SST*) as a maximum difference from the climatology to take into account the very different humidities across the world. An observation (all variables) is removed if any one of the three variables (*e*, *q* and *RH*) fail (Table 4.3). In total (as a percentage of the data remaining from each iteration), 0.09 % and 0.12 % of the data from iterations one and two are removed respectively. This increase is counter-intuitive but quantities and tests have been double checked and found to be without error. It is likely to be occurring due to the narrowing of the standard deviation with each iteration. Data removals for *e* and *q* are almost identical for the Outlier QC as might be expected considering their similarity quantitatively (McCarthy & Willett, 2006) and that *q* is derived from *e*. As *q* removals tend to be very slightly larger, future versions of HadCRUH could use only *q* and *RH* removals to save processing time.

Data removal for *q* is greater around the Tropics, coastal and data sparse regions (high latitudes and Southern Hemisphere above 45 ^{o}S). Data removal for *RH* is generally lower and more uniform across the oceans but removals are similarly higher around coastlines, in the high latitudes and other data sparse regions (Fig. 4.6). It may be that area specific thresholds of standard deviations are needed around coastlines, inland seas and some other regions.

The Neighbour Check QC is a simple 'keep or remove' homogeneity test at the individual observation level without any attempts to identify breakpoints or adjust the data. It compares each monthly mean anomaly of the observation variable with a neighbour composite monthly mean anomaly. Table 4.4 lists the set of criteria for a neighbour composite in terms of distance and time and the corresponding standard deviations within which the candidate value must lie (Rayner *et al.*, 2006). Only those values passing the Outlier QC are put through the Neighbour Check QC.

Much of the *RH* data are flagged for removal (Table 4.3) suggesting that the temporal and spatial boundaries of the neighbour checks are too large for use with *RH*. Consequently, Neighbour Check QC removals for *RH* are ignored. As with the Outlier QC, removals for e and q are sufficiently similar to just use q as an indicator for removal. In total (as a percentage of the data remaining in each iteration), 0.43 % and 0.32 % of the data from iterations one and two are removed respectively.

This is a conservative removal compared to that of *SST* which is ~4.5 % in a test run simultaneously for comparison. Data removal is largest in Northern Hemisphere coastal regions (Fig. 4.7) suggesting that the proximity of land may affect the spatial and temporal continuity of q. Data sparse regions also show relatively large data removal which may be due to poorer data quality in these regions or a result of having fewer observations.

Total data removal during the second stage quality control is 0.6 % and 0.4 % for iterations one and two respectively. Outlier and Neighbour Check QC tests are important but much of the very poor data has already been removed by the previous quality control tests. Removal by iteration two is less than 0.5 % and so there are no further iterations. Final climatologies of the resulting grids are made and used to re-grid the remaining raw data completing the marine component of HadCRUH.

A summary of all data removal can be found in Table 4.3. Total data removal is fairly consistent through time (Fig. 4.8). There is a clear annual cycle in data density, largely due to the seasonality of shipping routes. The fall in data density in 1998-1999, the first year of GTS records, may also affect the long-term homogeneity of the data set.

The geographical distribution of annual and seasonal average observation density in the final grids does not differ distinctly from the original distribution, showing no significant overall geographic bias in data removal (Figs. 4.2 and 4.9). Data coverage is poor in the Southern Hemisphere and in the high latitudes during winter. Due to the latter, only 70 ^{o}N to 70 ^{o}S are considered useful for regionally averaged analyses in HadCRUH, at least until coverage improves in higher latitudes.

4.3 AN ANALYSIS OF MARINE SURFACE HUMIDITY

4.3.1 The Climatology of Marine Surface Humidity

As mentioned in section 3.1.1, only q and RH are analysed, as with the land data, because these are considered of most interest to the scientific community. Similar to the land data, climatological marine q has a strong zonal structure peaking in the Tropics and decreasing towards the Poles (Fig. 4.10). It ranges from ~1 to ~20 g kg^{-1}. Climatologically, RH is found to vary very little over the oceans (Fig. 4.11), largely between 70 and 90 %. It is highest in the high latitudes, especially in the Summer hemisphere. In regions of poorer data coverage (high latitudes and Southern Hemisphere (45 ^{o}S+) values can go as low as 40 % and above 95 %.

4.3.2 Recent Changes in Marine Surface Humidity

The data are regionally averaged (section 3.4.2) at the monthly mean anomaly resolution for four regions (Global: 70 ^{o}N - 70 ^{o}S, Northern Hemisphere: 70 ^{o}N - 20 ^{o}N, Southern Hemisphere: 20 ^{o}S - 70 ^{o}S and Tropics: 20 ^{o}N - 20 ^{o}S). The timeseries for q (Fig. 4.12) exhibits distinct peaks in the Tropics for 1983 (less strong), 1987 and 1998 in common with the dates of some but not all large ENSO events and notably the same events as shown in the land data (*cf.* Fig. 3.32). Only a 1987 feature is apparent in the RH timeseries, but present in all zones (Fig. 4.13). This is positive in the marine data but negative in the land data.

For q, all regionally averaged trends except that of the Southern Hemisphere are positive and highly significant (at the 1 % level) (Global value = 0.07 g kg^{-1} 10yr^{-1}), and of greatest magnitude in the Tropics (0.10 g kg^{-1}). Trend sign is comparable to that over land (section 3.4.2), but the trend magnitude is generally somewhat smaller. Considering the Clausius-Clapeyron relation (section 1.2) under the assumption of near-constant *RH*, where the largest changes in absolute humidity are likely to be in regions of higher ambient *T* and unlimited water availability (such as over the oceans), these findings are unexpected if changes in *T* are similar over both land and ocean

Regionally averaged trends in *RH* are negative and highly significant (at the 1 % level) for all regions except the Tropics with a Global trend of -0.10 % $10yr^{-1}$. Importantly, this contradicts the assumption of constant *RH* over time (Allen & Ingram, 2002). In comparison with the land data, the Global trend is larger (land = -0.03 % $10yr^{-1}$) but the Southern Hemisphere trend is much smaller (marine = -0.11 % $10yr^{-1}$ *cf.* land = -0.34 % $10yr^{-1}$).

Of note in both variables, but especially *RH*, is a downwards shift in the mean after 1982. For *RH* it is apparent in all regions but strongest in the Southern Hemisphere. For *q* it is also strongest in the Southern Hemisphere and visible in the Northern Hemisphere. The magnitude of this shift is considerable, where for *RH*, the Global timeseries mean changes from 0.31 % to -0.11 % for the periods before and after the shift respectively. Consequently, if this is not a real climate phenomenon and actually due to problems with the data, *RH* trends could be much closer to zero and *q* trends more positive, which as discussed earlier would be closer to that expected based on Clausius-Clapeyron. This is discussed further in section 4.4.

Consistent with the land, all seasonal *q* trends (Table 4.5) are positive and significant (at a mix of 1 % and 5 % levels) except for the Southern Hemisphere, and largest in the Tropics and Northern Hemisphere Summer. Seasonal trends in *RH* (Table 4.5) are mostly negative and not significant but with positive trends in the Northern Hemisphere JJA and SON, and less consistent with the land data. The largest trends occur in the Northern Hemisphere Summer (positive and highly significant) (compared to Winter over land) and Southern Hemisphere Winter (negative and significant). Thus, there are some trends of significance at this smaller spatial and temporal scale but generally seasonal changes in marine *RH* are not significant.

At the grid-box scale, *q* trends (Fig. 4.14) are mostly positive, (76 % of grid-boxes) similar to the land data, with a range of -0.32 to 0.47 g kg^{-1} $10yr^{-1}$ (less positive than for the land data) and -2.15 (in MAM) to 1.93 (in SON) g kg^{-1} $10yr^{-1}$ for seasonally averaged trends. Negative trends are clustered mostly in the Southern Hemisphere where data coverage is poorer. Trends have widespread continuity both zonally and meridionally. Trends in *RH* (Fig. 4.15) are much more mixed and generally more negative (59 %) and range from -2.50 to 2.50 % $10yr^{-1}$ and -20.05 (in JJA) to 15.24 (in MAM) % $10yr^{-1}$ for seasonally averaged trends. The seasonal minima and maxima are very large. However, on further analysis at least 95 % of all grid-boxes with data present have trends within -5 to 5 % $10yr^{-1}$ for each season and the extremes are confined to high latitudes or the Southern Hemisphere above 45 ^{o} where data quality is poorer. *RH* is less zonally and meridionally consistent but has regional continuity. Drying in the south Pacific is common to both *q* and *RH*.

Trends in *q* are not as seasonally dependent as those for the land data (Fig. 4.14, *cf.* Fig. 3.34). This could be explained by the smaller annual temperature cycle amplitude over the oceans relative to the land. There is still some tendency towards moistening in the Summer season in the low to mid-latitudes at least for the Northern Hemisphere. The mid and high latitude Southern Hemisphere is very patchy in sign and magnitude with little spatial continuity implying poorer data quality there. Trends are consistently positive in the Tropics but consistently negative around New Zealand and over the north Pacific. The patch of drying in the north Pacific is notably large in spatial scale and magnitude in the DJF season.

For *RH*, seasonality is difficult to detect (Fig. 4.15). The large drying in the North Pacific in the DJF season matches that for *q*. As with *q*, the Southern Hemisphere mid to high latitudes appear to be of poor quality in terms of spatially incoherent trends. The poorer zonal continuity relative to *q* is evident in these seasonal analyses.

4.4 AN INVESTIGATION INTO THE 1982 HUMIDITY SHIFT

There appears to be a shift in the mean of the marine humidity data in 1982. No such shift is apparent in *SST* and *MAT* timeseries from HadSST2 and HadMAT1 (Rayner *et al.*, 2003; Rayner *et al.*, 2006) or the HadCRUH land data (*cf.* Figs. 3.32 and 3.33), isolating this as a marine humidity issue.

This is further supported by looking at land minus marine humidity for coastal grid-boxes (containing both land and marine data) averaged for the Globe (Fig. 4.16). For q there is a persistent moist bias of the marine data relative to the land pre-1982 and good agreement around zero there after. Interestingly, in 1998 there is a noticeable divergence from zero but a later return. This is the year in which: the marine humidity source is switched from ICOADS to GTS data; when the number of observations from buoys increases significantly; and also of a strong El Niño event which may be to some extent responsible. Notably, the impact of introducing large numbers of buoy observations is likely less than for temperature because only stationary buoys report humidity and as such any effects should be reasonably easy to detect spatially. Such a shift is not apparent in the regionally averaged marine timeseries however. Although the 1982 shift is clearest in the RH regionally averaged timeseries, it is less clear (although still present) in this coastal analysis. Indeed fitting a trend to the coastal difference series over the period gives a small and non-significant RH trend but a q trend comparable with that of the marine Global mean trend and highly significant (0.05 g kg^{-1} $10yr^{-1}$). If it is assumed that this shift is entirely non-climatic in origin, then this trend can be applied as a first-order correction to the marine humidity data resulting in an adjusted global marine trend of 0.12 g kg^{-1} $10yr^{-1}$ which is now slightly larger than for the land.

To look at the spatial distribution of the shift, t-tests are run on timeseries from four year time periods before and after 1982 for each grid-box to see if there are any regional patterns (Fig. 4.17). Significant differences are widespread, but with clear regions where the shift is not apparent and these regions differ seasonally in location.

After much discussion with Hadley Centre personnel and scrutiny of the ICOADS literature (Kennedy, J. *pers. comm.*., http://icoads.noaa.gov/e-doc/lmr), it was discovered that only whole number values of T_{dw} were reported to the GTS until 1st January 1982 due to space limitations. For T, recording decimal places was optional. For some countries such practices still continue. Indeed, by plotting the frequency of each decimal place for T and T_{dw} data for an equal period of time (nine years) before and after 1982 it is clear that this is the case for T_{dw} and not T giving strong support for a non-climatic origin of the shift (Fig. 4.18a). A similar approach is used to look at the possible 1998 shift (six years of data either side) (Fig. 4.18b) but no such changes to the decimal distribution exist. Only the 1982 shift will be considered from here onwards as its likely cause is clearest, and it is the more significant shift when considering long-term trends.

The majority of marine data can be attributed to a deck number. Thus it is possible to look at the decimal distribution of each deck present over the period. These are listed in Table 4.6 and relevant histograms of decimal distribution of nine years of T_{dw} data before and after 1982 shown in Fig. 4.19. Two decks (732 and 889) report all T_{dw} data as whole numbers before 1982 and to one decimal place afterwards. However, they only account for 4.2 % and 4.6 % of the data before and after 1982 respectively. Three further decks (892, 896 and 926) exhibit a bias towards whole numbers that is stronger before 1982 and these collectively account for 74.4 % and 79.4 % of the data before and after 1982 respectively.

There is no documentation available at present detailing rounding conventions. This is most likely at the discretion of each Country or fleet. To investigate the scale of the effect of rounding T_{dw} on resulting q and RH, three rounding techniques are applied to a random selection of 100 T and T_{dw} pairs: truncation; rounding up; and rounding up or down to the nearest integer from 0.5 (normal rounding). The mean bias for truncation, rounding up and normal rounding of RH respectively is -1.3 %, 1.0 % and 0.3 %. The mean bias for q is -0.46 g kg^{-1}, 0.42 g kg^{-1} and 0.02 g kg^{-1} respectively. When applied to post-1982 marine humidity data (1983-84) (Figs. 4.20 and 4.21), it is clear that a global practice of rounding up vastly overestimates the shift and truncation gives the opposite sign. However, as it is not an entirely global phenomena it is feasible that a mix of rounding practices applying to different ships were in place with a bias towards rounding up and normal rounding. This is still largely speculative and requires further investigation.

Overall, the findings point to a non-climatic cause of the 1982 shift which has implications for calculated trends over the period. However, no adjustments have been made for this shift in terms of the blended product of HadCRUH for two reasons. Firstly, the true cause is not yet conclusively identified. Secondly, the shift is not entirely global such that any adjustment should be applied only to affected regions. The finished product will provide better understanding of humidity, over the oceans and over the period of record, upon which an informed decision on adjustment can be made. Any adjustment would inevitably introduce error and affect uncertainty estimates making it vital to have a complete understanding of where, how and why the adjustments were made.

It is important to consider the implications of leaving the data unadjusted for all analyses however. Trends for the post-shift period 1982 to 2003 are calculated to see if there is any difference to trends for the whole period (Table 4.7). In all cases the post-1982 trends are more positive. For *RH*, trends for the Globe, Southern Hemisphere and Tropics are now positive and highly significant (at the 1 % level) and so still discordant with the assumption that *RH* remains largely constant over time (*cf.* section 4.3.2). Notably, the trend for the Northern Hemisphere, a region of high data density and likely higher data quality relative to other regions, is closer to zero and not significant.

4.5 CONCLUSIONS

The marine *q* and *RH* gridded components of HadCRUH have been constructed and are analysed at a range of spatial and temporal scales. Overall, humidity over the oceans has changed, in some regions significantly (relative to dataset variability), over the last 30 years. Key findings are summarised below:

- Climatologically, *q* varies from ~1 to 20 g kg^{-1} decreasing meridionally from the Tropics with strong zonal continuity. *RH* ranges mostly from ~70 to 90 %, peaking in the high latitude Summer Hemisphere.
- Overall, atmospheric water vapour content is increasing. Trends in *q* are highly significant (at the 1 % level) when averaged over the: Globe (0.07 g kg^{-1} $10yr^{-1}$); Northern Hemisphere (0.08 g kg^{-1} $10yr^{-1}$); and Tropics (0.10 g kg^{-1} $10yr^{-1}$). These are smaller than over land.
- Trends in *q* are largest in the Tropics and Northern Hemisphere Summer, where ambient *T* is high relative to other regions/seasons (water is assumed to be always available across all except ice covered regions). Thus, over large scales at least, this is in accordance with the Clausius-Clapeyron relation under the assumption of near-constant *RH*.
- Trends in *RH* are all negative and highly significant (at the 1 % level) when averaged over the: Globe (-0.10 % $10yr^{-1}$); Northern (-0.10 % $10yr^{-1}$); and Southern (-0.11 % $10yr^{-1}$) Hemispheres. This contradicts assumptions of near-

constant *RH* over time. Seasonal trends however are mostly negative and not significant. They are positive for the Northern Hemisphere JJA and SON.

- There is a shift unique to marine humidity giving a positive bias to pre-1982 data (found to be significant almost globally at the 0.05 % level). This is most likely of non-climatic origin. A caveat for all marine trends thus pertains such that trends are at present likely underestimated. A first-order adjustment based on coastal grid-box timeseries yields a new global marine q trend of 0.12 g kg^{-1} 10yr^{-1} which is slightly larger than that over land.

The marine data have been proven to be sufficient for HadCRUH providing reasonable spatial and temporal coverage to give a useful representation of the global climate from 1973 to 2003. Issues of: uncertainty, bias and error; improvements to the quality control process; and further investigation into the 1982 shift, and possible 1998 shift, including possible adjustment if these are truly non-climatic phenomena are desirable for future versions of HadCRUH.

4.6 TABLES AND FIGURES FOR CHAPTER 4

Problem	Discussion and Implications	Error	Source
Solar heating bias	The effect on humidity is thought not to be greater than 0.1 g kg^{-1}	---	Kent & Taylor, 1996
Screen vs Psychrometer	Psychrometer readings of T_{dw} tend to be lower and are generally thought more reliable than that from Screens – North Atlantic study region	Linear correction to screen T_{dw} values of 1 ^{o}C at -5 ^{o}C and 0 ^{o}C at 26 ^{o}C transfer to corrections in q of 0.22 and 0.0 g kg^{-1} respectively	Kent *et al.*, 1993
Wick contamination	Wet-bulb thermometers at sea were thought to be at risk from contamination by salt and algae but this was found to be negligible	---	Taylor, P. K., *pers. comm.*
Random observation error bias correction	Causes of random observation error: - wet-bulb covering (ice, water or a mixture) - stem heat conduction - wick drying - poor ventilation - moisture on the dry bulb After removal of outliers > 4.5 σ the random observation error in q over 45 ^{o}S – 75 ^{o}N ranged between 0.6 ± 0.04 g kg^{-1} to 1.8 ± 0.3 g kg^{-1}	Mean error estimate for q corrected to 10 m neutral recording height = 1.1 ± 0.2 g kg^{-1}	Kent *et al.*, 1999; Kent & Berry, 2005

Table 4.1: Documented issues with marine humidity measurements and error corrections.

QC Test	Description – an observation is rejected if:
BASE QC – GOOD POSITION	The platform position at the time of observation is not consistent with its track – track is calculated using ship and wind speed.
BASE QC – GOOD LOCATION	The longitude and latitude of the platform are not plausible (-90 ^{o}-90 ^{o} latitude, -180 ^{o}-180 ^{o} longitude).
BASE QC – GOOD DATE	The date given is not plausible – 1-31 days, 1-12 months.
BASE QC – OVER SEA	The location is not over a water body.
GOOD DEWPOINT TEMPERATURE QC (T_{dw} QC)	The T_{dw} value is greater than the T value.
GOOD RELATIVE HUMIDITY QC (*RH* QC)	The *RH* value is less than 40 %.
GOOD TEMPERATURE QC (T QC)	The T value fails its outlier or neighbour check.

Table 4.2: First stage quality control test descriptions.

QC Test		1st Stage QC	2nd Stage QC – 1st Iteration	2nd Stage QC – 2nd Iteration
No. of obs before QC		47 032 905	39 657 025	39 404 909
No. of obs after QC		39 657 025	39 404 909	39 235 951
% removed in processing		15.7 %	0.6 %	0.4 %
Base QC	Bad location	0 %	---	---
	Over land	1.6 %	---	---
	Bad date	0.1 %	---	---
	Bad position	2 %	---	---
Base QC total removal		3.6 %	---	---
Bad T		11.6 %	---	---
T_{dw} >T		0.3 %	---	---
RH < 40%		0.7 %	---	---
e neighbour fail		---	0.42 %	0.32 %
e climatology fail		---	0.07 %	0.09 %
q neighbour fail		---	0.43 %	0.32 %
q climatology fail		---	0.08 %	0.09 %
RH neighbour fail (not used)		---	55.15 %	3.18 %
RH climatology fail		---	0.03 %	0.04 %
Overall neighbour fail		---	0.43 %	0.32 %
Overall climatology fail		---	0.09 %	0.12 %

Table 4.3: Summary of data removal from all quality control tests. Whole numbers indicate the number of observations. Percentages indicate that removed at that stage/iteration.

Spatial/Temporal Distance from Candidate Station / **No. of Neighbours in Composite Mean**	**2 pentads and 1 º (111km)**	**2 pentads and 2 º (222km)**	**4 pentads and 1 º (111km)**	**4 pentads and 2 º (222km)**
> 100	(1) 2.5 σ	---	(6) 2.5 σ	---
16-100	(2) 3.0 σ	---	(7) 3.0 σ	---
6-15	(3) 3.5 σ	---	(8) 3.5 σ	---
1-5	(4) 4.0 σ	(5) 4.0 σ	(9) 4.0 σ	(10) 4.0 σ

Table 4.4: Neighbour composite criteria for the Neighbour Check QC and the corresponding standard deviation range within which the candidate value must lie. Numbers in brackets denote order of preference for criteria.

REGION	q (g kg^{-1} 10yr^{-1})			
	DJF	**MAM**	**JJA**	**SON**
GLOBAL (70 °N - 70 °S)	0.06*	0.06**	0.07**	**0.08****
NORTHERN HEMISPHERE (20 °N - 70 °N)	0.06*	0.06**	**0.14****	0.12**
TROPICS (20 °S - 20 °N)	**0.13***	0.12**	0.10*	0.11*
SOUTHERN HEMISPHERE (70 °S - 20 °S)	-0.01	0.00	**-0.03**	-0.01
	***RH* (% 10yr^{-1})**			
GLOBAL (70 ° N - 70 ° S)	**-0.14**	-0.08	-0.05	**-0.14**
NORTHERN HEMISPHERE (20 °N - 70 °N)	-0.21	-0.10	**0.37****	0.12*
TROPICS (20 °S - 20 °N)	-0.05	-0.09	-0.08	**-0.14**
SOUTHERN HEMISPHERE (70 °S - 20 °S)	-0.22	-0.06	**-0.29***	-0.26

Table 4.5: Regionally averaged trends in *q* and *RH* for each season. Values of largest magnitude for each region are in bold. Shaded boxes are where trends are strongest in the land data. Trends are created using the LSR technique (Box 3.4). Significance at 5 % is shown with a * and at the 1 % with **.

Deck	Description	Presence of Decimals	% Data Before 1982	% Data After 1982
254	UK Met Office Marine Data Bank	No bias present	4.3	2.7
667	Inter-American Tropical Tuna Commission (IATTC)	No humidity data	0.0	0.0
714	Canadian Marine Environmental Data Service (MEDS)	No humidity data	0.0	0.0
732	Russian Marine Meteorological Data Set (MORMET)	100 % whole number pre-1982 and 100 % decimal after	1.1	1.1
780	Levitus World Ocean Atlas / Data Bank (WOA / WODB)	No bias present	0.4	0.4
883	US National Meteorological Centre (NMC, now NCEP) Data	No bias present	0.0	0.1
889	Autodin (US. Department of Defense Automated Digital Network)	100 % whole number pre-1982 and 100 % decimal after	3.1	3.5
892	US National Centre for Environmental Prediction (NCEP) Ship Data	Whole number bias before and after 1982 but larger before.	16.1	18.2
896	NCEP Miscellaneous (OSV, platform, rig) Data	Whole number bias before and after 1982 but larger before.	1.1	2.1
926	International Maritime Meteorological (IMM) Data	Whole number bias before and after 1982 but larger before.	52.9	54.5
927	International Marine (US or foreign-keyed ship data)	Whole number bias before and after 1982 but slightly larger after.	20.9	17.4

Table 4.6: Ship decks present around the period of the 1982 shift. Dark shading identifies decks where all pre-1982 data is recorded as integers. Lighter shading identifies decks where an integer bias is present in all data but stronger before 1982. Sourced from the Hadley Centre Marine Data Bank, Brohan, P. *pers. comm.*.

Trend Period and Variable	Global	Northern Hemisphere	Tropics	Southern Hemisphere
1982-2003 *q*	0.12**	0.14**	0.11*	0.09**
1973-2003 *q*	0.07**	0.08**	0.10**	0.01
1982-2003 *RH*	0.14**	-0.04	0.18**	0.25**
1973-2003 *RH*	-0.10**	-0.10**	-0.11	-0.11**

Table 4.7: Regionally averaged trends for the entire record and post 1982 record. Trends for q are in g kg^{-1} 10yr^{-1} and trends for RH are in % 10yr^{-1}. Trends are significance tested using REML (Box 3.4) where ** denotes significance at the 1 % level and * denotes significance at the 5 % level.

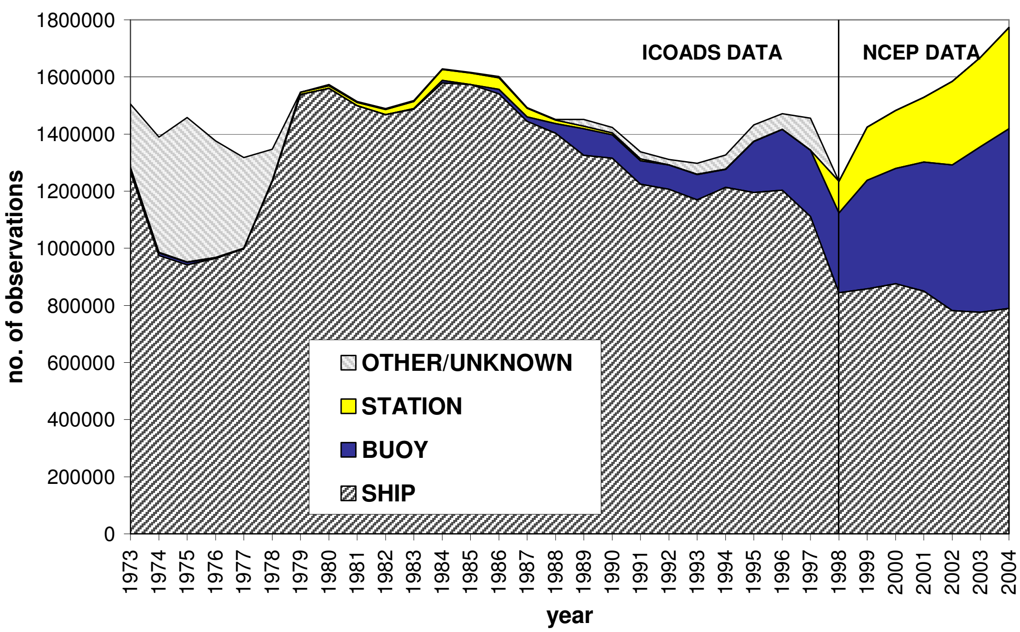


Figure 4.1: Frequency of marine humidity observations by platform type for 1973 to 2003. SHIP refers to merchant ships, military ships, ocean station vessels (on or off station), lightships and general ships. BUOY refers to moored buoys, drifting buoys and ice buoys. STATION refers to ice, oceanographic, coastal or island stations and fixed ocean platforms (rigs). OTHER refers to unknown, bathythermographs (MBT and XBTs) and Coastal Marine Automated Network platforms.

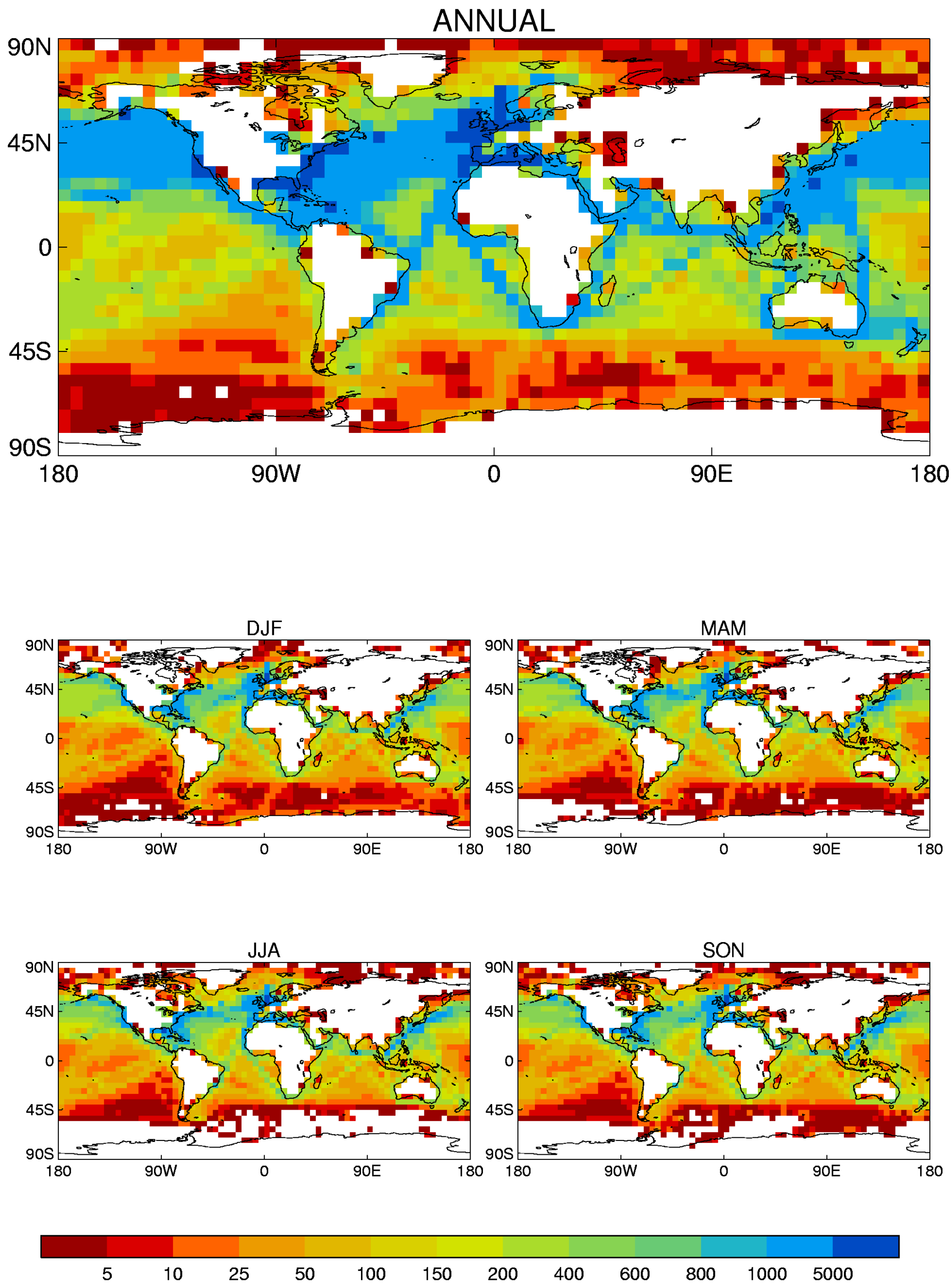


Figure 4.2: Average sampling density of marine humidity observations per year and per season from 1973 to 2003. All grid-boxes with a mean number of observations of less than 1 are shown as blank.

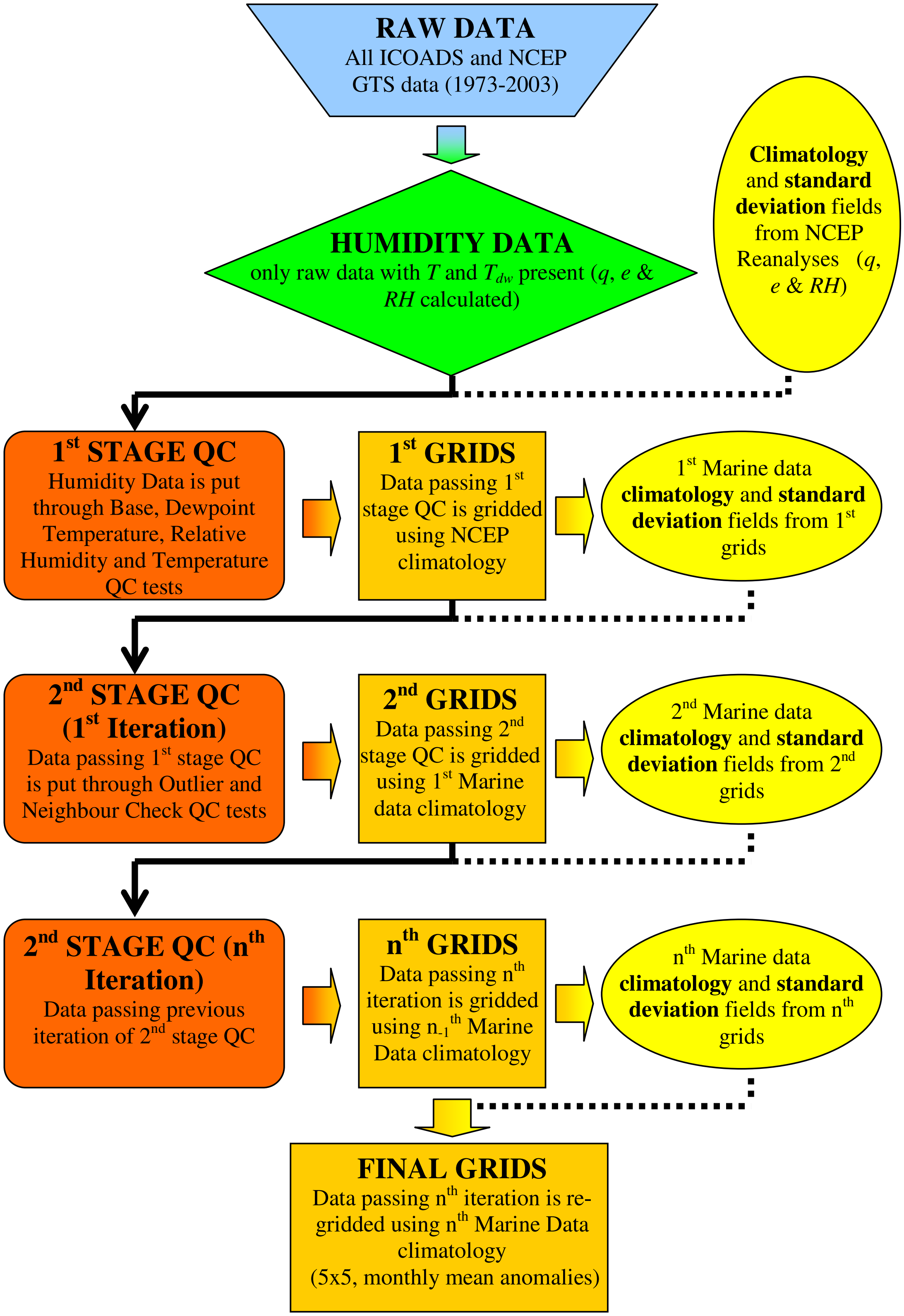


Figure 4.3: Flow diagram of the marine data system for HadCRUH.

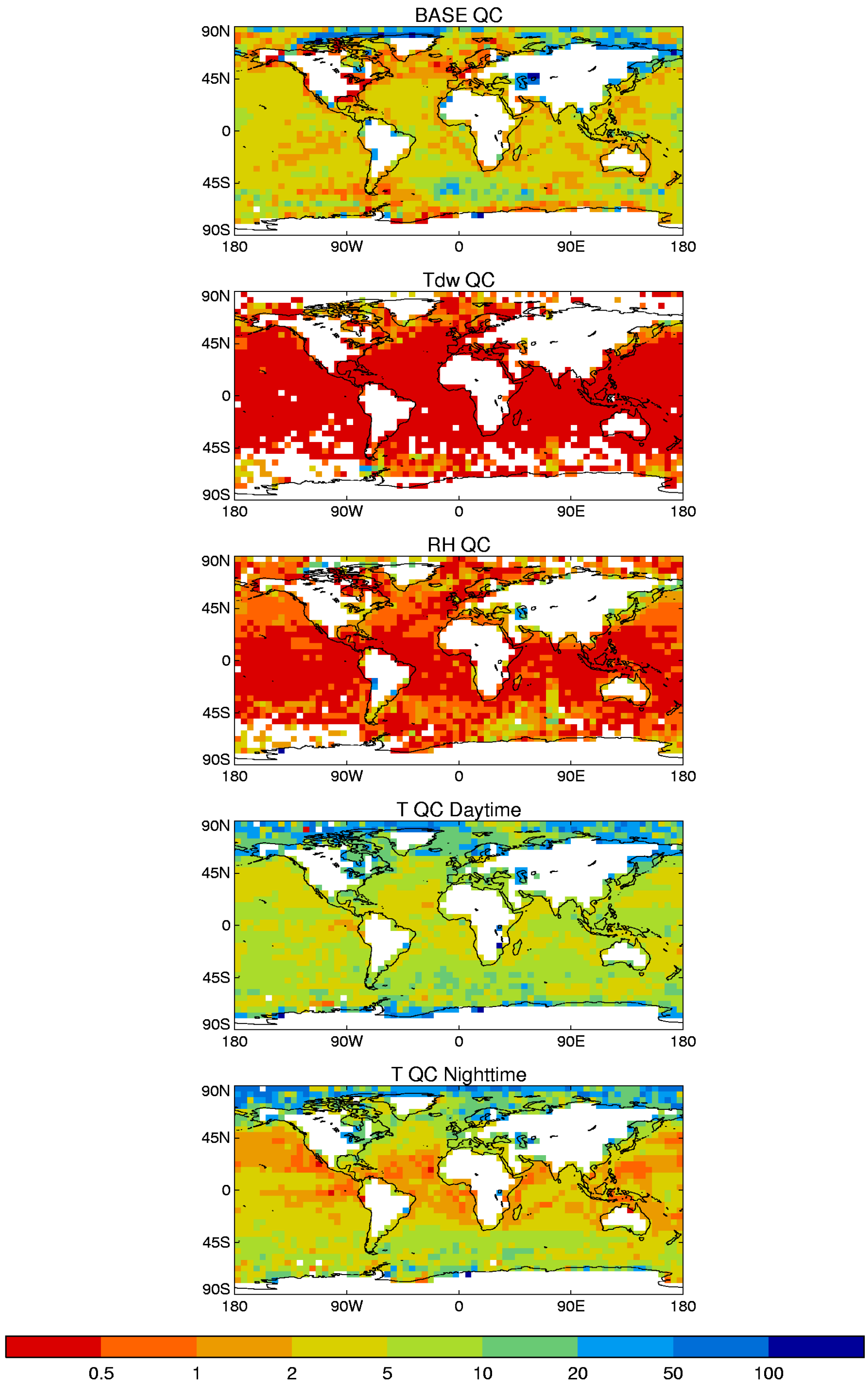


Figure 4.4: Percentage of all humidity observations failing quality control tests over the period 1973 to 2004. See Fig. 4.2 for a comparison with original data coverage and Table 4.2 for a description of tests.

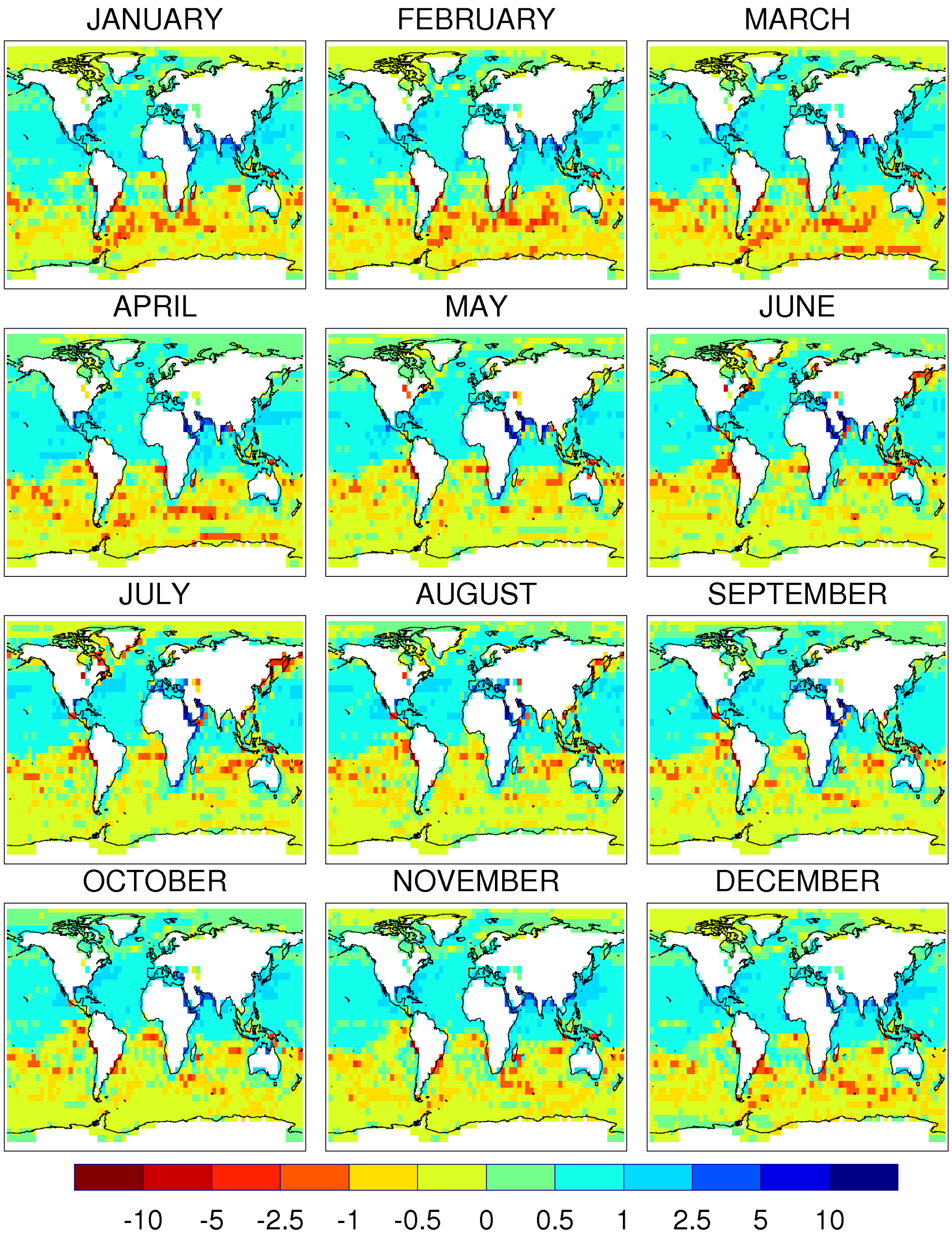


Figure 4.5: Difference between observation based monthly mean climatologies (1974 to 2003) and those derived from NCEP Reanalyses for *q*. NCEP Reanalyses climatologies are subtracted from observation based climatologies. Units are in g kg^{-1}.

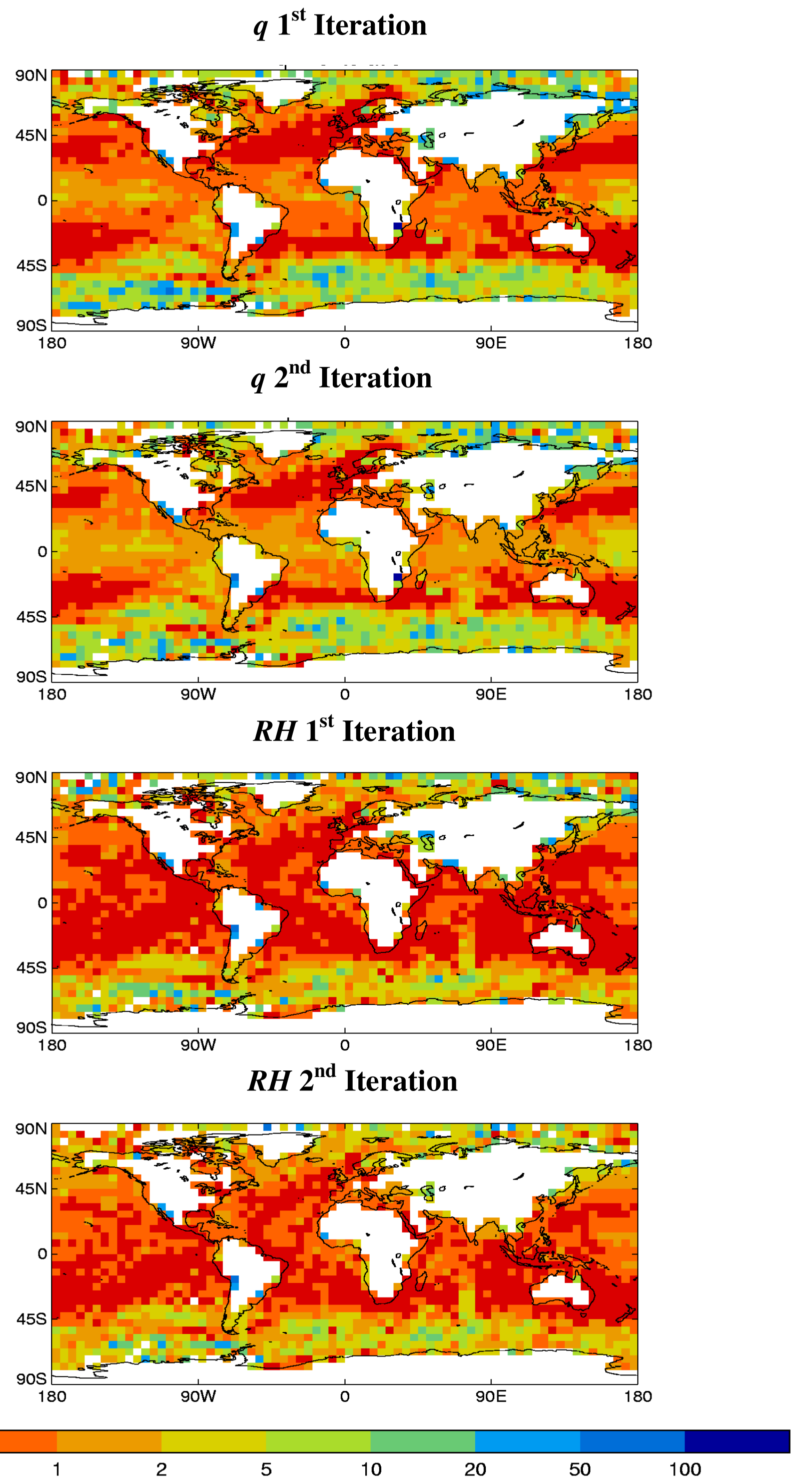


Figure 4.6: Percentage of humidity observations failing the Outlier QC test over the period 1973 to 2004. This includes observations that have already been flagged in other tests. See Fig. 4.2 for a comparison with original data coverage.

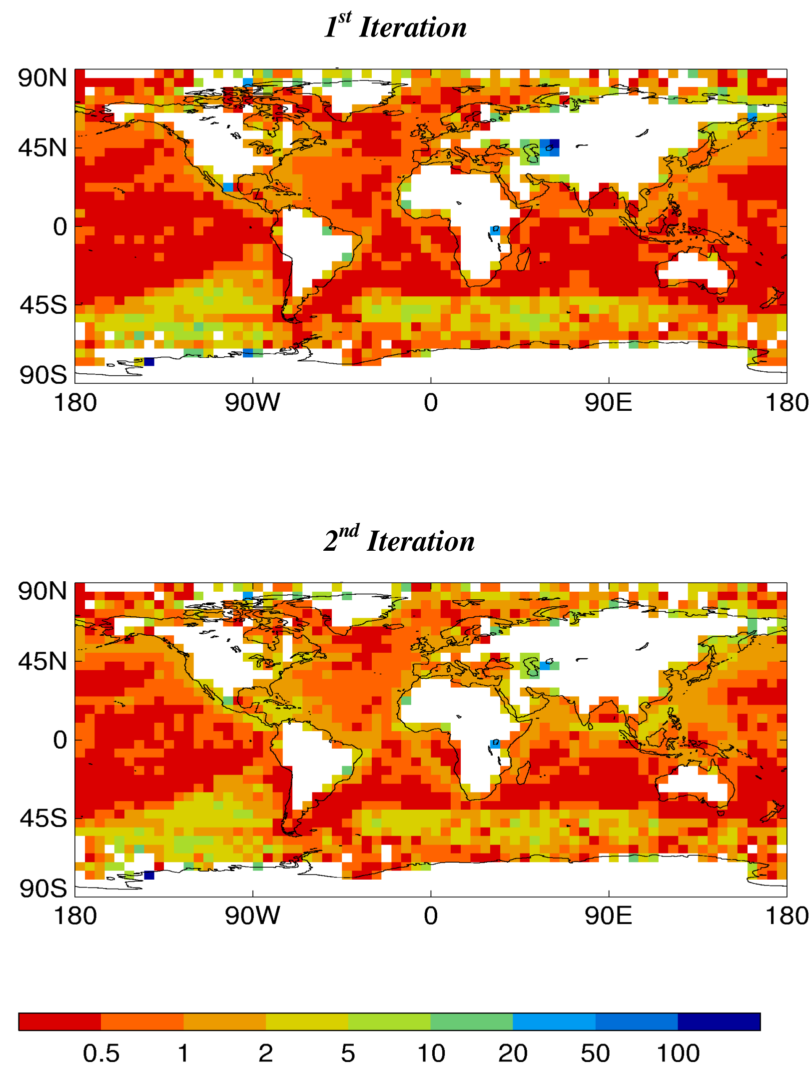


Figure 4.7: Percentage of humidity observations failing the *q* Neighbour Check QC test over the period 1973 to 2004. This includes observations that have already been flagged in other quality control tests. See Fig. 4.2 for a comparison with original data coverage.

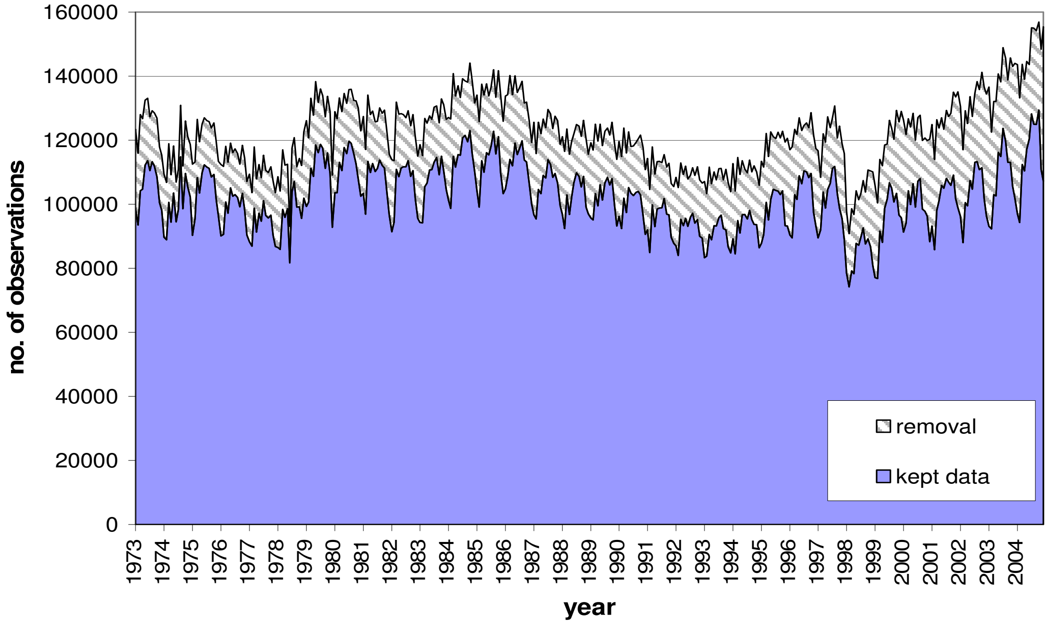


Figure 4.8: Observation density and removal by month.

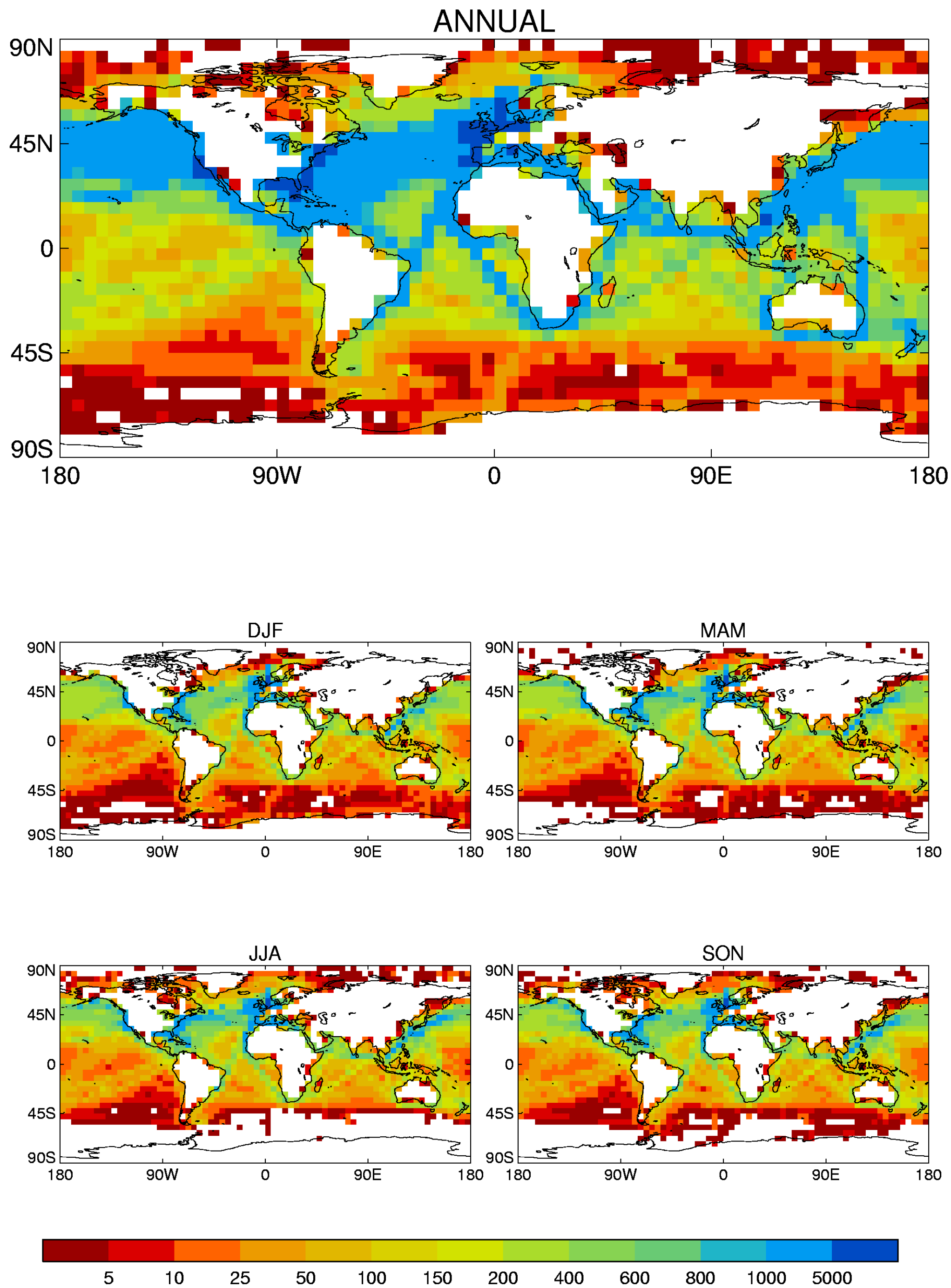


Figure 4.9: Average sampling density of the quality controlled and homogeneity checked humidity observations per year and per season from 1973 to 2003. See Fig. 4.2 for a comparison with original data coverage. All grid-boxes with a mean number of observations of less than 1 are shown as blank.

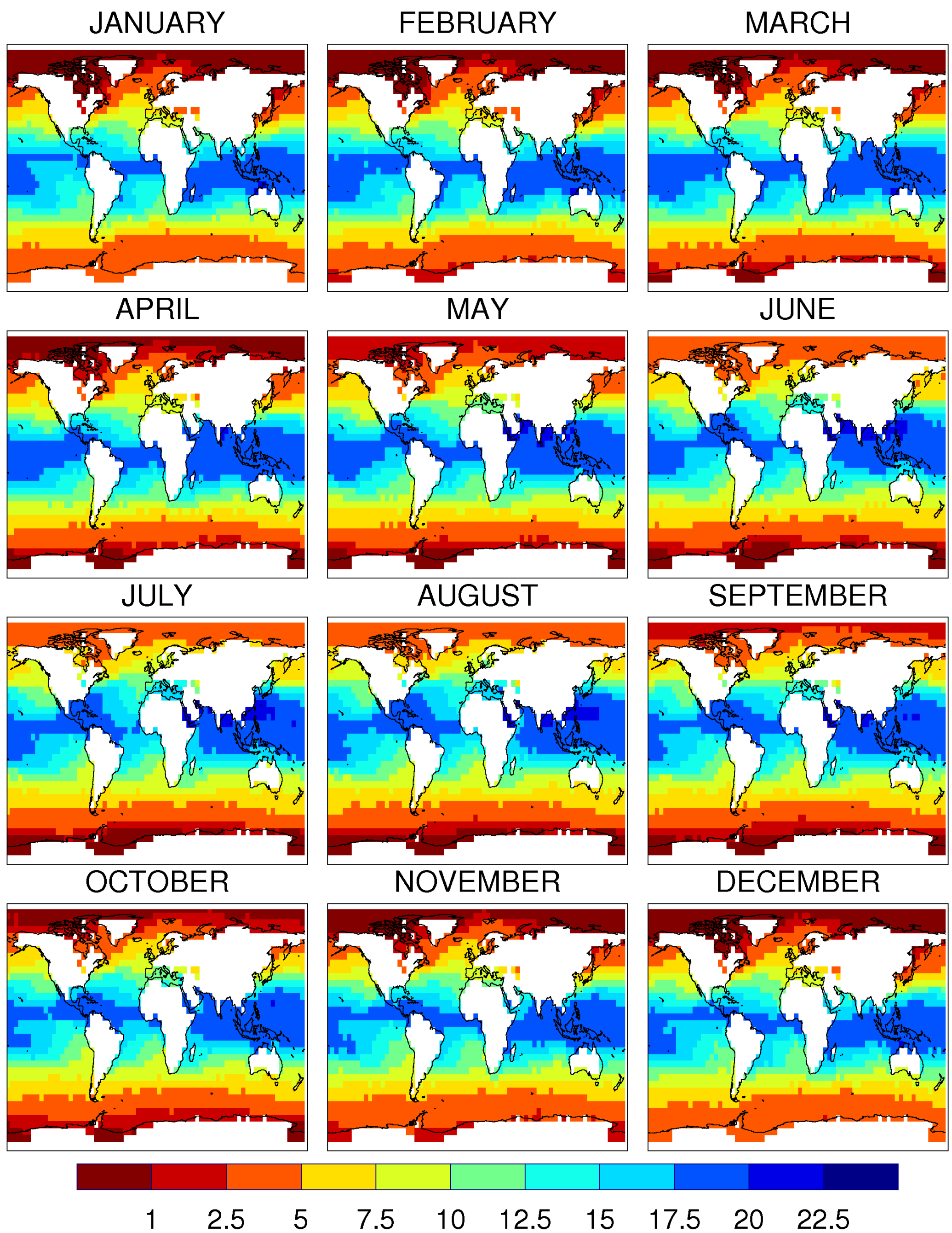


Figure 4.10: Monthly climatologies derived from observational marine data monthly mean *q* fields (gaps filled with NCEP Reanalyses) over the period 1974 to 2003. Units are in g kg^{-1}.

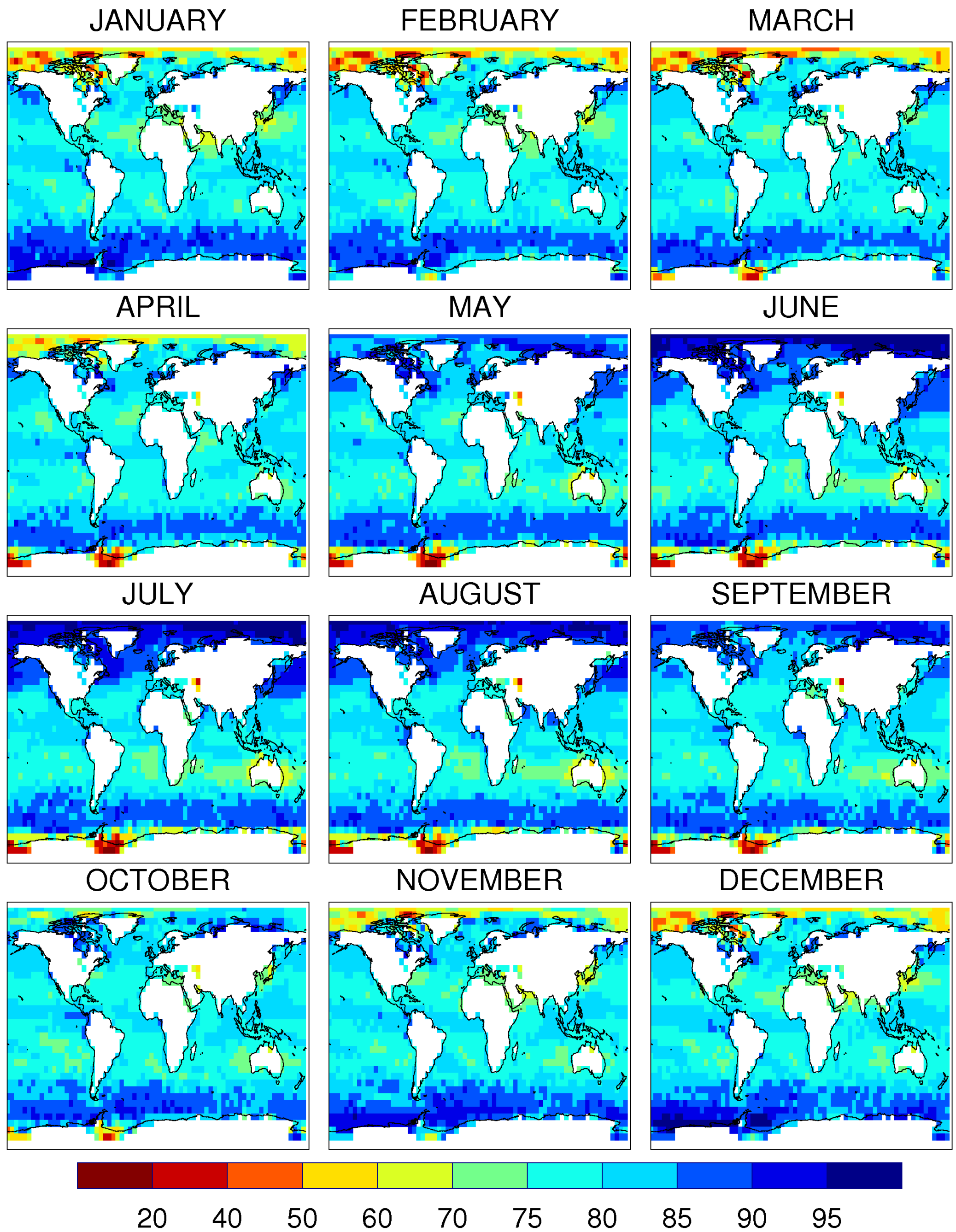


Figure 4.11: Monthly climatologies derived from observational marine data monthly mean *RH* fields (gaps filled with NCEP Reanalyses) over the period 1974 to 2003. Units are in %.

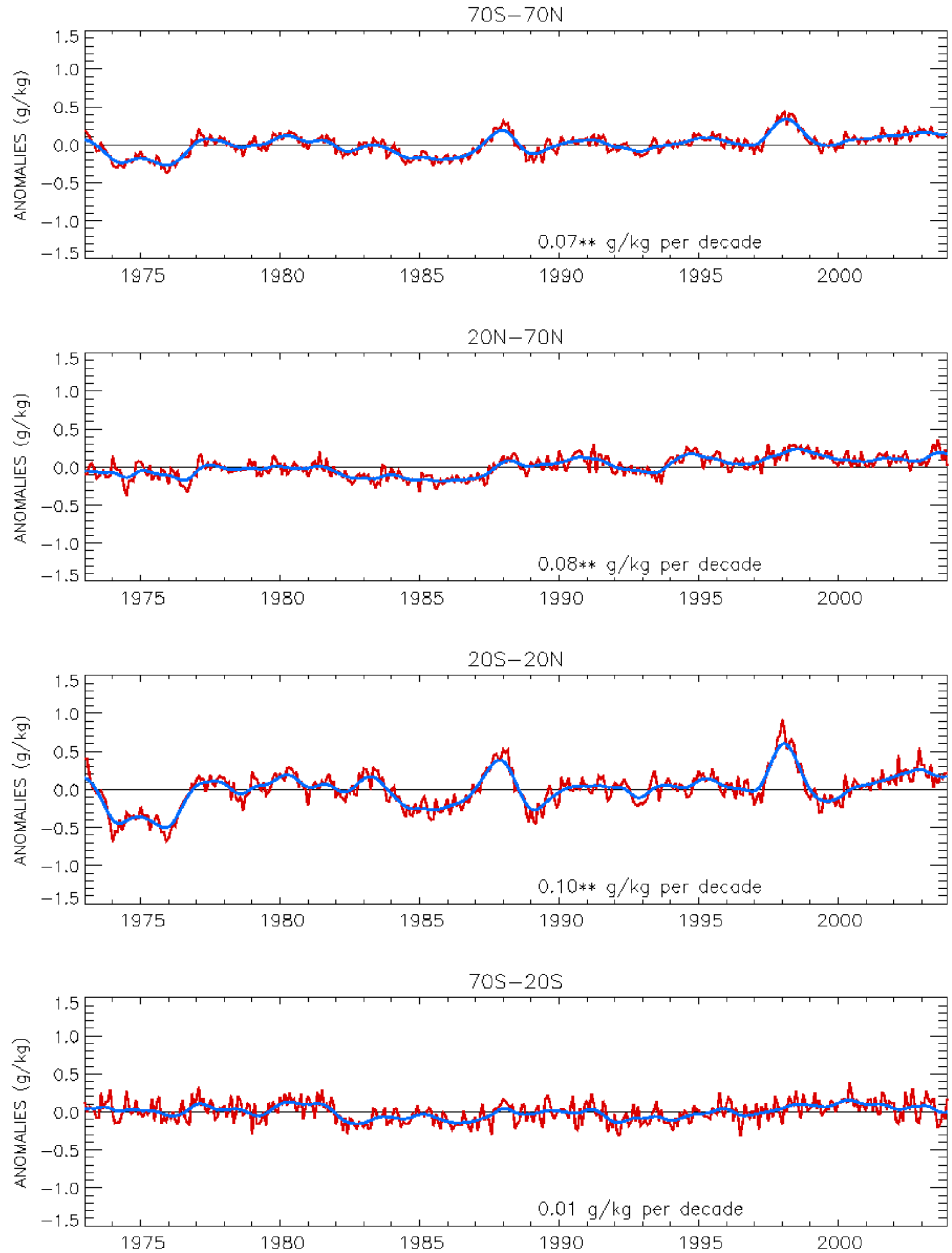


Figure 4.12: Regionally averaged monthly mean anomaly *q* timeseries and trends from 1973 to 2003. Trends are significance tested using REML (Box 3.4) where ** show significance at the 1 % level, * shows significance at the 5 % level. Red lines are the monthly mean anomaly timeseries and blue lines are the 21 point Gaussian weighted filtering for the low frequency component of the timeseries.

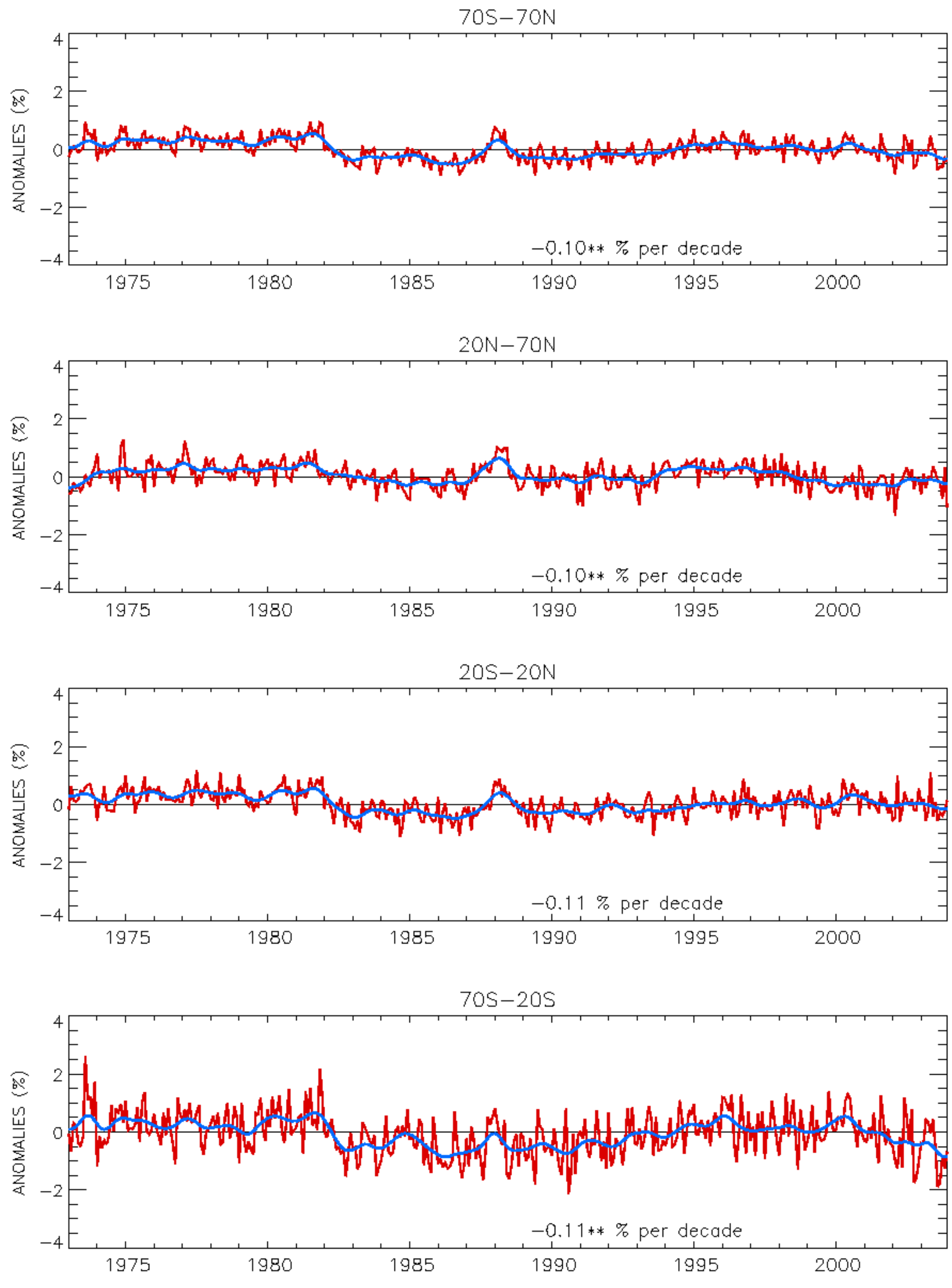


Figure 4.13: Regionally averaged monthly mean anomaly *RH* timeseries and trends from 1973 to 2003. Trends are significance tested using REML (Box 3.4) where ** show significance at the 1 % level, * shows significance at the 5 % level. Red lines are the monthly mean anomaly timeseries and blue lines are the 21 point Gaussian weighted filtering for the low frequency component of the timeseries.

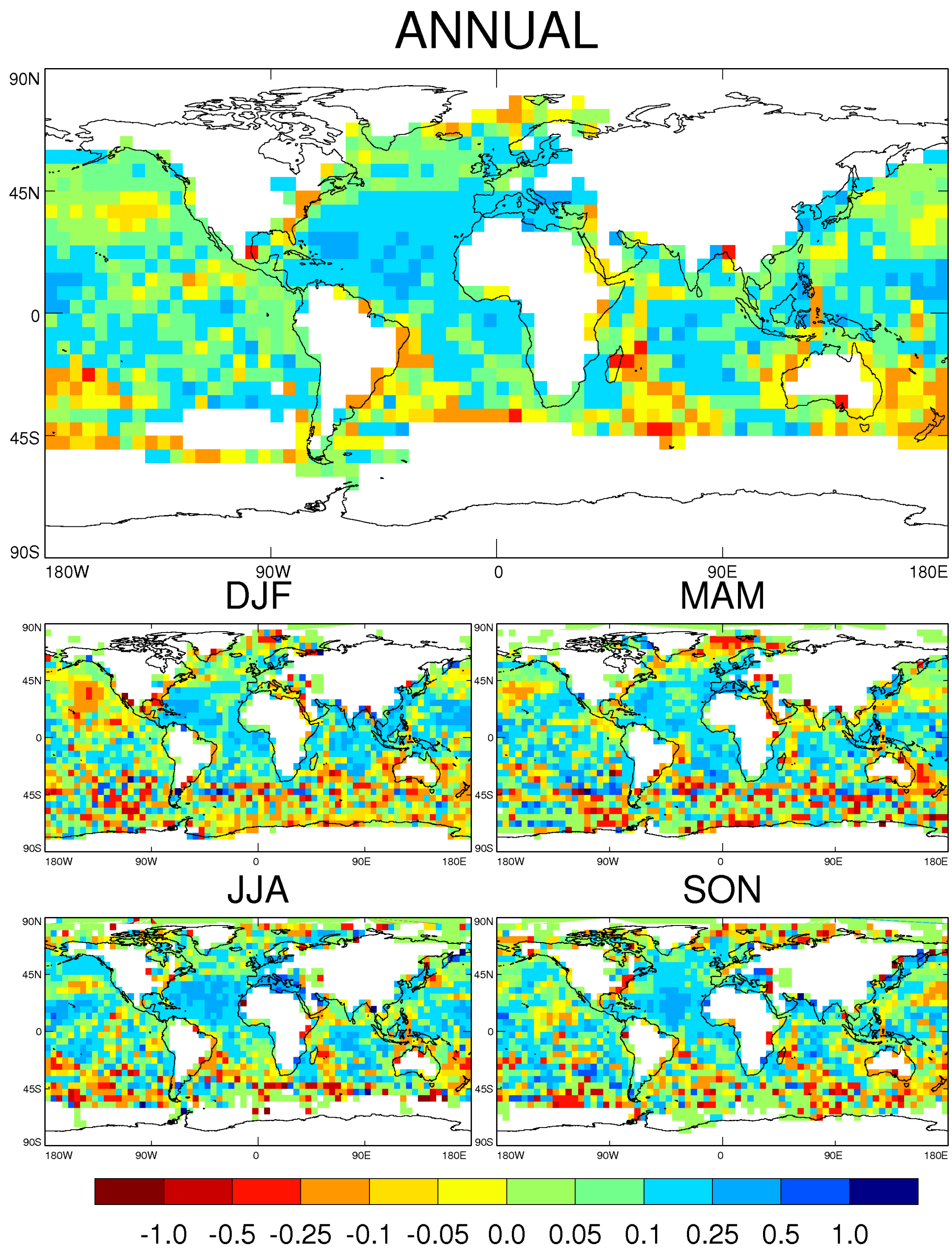


Figure 4.14: Grid-box trends for *q* for the period 1973 to 2003 calculated for the whole annual cycle and seasonal averages. Units are in g kg^{-1} 10yr^{-1}. Trends are calculated using the MPS method (Box 3.4). At least 50 % of months (seasons) must be present to calculate trends for the whole (seasonal) timeseries. As only two months per season are required to create seasonal averages a grid-box may have a trend present in all seasons but not when trends are calculated over the whole dataset.

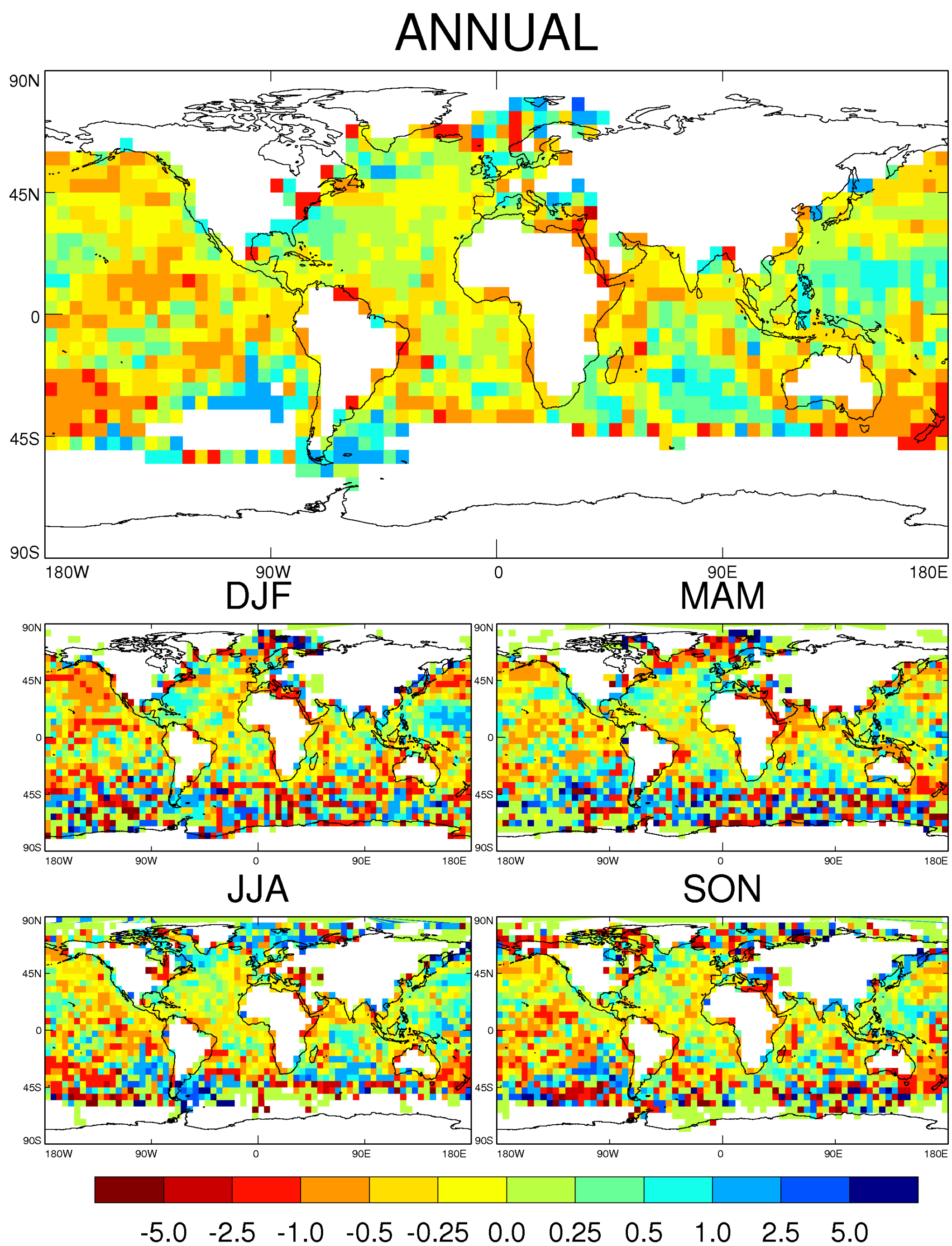


Figure 4.15: Grid-box trends for *RH* for the period 1973 to 2003 calculated for the whole annual cycle and seasonal averages. Units are in % $10yr^{-1}$. Trends are calculated using the MPS method (Box 3.4). At least 50 % of months (seasons) must be present to calculate trends for the whole (seasonal) timeseries. As only two months per season are required to create seasonal averages a grid-box may have a trend present in all seasons but not when trends are calculated over the whole dataset.

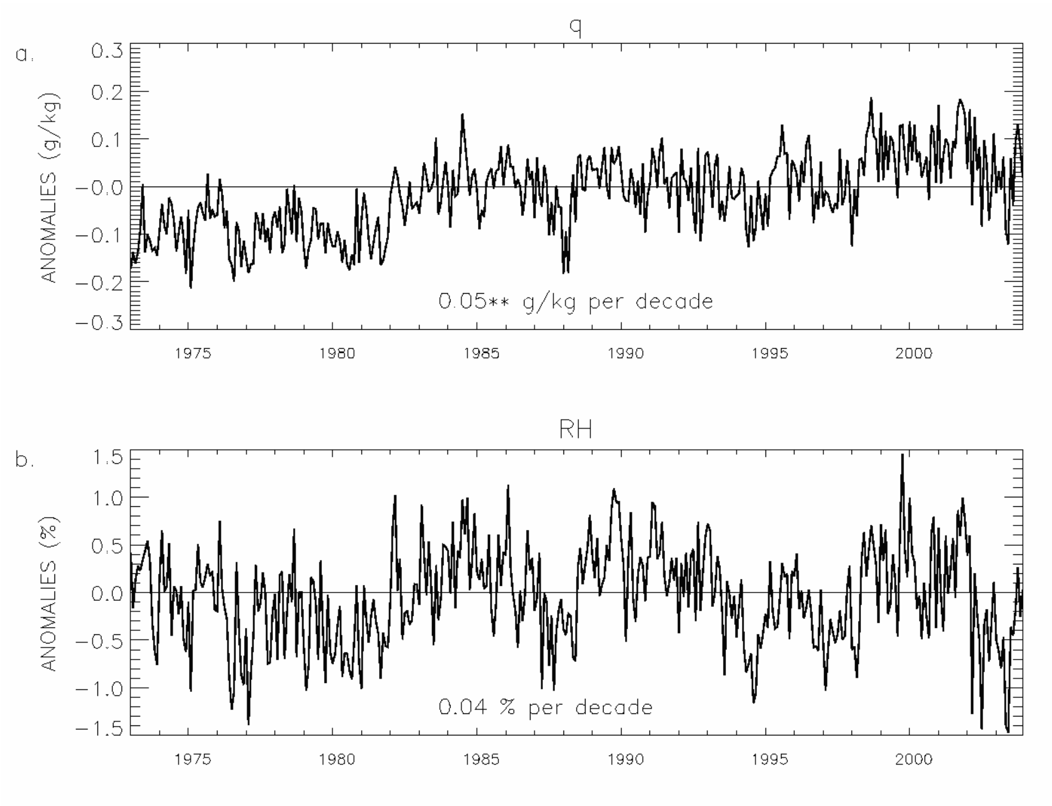


Figure 4.16: Globally averaged difference series and trends of (land – marine) monthly mean anomaly *q* and *RH* for coastal grid-boxes. Trends are fitted using the REML method (Box 3.4) where ** shows significance at the 1% level.

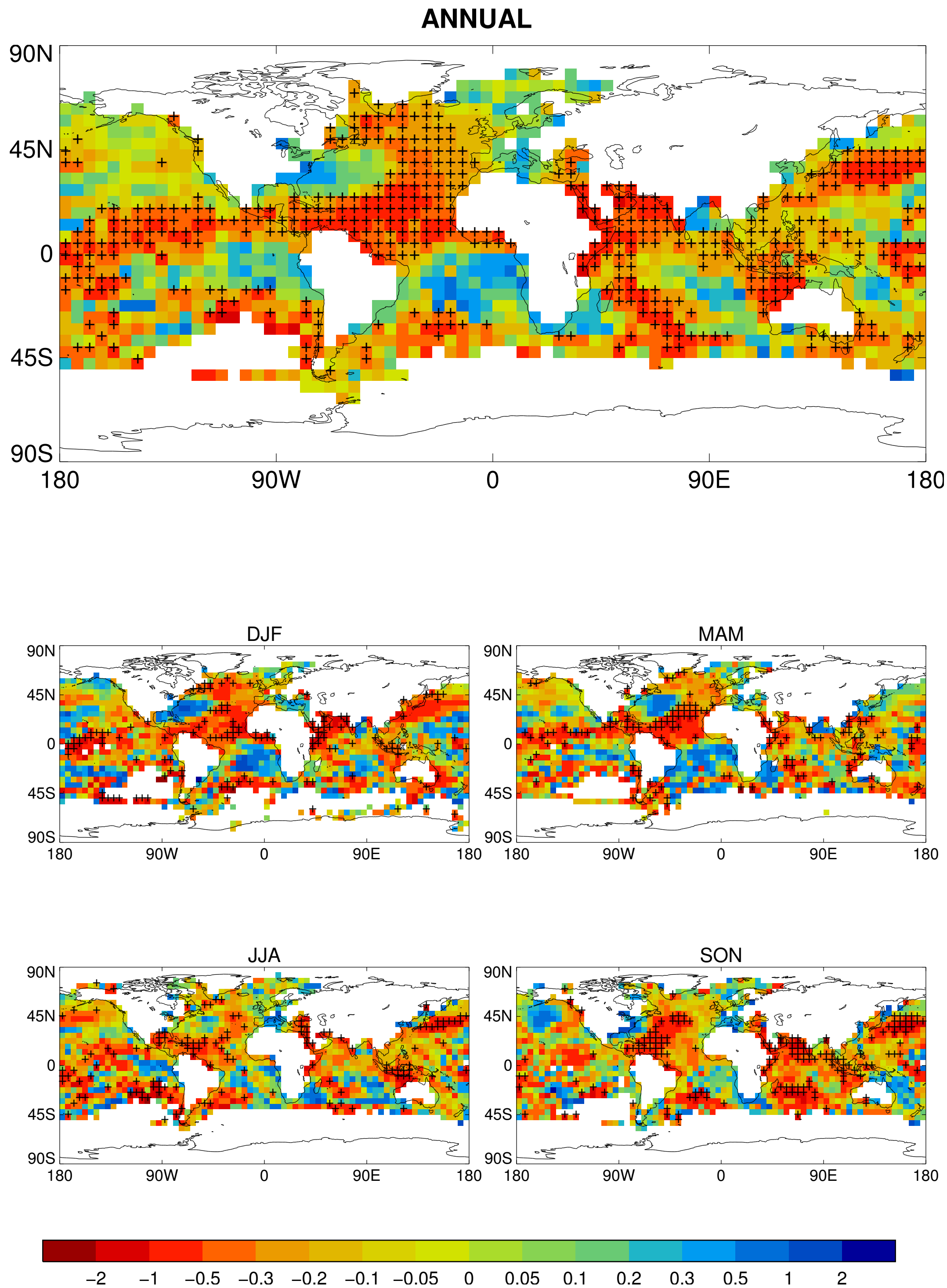
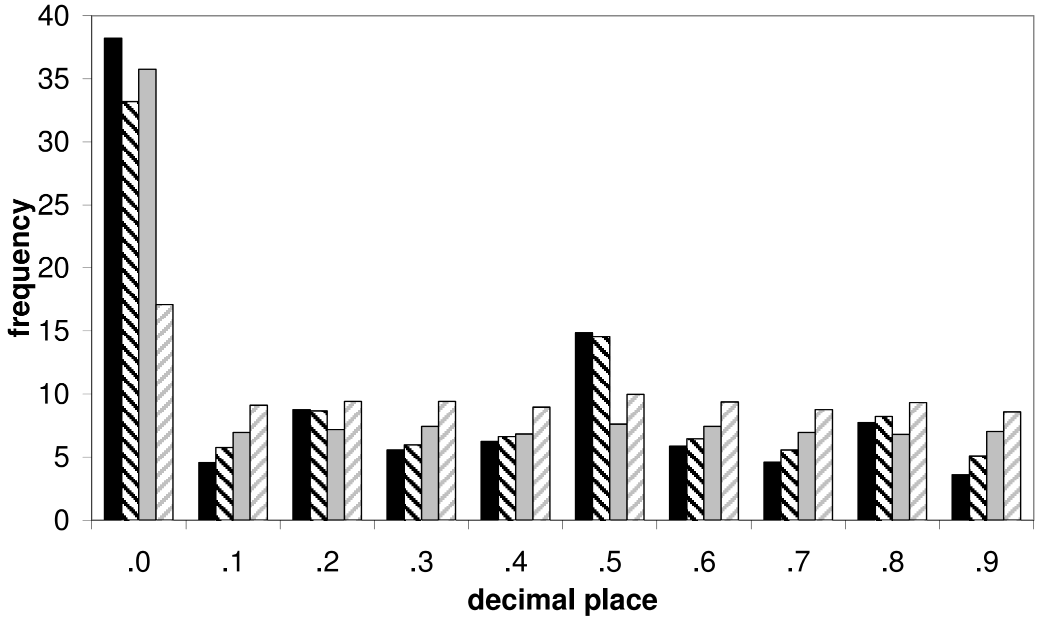


Figure 4.17: Difference maps of mean anomaly *q* 1983-1986 minus 1978-1981. Units are in g kg^{-1}. Crosses show grid-boxes with significantly different timeseries (at the 0.05 level (t test)) before and after the shift.

a)



b)

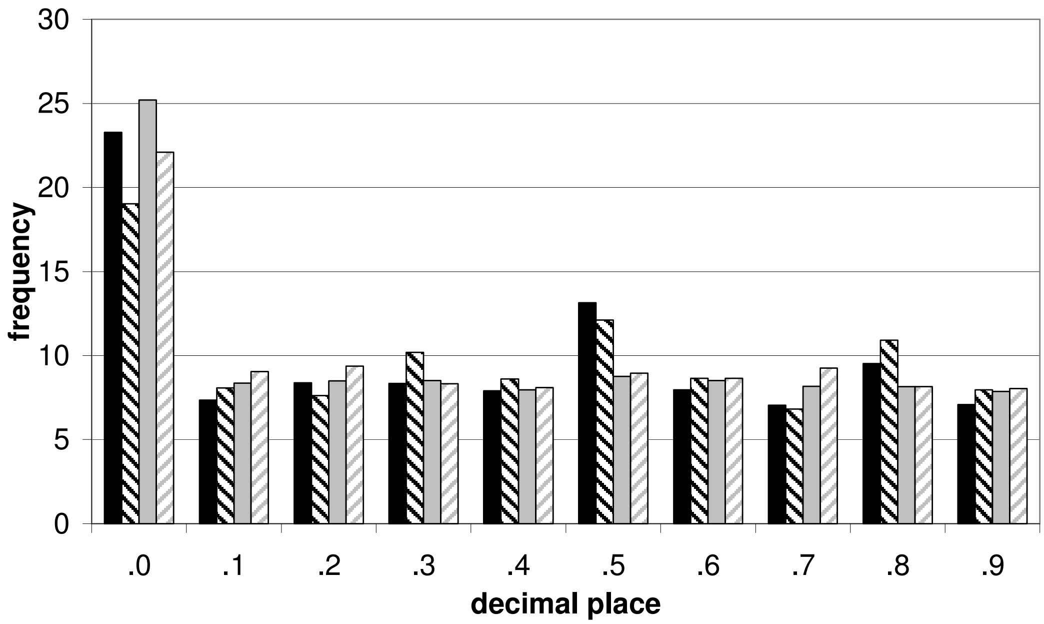


Figure 4.18: Decimal place distribution in the marine data. Plot a) shows data for nine years before and after 1982. Plot b) shows data for six years before and after 1998. Black (grey) solid bars show T (T_{dw}) data before the shift and black downwards (grey upwards) slashed bars show T (T_{dw}) data after the shift.

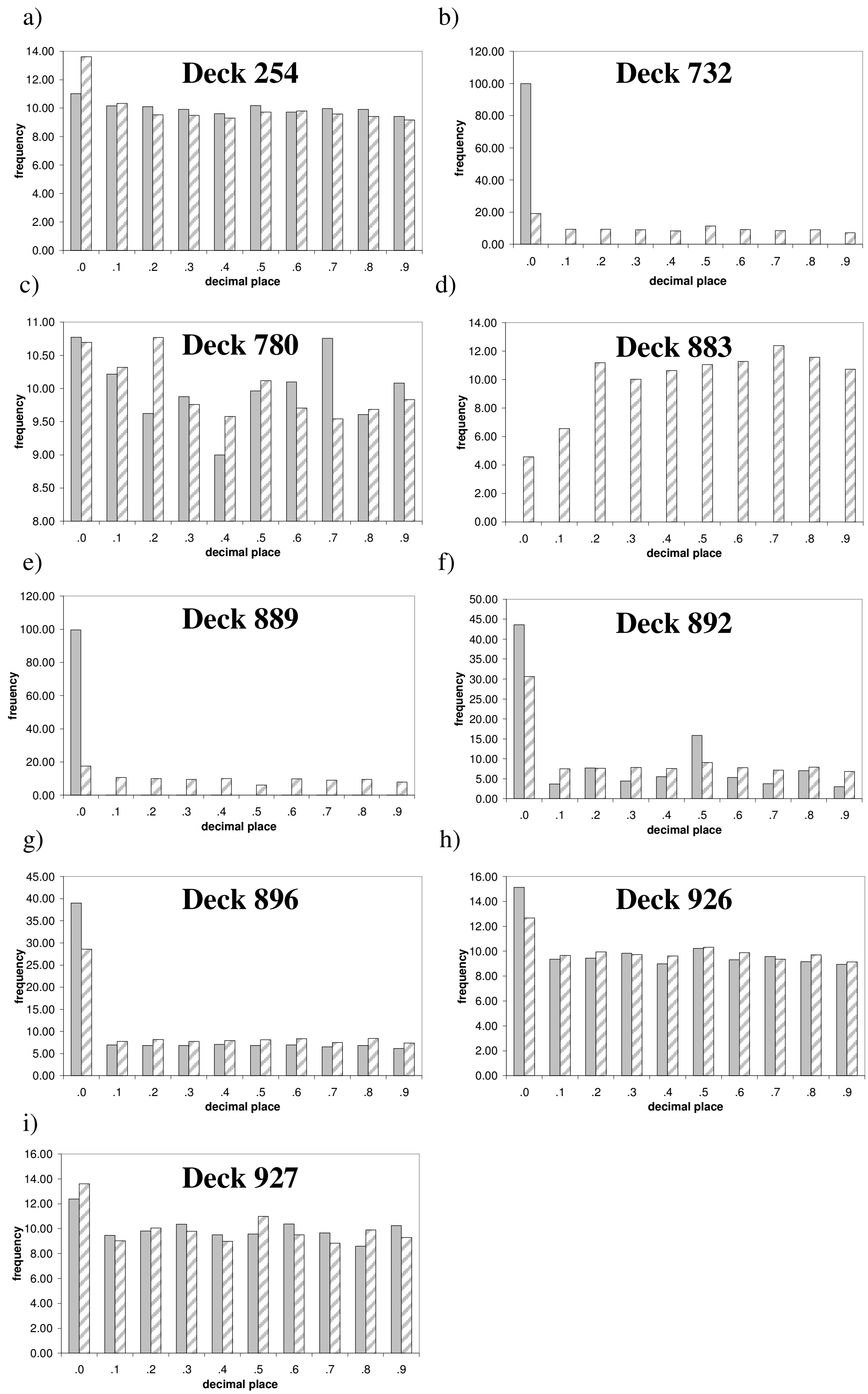


Figure 4.19: Decimal place distribution for the marine humidity data by deck. Grey solid (dashed) bars show T_{dw} data for nine years before (after) 1982. Frequency is in % of data from the deck in question. Decks are described in Table 4.6.

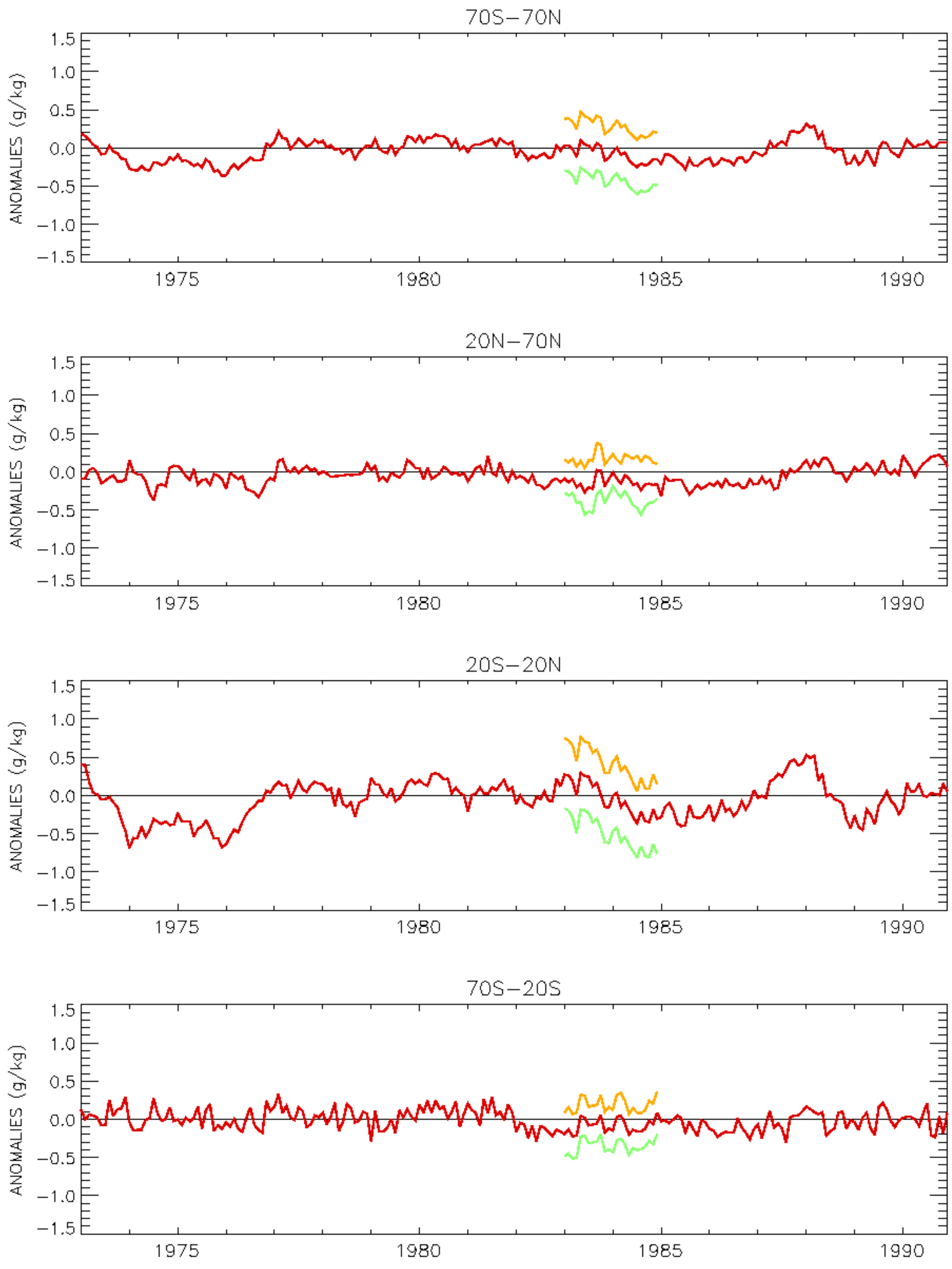


Figure 4.20: A comparison of the regionally averaged *q* monthly mean anomaly timeseries with truncation and rounding up applied to 1983-1984 data sections. Orange lines denote rounded up and green lines denote truncation of all available humidity data at the time.

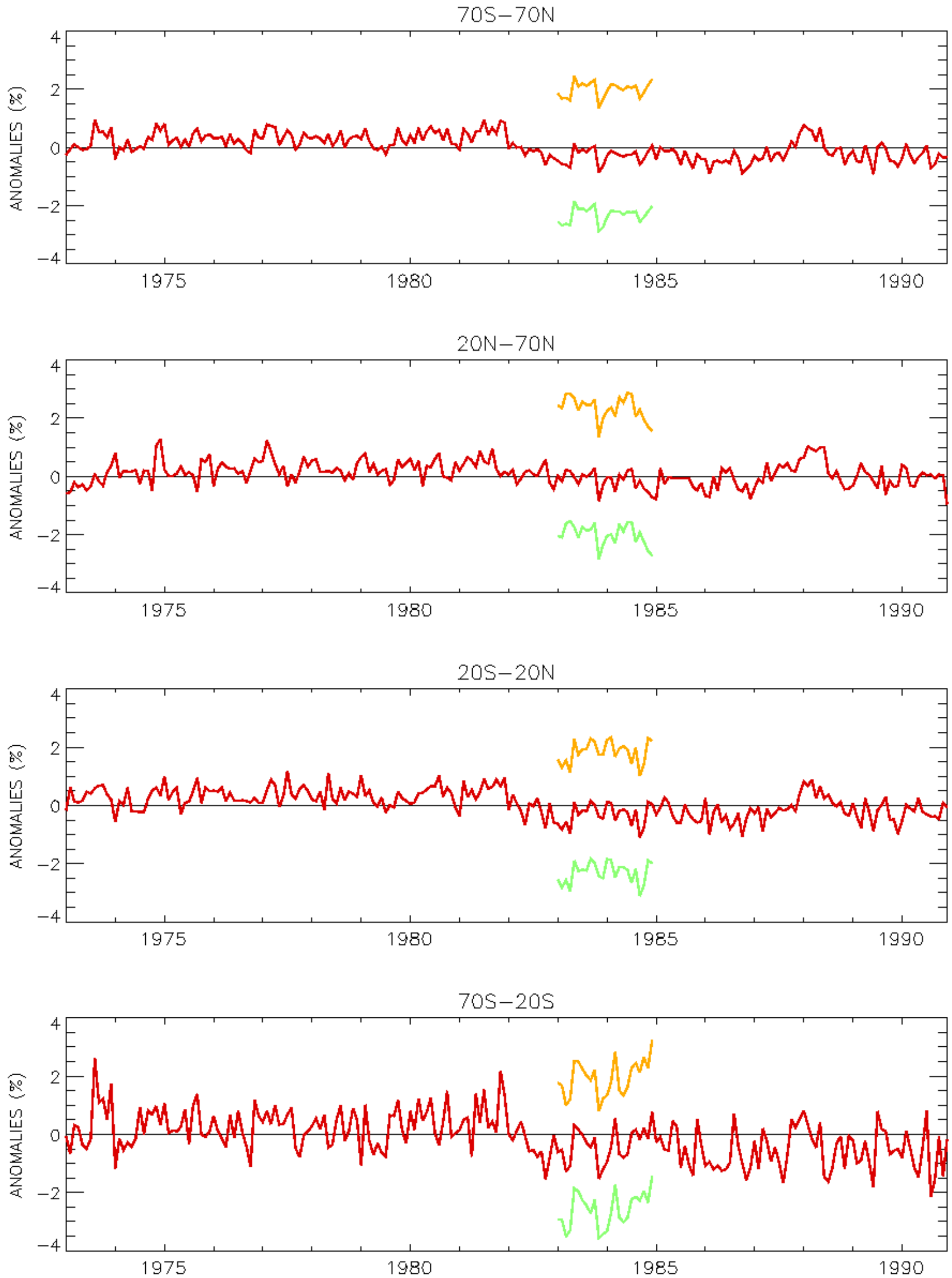


Figure 4.21: A comparison of the regionally averaged *RH* monthly mean anomaly timeseries with truncation and rounding up applied to 1983-1984 data sections. Orange lines denote rounded up and green lines denote truncation of all available humidity data at the time.