## Chapter 6

# Conclusions and Further Work

### SUMMARY OF KEY FINDINGS

- Following substantial quality control and homogenisation efforts to the global land and marine data, HadCRUH represents a plausible estimate of near-global surface humidity values from 1973 to 2003 and is suitable for climate studies.
- For the period 1973 to 2003 there has been a significant increase in surface water vapour over the Globe, Tropics and Northern Hemisphere. The Northern Hemisphere signal is detectable above HadCM3 climate model expectations of natural internal variability. Trends are larger over land, and in the Tropics and Northern Hemisphere Summer. Relative humidity trends remain essentially indistinguishable from zero on the largest spatial and temporal scales, but are sometimes significant at regional scales.
- Increases in specific humidity are largely consistent with observed changes in temperature and the magnitude of change as theorised by the Clausius-Clapeyron relation under the assumption of constant relative humidity. Temperature and specific humidity correlate strongly on all scales increasing confidence in the veracity of both HadCRUH and the global temperature record (from HadCRUT3).
- There appears to be a strong positive bias in marine humidity data prior to 1982. This most likely relates to a known change in reporting practice for dewpoint temperature at this time. As a result, trends in both specific humidity and relative humidity are likely to be underestimated over parts of the oceans.
- The anthropogenic forcings only ensemble from HadCM3 captures global surface humidity changes far better than the natural forcings ensemble. This points to a primarily anthropogenic origin of recently observed humidity changes, and reduces the chances that it is entirely natural in origin. *Thereby this work further supports the reality of recent human-induced climate change.*

## 6.1 PRINCIPAL FINDINGS IN THE LAND DATA

For the period 1973 to 2003 over land, climatological $q$ varies from 0 to 20 g $kg^{-1}$ and $RH$ ranges from 10 to 90 %. While $q$ has strong zonal homogeneity and decreases meridionally from its maximum in the Tropics, $RH$ has no such patterns. It is less spatially coherent than $q$ but does still exhibit some regional structure. ENSO signals are apparent in the timeseries, especially in $q$ in the Tropics. Not all major ENSO events however, are captured by the large scale regionally averaged timeseries.

There is an overwhelming picture of statistically significant moistening in $q$ for regional averages of the Globe (0.11 g $kg^{-1}$ $10yr^{-1}$), Northern Hemisphere and Tropics and for all seasons. There is no trend over the Southern Hemisphere. Conversely, trends in $RH$ are of either sign and not significant except for in the Southern Hemisphere (where they are negative) and some seasonally averaged trends. The zonal structure of $q$ and more regional structure of $RH$ are also apparent in trend analyses.

## 6.2 PRINCIPAL FINDINGS IN THE MARINE DATA

For marine data, climatological $q$ varies from 1 to 20 g $kg^{-1}$ and $RH$ ranges mostly from 70 to 90 %, both bearing structural similarities to the land data. ENSO signals in $q$ in the Tropics are apparent for some events (similar to those shown in the land data) and stronger than over land but much less apparent in $RH$ and of opposite sign to the land data.

As is the case for the land, overall, there is a picture of significant moistening in $q$, when regionally averaged over the Globe (0.07 g $kg^{-1}$ $10yr^{-1}$), Northern Hemisphere and Tropics, and for all seasons. Trends are mostly less than over land. Trends in $RH$ are all negative with the exception of the Northern Hemisphere Summer and Autumn, and significant when averaged for the Globe, Northern and Southern Hemispheres, and for some seasonally averaged trends.

A significant (in ~50 % of grid-boxes) negative shift occurs in the marine humidity in 1982. This is likely to be linked to the change in practice from recording $T_{dw}$ with integer values prior to January 1982 to quoting the first decimal place thereafter. The magnitude of the shift seen might be expected had there been a practice of rounding up $T_{dw}$ values prior to this date in some but not all ship decks. The cause is however, not conclusively proven. Error adjustments for globally averaged *q* and *RH* have been calculated based on the trend in land-marine differences for those coastal grid-boxes containing both types of measurement. These imply more positive trends in marine *q* and *RH*. Notably this results in *RH* trends being closer to zero than previously found, consistent with common assumptions of constant *RH* over large scales.

## 6.3 PRINCIPAL FINDINGS OF CLIMATE ANALYSES OF HADCRUH

The quality control and homogenisation processes applied to the land and marine components of HadCRUH, have been sufficient to develop a near-global dataset that exhibits plausible features and trends in humidity. Spatial coverage of HadCRUH for both *q* and *RH* is good within 40 $^{o}$S to 60 $^{o}$N excepting central Africa, the Amazon and parts of the Middle East.

Considering changes in *q*, all regionally averaged trends over the large scale, including seasonally averaged, are positive and significant except for the Southern Hemisphere where trends are indistinguishable from zero and not significant. Data coverage is poorer there, particularly over the Southern Ocean. Trends are strongest in the Tropics and seasonally in the Northern Hemisphere Summer. At the smaller regional scale, all regions except Central America, western and central North America, East Africa, and Australia show significant positive trends. For *RH*, regionally averaged trends over the large scale are mostly negative and not significant except for the Tropics and when seasonally averaged over the Northern Hemisphere where positive and significant trends are found in Summer and Autumn and significant negative trends are found in Spring. At the smaller regional scale, trends are of mixed sign and significant in 12 regions. This suggests that the assumption of constant *RH* is not robust for all regions and on all spatial and temporal scales.

*T* and *q* correlate consistently strongly ($r > 0.6$) at monthly mean anomaly resolution on all spatial scales. As two independent datasets, this substantiates both HadCRUT3 and HadCRUH as realistic climate records. At both the large and regional scale the magnitude of $q$ trends are generally as expected from trends in $T$ from HadCRUT3 and the Clausius-Clapeyron relation.

The observations, for the Northern Hemisphere regional average and smaller regions within and over the Tropics, show a climate signal detectable beyond that expected from natural internal variability provided by a range of overlapping segments from a HadCM3 control simulation with fixed external forcing. Humidity, $q$ in particular, is a previously unused diagnostic for formal climate change detection and attribution studies. Such studies include techniques to optimise the climate signal above background noise and so the fact that a number of regions already exhibit trends outside natural internal variability is encouraging.

The HadCM3 natural forcing only ensemble tends to underestimate the observed changes or produce negative trends. The anthropogenic only forcing ensemble best mimics the observations, performing better than anthropogenic/natural combined ensemble in general, but tending to overestimate trends slightly. These results have now been confirmed in an initial formal detection and attribution study based upon data used above (Willett *et al.*, *in prep.*) which finds a robust anthropogenic signal in the HadCRUH observations, and with much lower confidence a natural forcings signal.

## 6.4 AVENUES OF FURTHER WORK

Within the time limitations of a three year project only so many goals can realistically be reached in terms of dataset quality, resolution and analysis. Should further funding become available there are three key areas for focus: improving dataset quality and coverage; investigating biases and calculating appropriate adjustments; and creating new products from HadCRUH.

Primarily, HadCRUH can be improved by the addition of more data. From personal communication with NCDC (Vose, R), the land data source (ISH dataset) is not only regularly updated but there are also plans for the addition of new data in some regions including: Brazil; West Africa; Canada; Australia; Antarctica; and Russia. Problems with data sparsity over the US, likely a result of station records split into separate periods reporting under different station IDs, are also being looked into and should be improved in the near future. There is also the potential for merging with other datasets such as HAHN or NEW to improve coverage (section 3.1.4) (Fig. 6.1). The marine data are continuously updated and there are plans for digitisation of more records (MARCDAT-II, 2005). In addition to data coverage improvements since 1973 these developments may also lead to improved coverage before 1973, permitting the extension of HadCRUH further back in time.

Improvements to the quality control procedure and homogeneity tests would add value to HadCRUH and there is scope for improvement in all quality control tests applied. There are four key areas for improvement of the quality control procedure:

- *accounting for wind ventilation effects*,

  For both the land and marine data, wind information may need more attention. As discussed for the land data, no quantifiable problems are found but for a few case-study stations with known quality wind data, a more in depth study may prove interesting.

- *permitting low marine RH values in coastal regions,*

  All observations with an *RH* below 40 % are not included in HadCRUH for reasons discussed in section 4.2.2. However, in subsequent analysis this is found to remove proportionally more observations around coastlines and inland seas. Future versions could include a caveat on this test such that observations within a certain proximity to land that may experience very dry continental air are not excluded.

- *improving the marine Outlier QC,*

  The proportion of marine observations failing the Outlier QC test is higher around coastlines and in high latitudes. A caveat could be implemented here for observations within a certain proximity to land where variability may be greater. The effects of data sparsity on variance (in the case of high latitudes) could be investigated to create a more appropriate test for regions particularly sensitive to data removal from these tests.

- *and the application of marine Zero DPD and Repeats QC.*

  Zero *DPD* (wick drying) or Repeats QC tests are applied to the land data but not the marine data. Analyses show that considerable data would be removed by such a test (Table 6.1). The process of detection and removal is complex computationally with the data in its current form and so not possible for version one but could be considered for future versions.

The improved understanding of surface humidity (in space and time) provided by HadCRUH can now feed in to a more sensitive and informed Neighbour Check QC (marine data) and homogenisation (land data). Ideally this would bring in metadata where available including natural events (volcanoes and ENSO events) that might impact climate. Marine metadata for some platforms are now available electronically (http://icoads.noaa.gov/metadata/). This is obviously a large task requiring much effort. For the land and marine data, homogenisation (Neighbour Check QC) in *RH* could be improved. The land data, while sufficient for version one, could benefit from an entirely independent *RH* homogenisation with a more finely tuned outlier test, possibly latitudinally sensitive. For the marine data, the *RH* Neighbour Check QC removes far too much data and so is not implemented. This could be improved with better knowledge of *RH* variance over oceans.

Bias estimates and corrections are widespread in datasets to account for such things as increasing platform heights (marine data only), instrument exposure changes and urbanisation (land data only). HadCRUH version one is deliberately created without such corrections for future use as a standard with which to create and compare adjusted data. Bias estimates for marine humidity have already been investigated to some extent by studies of the National Oceanography Centre, Southampton (Kent *et al*., 1999; Kent *et al.*, 1993; Kent & Berry, 2005) and correction estimates produced. Further work to assess the application of such corrections to HadCRUH marine data and investigations for the land data are desirable.

In addition, further research is needed on the potential biases highlighted for the marine data over the course of this project: the 1982 shift; nighttime observational bias; and the shift from predominantly ship to predominantly buoy observations. The 1982 shift has serious implications for users of marine humidity record as it stands, and it is imperative that this outstanding issue be satisfactorily resolved by ascertaining the source and a more accurate (and possibly deck specific) adjustment factor. Although a bias introduced by disproportional sampling across the diurnal cycle is thought unlikely it cannot yet be ruled out. Further analysis and possible bias corrections are desirable. The shift from ship to buoy measurements (Fig.4.1) has lead to significant biases in *SST* and *MAT* observations (Rayner *et al*., 2006; Rayner *et al*., 2003) and as such the same is foreseeable for humidity.

There is much potential for new products developing from HadCRUH. There has been interest in the provision of higher spatial and temporal resolution datasets. As all land data originate from at least six hourly observing frequency and with the addition of more data (discussed above), for specific regions at least, especially Europe, this should be possible. In addition, it is increasingly important to provide uncertainty estimates and dataset builders have gone to great efforts to quantify and represent this (Brohan *et al*, 2006; Thorne *et al.*, 2005a). Following the methods of Brohan *et al*. (2006) such a product should be possible for HadCRUH, taking into account errors in: measurement; homogenisation adjustment; climatology; conversion between humidity variables; sampling density and bias estimates.

The relationship between humidity and human heat stress (sections 1.1, 1.2 and 5.2.4) and the quite feasible updates and resolution improvements to HadCRUH provide a unique and timely opportunity for creating a humidity and temperature based heat stress index. Currently, the UK heat stress index is based on temperature alone whereas the US are already including humidity. If combined with forecasting capabilities and historical epidemiological records, a real time health warning product for both the general public and Health Services (for predicting increases in hospital admissions) could be possible. Investigative work into feasibility and options for such a product will be necessary before any conversion can go ahead. However, in theory, the creation of such a heat stress index of societal benefit, enabling greater understanding of health and climate should be possible.

## 6.5 CONCLUSIONS

HadCRUH has shown that near-surface water vapour is increasing significantly across the globe and that this is highly likely to be primarily of human origin. The dataset represents a useful tool for the climate science community in that it provides a plausible representation of recent changes in surface specific humidity and relative humidity. Hence, it can now be used for: more formal model validation; climate change detection and attribution; and energy budget studies. It will be made available to the general research community through the Climatic Research Unit website and www.hadobs.org. It also has potential as the basis for a product of societal benefit in terms of human heat stress impacts monitoring.

# 6.7 TABLES AND FIGURES FOR CHAPTER 6

| Test | % of humidity observations failing | % of humidity observations failing that pass all other quality control tests |
|---|---|---|
| **Zero *DPD* > 12 hrs** | 2.7 | 2.2 |
| **Repeated *T* value > 12 hrs** | 11.8 | 10.5 |

**Table 6.1: Summary of potential marine data removal for Zero *DPD* and Repeats QC tests.** Tests are implemented on each observing platform with a call sign (identification number) present.

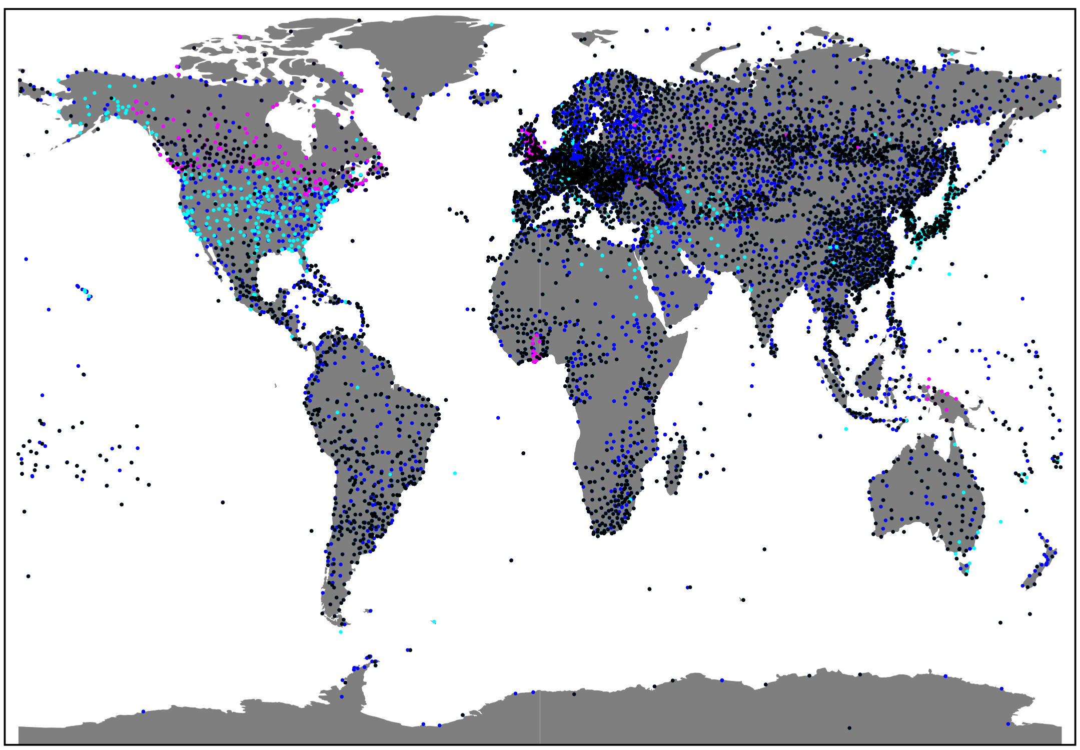

**Figure 6.1: Potential additional data to HadCRUH.** Stations from the ISH dataset are shown in sky blue and black where black dots identify stations that also feature in other datasets. All other non-black dots denote additional data provided by the NEW (pink) and HAHN (royal blue) datasets over the 1973 to 2003 period.