A comparison of large scale changes in surface humidity over land in observations and CMIP3 general circulation models

Katharine M Willett[1,2]**, Philip D Jones**[2]**, Peter W Thorne**[1,3] **and
Nathan P Gillett**[2,4]

[1] Meteorological Office Hadley Centre, Exeter, UK

[2] Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK

[3] Cooperative Institute for Climate and Satellites—NC, NOAA's National Climatic Data Center, Asheville, NC, USA

[4] Canadian Centre for Climate Modelling and Analysis, University of Victoria, Victoria, BC, Canada

E-mail: Kate.Willett@metoffice.gov.uk

Abstract
Observed changes in the HadCRUH global land surface specific humidity and CRUTEM3 surface temperature from 1973 to 1999 are compared to CMIP3 archive climate model simulations with 20th Century forcings. Observed humidity increases are proportionately largest in the Northern Hemisphere, especially in winter. At the largest spatio-temporal scales moistening is close to the Clausius–Clapeyron scaling of the saturated specific humidity ($\sim$7% K^{-1}). At smaller scales in water-limited regions, changes in specific humidity are strongly inversely correlated with total changes in temperature. Conversely, in some regions increases are faster than implied by the Clausius–Clapeyron relation. The range of climate model specific humidity seasonal climatology and variance encompasses the observations. The models also reproduce the magnitude of observed interannual variance over all large regions. Observed and modelled trends and temperature–humidity relationships are comparable except for the extratropical Southern Hemisphere where observations exhibit no trend but models exhibit moistening. This may arise from: long-term biases remaining in the observations; the relative paucity of observational coverage; or common model errors. The overall degree of consistency of anthropogenically forced models with the observations is further evidence for anthropogenic influence on the climate of the late 20th century.

Keywords: specific humidity, GCMs, climate change, Clausius–Clapeyron

1. Introduction

Saturation specific humidity (q_s) is governed by the Clausius–Clapeyron relation. For a 1 K increase in temperature (T) the water holding capacity of the atmosphere increases by $\sim$7%; increasing with decreasing temperature (Sun and Held 1996, Held and Soden 2006). Essentially, if relative humidity (RH) remains constant then specific humidity (q) should also increase with a scaling of approximately 7% K^{-1}. Indeed, quasi-constant RH as the climate changes, at least over large spatial and temporal scales, is often assumed (Manabe and Wetherald 1967) and is an emergent property of general circulation models (GCMs) of the climate system (Held and Soden 2000, Allen and Ingram 2002, Ingram 2002). However, more recently it has come under debate; decreasing surface RH over land has been observed since 2003 (Simmons *et al* 2010).

Actual changes in surface q and to what extent these are governed by changes in T and a quasi-constant RH constraint are key to our understanding of the climate system. Surface humidity is the principal source for the free-troposphere where water vapour has important implications for earth radiation and energy budgets and therefore climate sensitivity (Trenberth *et al* 2005) and the hydrological cycle (Allen and Ingram 2002, Wentz *et al* 2007). At the surface, humidity is an important factor in human thermal comfort (Taylor 2006).

The World Climate Research Programme's (WCRPs) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset affords for the first time an opportunity to evaluate recent changes in surface humidity in a suite of Climate of the 20th Century forced GCMs. This dataset has been used for the analysis of numerous other climate variables (e.g. surface T trends (Stone *et al* 2007); total atmospheric moisture content (Santer *et al* 2007); precipitable water (Mears *et al* 2007); tropical lapse rates (Santer *et al* 2008); sea surface temperature changes in tropical cyclogenesis regions (Santer *et al* 2006a)). In these studies, only the anthropogenically forced models have shown reasonable consistency with the most recent observational datasets.

Following the above studies, this paper expands upon previous analyses of the Hadley Centre and Climatic Research Unit Surface Humidity record HadCRUH (Willett *et al* 2007, 2008—henceforth W07 and W08 respectively; Simmons *et al* 2010) by undertaking a multi-model comparison. It uses the gridded near-global homogenized surface specific humidity product HadCRUH land component (W08) and the Hadley Centre and Climatic Research Unit land surface air temperature product CRUTEM3 (Brohan *et al* 2006). Only land surface data are considered due to issues with marine humidity data pre-1982 (W08). Previously, a dataset of recent changes in surface q and RH was created and assessed (W08) and compared to a single fully coupled GCM, with natural and anthropogenic forcings for the purpose of detection and attribution (W07). Here, firstly recent proportional changes in observed q relative to climatology, along with simultaneous changes in surface T, are assessed. Secondly, a comparison is undertaken of the observed and CMIP3 archive Climate of the 20th Century (forced with historical anthropogenic and natural emissions) mean state and changes in q alongside T. To investigate the effect of spatial completeness on trend analyses, following Simmons *et al* (2010), both full model coverage and coverage sub-sampled to HadCRUH field availability are considered. This study therefore addresses the extent to which RH remains constant over the land surface of the Earth in the observations and the CMIP3 models.

Ideally the model assessment would include both q and RH. However, surface RH is not available for any Climate of the 20th Century runs from the CMIP3 archive. Its derivation from pressure level RH fields makes an inadequate comparison as it is a fundamentally different quantity from observed surface RH. Furthermore, calculating surface RH from monthly mean surface q and T is unsatisfactory owing to non-linearity in the conversion formulae (McCarthy and Willett 2006). It would yield values fundamentally different from the monthly mean of instantaneous RH. Given recent findings regarding changes in surface RH over land (Simmons *et al* 2010) it would be very useful if surface RH were to become a mandatory field in future CMIP archives enabling a surface RH multi-model intercomparison study to be conducted.

Section 2 describes the datasets and models used in the present study. Section 3 describes observed changes in surface q with T across the land surface of the globe. Section 4 compares all findings from the observational data to the full suite of state of the art GCMs from which the relevant parameters are available (q was not a mandatory variable). Section 5 provides a summary.

2. Dataset description

All observed humidity data are from HadCRUH: a quality controlled and homogenized land and marine monthly mean anomaly $5° \times 5°$ gridded dataset of q and RH for the period 1973–2003 (W08). This is based around a 1974–2003 climatology chosen to maximize the number of contributing stations and spatial coverage. The land data originate from the National Climatic Data Center (NCDC) Integrated Surface Dataset (ISD; Lott *et al* 2001). This is a global compilation of station SYNOP reports and high temporal resolution records from various sources. Most CMIP3 Climate of the 20th Century model runs finish in 1999 so in order to compare like with like the observational dataset is curtailed here and renormalized to a 1974–1999 climatology such that the grid-box mean is zero over this period. This period is chosen to both make the climatology period length as close to the widely used standard of 30 years as possible (New *et al* 1999), and maintain consistency with the original observational data as far as possible. The authors strongly recommend that CMIP5 historically forced runs are continued as close to the present day as possible to enable up to date comparisons in the future. The data have been homogenized using a neighbour-based breakpoint detection and adjustment algorithm and only long-term stations are included in the product. Grid-box averages were created by taking a straight average of all stations contributing data in any given month within the grid-box region. No spatial or temporal infilling was performed. Further details are given in W08.

CRUTEM3, the land component of HadCRUT3 (Brohan *et al* 2006 and references therein) is used to provide simultaneous monthly mean anomaly surface T on the same $5° \times 5°$ grids for comparison. The land component differs from HadCRUH in source datasets (monthly summaries of dry-bulb temperature versus synoptic level dewpoint temperature reports), although many observing stations are common. Methods for quality control and homogenization also differ substantially such that, over land at least, the two datasets can be considered structurally independent (Thorne *et al* 2005). CRUTEM3 data for the 1973–1999 period are extracted and renormalized to the climatology period of 1974–1999. The more globally complete CRUTEM3 dataset is sub-sampled to HadCRUH field availability. Both HadCRUH (blended land and marine) and CRUTEM3 are freely available at www. hadobs.org.

Table 1. CMIP3 GCMs providing surface specific humidity and surface temperature. Abbreviations are as follows: (GHG) well mixed greenhouse gases; (OZ) tropospheric and stratospheric ozone; (SD) sulfate aerosol direct effects; (SI) sulfate aerosol indirect effects; (BC) black carbon; (OC) organic carbon; (MM) mineral dust; (SS) sea salt; (LC) land use change; (SO) solar irradiance; and (V) volcanic aerosols Source: Santer *et al* (2006b).

Model	Description	Members	Forcings
BCCR	BCCR-BCM2.0—Bjerknes Centre for Climate Research (BCCR), Univ. of Bergen, Norway	1	GHG, SD
CCCMAt47, CCCMA t63	CGCM3.1(T47) and CGCM3.1(T63)—Canadian Centre for Climate Modelling and Analysis, Univ. of Victoria, Canada	5 t47 runs and 1 t63 run	GHGs, SD
CNRM	CNRM-CM3—Center for National Weather Research, Meteo-France, Toulouse, France	1	GHG, OZ, SD, BC
CSIRO3.0, CSIRO3.5	CSIRO-Mk3.0 and CSIRO-Mk3.5—Commonwealth Scientific and Industrial Research Organization, Australia	2 3.0 runs and 3 3.5 runs	GHG, SD
GISSAOM, GISSEH, GISSER	GISS-AOM, GISS-EH, GISS-ER—NASA Goddard Institute for Space Studies, Colombia Univ. NY, USA	2 AOM runs, 5 EH runs and 9 ER runs	AOM: GHG, SD, SS, EH and ER: GHG, OZ, SD, SI, BC, OC, MD, SS, LC, SO, V
INMCM	INM-CM3.0—Institute for Numerical Mathematics, Russia	1	GHG, SD, SO
IPSL	IPSL-CM4—Sciences de l'environnement, Institut Pierre Simon Laplace, France	2	GHG, SD, SI
MIROCH, MIROCM	MIROC3.2(hires), MIROC3.2(medres)—Center for Climate System Research (The University of Tokyo); National Institute for Environmental Studies; Frontier Research Center for Global Change (JAMSTEC), Japan	1 hires, 3 medres	GHG, OZ, SD, SI, BC, OC, MD, SS, LC, SO, V
MRI	MRI-CGCM2.3.2—Meteorological Research Institute, Ibaraki, Japan	5	GHG, SD, SO, V
NCAR	CCSM3—National Center for Atmospheric Research, Boulder, USA	8	GHG, OZ, SD, BC, OC, SO, V

All CMIP3 archive models (Meehl *et al* 2007) with surface specific humidity (huss) and surface temperature (tas) available for the 'Climate of the 20th Century' forcing scenario are obtained from www-pcmdi.llnl.gov/. The IAP/LASG FGOALS_g1.0 runs have not been included following discovery of an addendum submitted to the CMIP3 data portal (www-pcmdi.llnl.gov/ipcc/model_documentation/more_info_iap_fgoals.pdf) with a disclaimer recommending that IAP is not used for mid- to high-latitude climate studies. This notes cold biases in the tropical Pacific, an overly strong ENSO, overestimated sea ice extent in the high-latitudes and a weakened Atlantic Meridional Overturning Circulation. A revised version is now referred to in the addendum but has not been included here. This results in 15 models and a total of 49 ensemble members. These include greenhouse gas and aerosol forcings in all cases, and in some cases volcanic, solar and other forcings (table 1). While the natural variability of the climate system e.g. ENSO does not align with observed timings and magnitude, anthropogenic, volcanic and solar forcings are chronologically aligned to those experienced historically. For this reason it is important that model and observed dataset periods and climatologies are identical. All runs are retrieved

as monthly mean values from 1973 to 1999 (December) and are regridded to match HadCRUH resolution ($5° \times 5°$) by simple averaging over each grid-box. This technique is chosen to cope with the irregular gridding at high-latitudes in some models and while not ideal should be sufficient for the resolution of analyses made herein. Model data are resampled to match HadCRUH spatio-temporal availability and then anomalies are created from a 1974 to 1999 climatology for each grid-box.

3. Recent changes in observed surface humidity from 1973 to 1999

The largest magnitude changes (in g kg^{-1}) in *q* have previously been found over the Tropics and summer hemisphere (Dai 2006; W08). Here it is shown that proportional changes (% relative to climatology) are largest in the Northern Hemisphere in winter (December, January and February—DJF (figures 1(a) and (b))). At the grid-box level these increases can be very large ($>20\%$ between 1973 and 1999), particularly in continental interiors.

Table 2. Large scale changes in specific humidity and temperature between 1973 and 1999. Changes are calculated by fitting trends to the data using the median of pairwise slopes technique (Sen 1968, Lanzante 1996). Error bars are obtained from the 90th percentile confidence intervals around the median of pairwise slopes. Confidence intervals for changes in specific humidity with temperature are created by combining the individual confidence intervals following Stuart and Ord (1987). Values in bold are significantly different from zero at the 90% level of confidence.

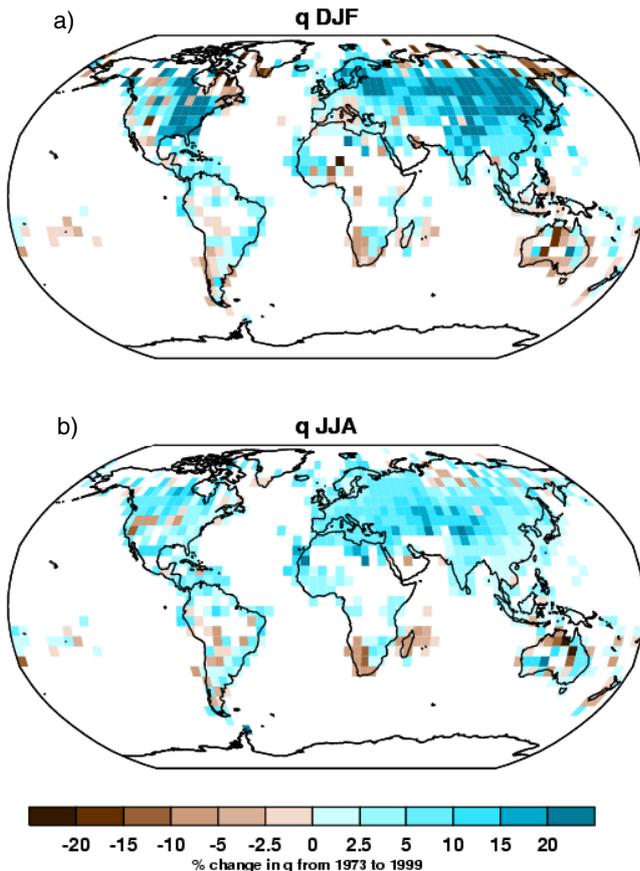| Region	Total change in q (%)	Total change in T (°C)	Change in q per °C increase in T (%)	Correlation between T and q
Globe (70°S–70°N)	**4.11 ± 0.78**	**0.56 ± 0.09**	**7.35 ± 0.97**	0.72
Northern Hemisphere (20°N–70°N)	**5.98 ± 1.18**	**0.68 ± 0.15**	**8.81 ± 1.46**	0.67
Tropics (20°S–20°N)	**2.45 ± 0.63**	**0.45 ± 0.08**	**5.45 ± 0.68**	0.90
Southern Hemisphere (20°S–70°S)	−0.27 ± 1.08	**0.31 ± 0.10**	−0.85 ± 3.56	0.42

a)

b)

Figure 1. Seasonal changes in surface specific humidity from 1973 to 1999 shown as total change over the period as a percentage relative to the 1974–1999 climatology. Changes are calculated based on trend fitting using the median of pairwise slopes technique (Sen 1968, Lanzante 1996).

Large scale averages over all seasons show the greatest q increase in the extratropical Northern Hemisphere at 5.98 ± 1.18% (table 2), coincident with the largest increase in T. For the Tropics and Globe, increases in q are 2.45 ± 0.63% and 4.11 ± 0.78% respectively. There is no significant change in the extratropical Southern Hemisphere; there is evident drying over the arid areas in this region (figures 1(a) and (b)). There is significant warming in all regions in CRUTEM3; although this warming is smallest in the extratropical Southern Hemisphere. The large gaps in data, and lower station density over Central Africa, the Amazon and Central Australia increase

the uncertainty in the Southern Hemisphere (Tropics and extratropics) observations of both T and q.

As T increases so does the water holding capacity of the atmosphere (q_s). Hence, q also increases where evaporation is not limited by water availability and where atmospheric conditions are conducive to evaporation in terms of windiness. It is expected that q_s will increase by ∼7% K^{-1} following Clausius–Clapeyron theory and therefore in regions where RH remains constant we may expect to observe similar scale increases in q. Indeed, such changes have been observed by remote sensing of column water vapour (Mears *et al* 2007, Trenberth *et al* 2005, Trenberth 2007).

For HadCRUH, this is explored in more detail using 23 of the Giorgi and Francisco (2000) regions (table 3) which were used in slightly modified form in the IPCC 4th Assessment (Christensen *et al* 2007). For all regions monthly mean T and q correlate positively; for 17 of the 23 regions the correlation r value is greater than 0.5 (table 3). Fifteen regions exhibit significantly increasing q consistent with quasi-constant RH (figure 2, table 3) in that trends lie between 6 and 8% K^{-1} taking into account confidence intervals. Thus there is support in the observations for Clausius–Clapeyron scale increases (∼7% K^{-1}) in q over most but not all regions over the 1973–1999 period. The Southern Hemisphere regions of South Africa, Northern Australia and Southern Australia show an inverse relationship with T in terms of total change but these are not significant. These are typically desert regions where water, and hence evaporation, is limited. Simmons *et al* (2010) point to faster warming over land relative to the oceans as a causal mechanism for the observed very recent drying (outside our period) in relative terms (decreasing RH). Thus, while being water-limited may explain slower than Clausius–Clapeyron scaling to some extent, it is unlikely to be the full story and further research is needed. Three regions (the Tibetan Plateau, Southern Asia and the Caribbean) exhibit q increases which significantly exceed 8% K^{-1}. This implies local increases in RH in these regions.

Over larger scales (table 2), T–q correlations greater than 0.5 are found in all regions except for the extratropical Southern Hemisphere. The strongest correlation is found over the Tropics ($r = 0.9$), a warm and moisture-abundant region. RH and T are not correlated at monthly mean anomaly resolution (not shown). Increases in q scaling with T are both significant and consistent with 7% K^{-1} for the Globe (7.35 ± 0.97% K^{-1}). Scaling is significantly positive
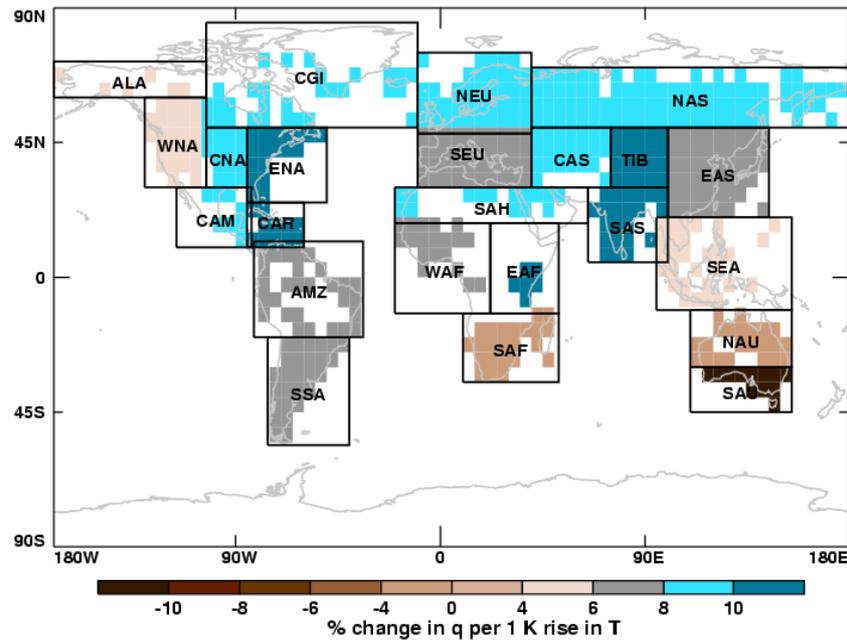
Figure 2. Regional changes in surface specific humidity relative to changes in temperature from 1973 to 1999 shown as the percentage of specific humidity relative to climatology (1974–1999) per 1 K rise in temperature (% K^{-1}). Changes are calculated using trends fitted by the median of pairwise slopes technique (Sen 1968, Lanzante 1996). Regions are based on Giorgi and Francisco (2000). They are described along with total changes in specific humidity (%) and temperature (°C), changes in specific humidity per 1 K change in temperature and monthly mean anomaly timeseries correlations of temperature and specific humidity in table 3.

Table 3. Regional scale changes in specific humidity and temperature from 1973 to 1999. Changes and confidence intervals are calculated as described in table 2. Values in bold are significantly different from zero at 90% confidence. Italicized values are also consistent with Clausius–Clapeyron scaling in that within the confidence intervals they lie between 6 and 8 % K^{-1} scaling. The spread from 6 to 8% is chosen to reflect that the 7% K^{-1} scaling is only an approximation.

Region acronym	Region description	Total change in q (%)	Total change in T (°C)	Change in q per °C increase in T (%)	Correlation between T and q
ALA	Alaska	**6.13 ± 5.04**	**1.22 ± 0.60**	*5.02 ± 2.09*	0.76
CGI	Canada, Greenland and Iceland	**4.86 ± 3.19**	**0.52 ± 0.36**	*9.32 ± 2.70*	0.73
WNA	Western North America	**3.88 ± 2.90**	**0.84 ± 0.42**	*4.61 ± 1.84*	0.72
CNA	Central North America	1.87 ± 4.06	0.22 ± 0.48	8.33 ± 8.82	0.68
ENA	Eastern North America	**3.59 ± 2.66**	**0.33 ± 0.25**	*10.85 ± 4.59*	0.79
CAM	Central America	**3.74 ± 1.78**	**0.41 ± 0.16**	*9.10 ± 4.10*	0.58
AMZ	Amazon	**3.20 ± 0.65**	**0.51 ± 0.11**	*6.31 ± 0.80*	0.83
SSA	Southern South America	**2.23 ± 2.05**	**0.35 ± 0.21**	*6.37 ± 4.01*	0.67
NEU	Northern Europe	**5.14 ± 3.13**	**0.53 ± 0.32**	*9.66 ± 2.63*	0.87
SEU	Southern Europe	**5.34 ± 1.78**	**0.69 ± 0.19**	*7.74 ± 1.40*	0.85
SAH	Sahel	**8.74 ± 2.42**	**0.89 ± 0.19**	*9.76 ± 2.43*	0.57
WAF	Western Africa	**4.71 ± 1.87**	**0.68 ± 0.18**	*6.96 ± 2.61*	0.44
EAF	Eastern Africa	**3.03 ± 1.40**	0.14 ± 0.14	*21.96 ± 21.85*	0.37
SAF	Southern Africa	−1.52 ± 1.80	**0.64 ± 0.14**	−2.37 ± 2.93	0.17
NAS	Northern Asia	**7.73 ± 3.70**	**0.80 ± 0.46**	*9.68 ± 1.90*	0.74
CAS	Central Asia	**5.70 ± 3.00**	**0.61 ± 0.47**	*9.38 ± 4.23*	0.67
TIB	Tibetan Plateau	**10.88 ± 3.19**	**0.69 ± 0.38**	15.77 ± 6.02	0.49
EAS	Eastern Asia	**5.91 ± 2.26**	**0.75 ± 0.23**	*7.87 ± 1.81*	0.66
SAS	Southern Asia	**4.97 ± 1.22**	**0.43 ± 0.14**	11.62 ± 3.33	0.56
SEA	South East Asia	**1.84 ± 0.71**	**0.42 ± 0.11**	*4.38 ± 0.95*	0.85
NAU	Northern Australia	−1.02 ± 3.18	**0.61 ± 0.22**	−1.68 ± 5.40	0.21
SAU	Southern Australia	**−3.18 ± 2.25**	0.10 ± 0.22	−31.9 ± 84.78	0.50
CAR	Caribbean	**4.59 ± 0.84**	**0.40 ± 0.11**	11.58 ± 2.18	0.80

and slightly higher in the extratropical Northern Hemisphere (8.81 ± 1.46% K^{-1}) but lower in the Tropics (5.45 ± 0.68% K^{-1}). Extratropical Southern Hemisphere scaling is small and negative but not significant. Unfortunately, over the

Southern Hemisphere especially, HadCRUH and CRUTEM3 (after HadCRUH resampling) omit key regions of hydrological extremes such as Central Africa and much of Amazonia. This suggests that this region may not be well represented in the

Table 4. Fractions of grid-box climatological mean (1974–1999) and σ of specific humidity in the first ensemble members for each model within 2σ of HadCRUH climatology and within 20% of HadCRUH σ respectively. Members having >66.6% of grid-boxes within these ranges are in bold. Ensemble members are considered dry/moist/high/low if >66.6% of those grid-boxes outside these ranges are dry/moist/high/low.

Model	Grid-boxes within 2σ of observed climatology (%)		Grid-boxes within 20% of observed σ		Bias relative to the observations JJA/DJF	
					Mean (D = dry, M = moist, — = no bias)	σ (H = high, L = low, — = no bias)
	JJA	DJF	JJA	DJF		
BCCR	63.0	**68.9**	39.0	39.7	D/D	—/L
CCCMAt47	58.5	**68.2**	52.4	50.7	M/M	L/L
CCCMAt63	62.2	**70.6**	51.5	39.4	—/—	L/L
CNRM	53.5	**73.8**	46.0	47.5	M/—	H/—
CSIRO3.0	**66.6**	**68.9**	27.3	33.5	D/—	—/—
CSIRO3.5	36.4	47.2	38.3	37.7	M/M	H/H
GISSAOM	64.0	55.9	38.7	33.5	—/M	L/—
GISSEH	52.4	64.8	39.6	33.0	—/—	—/L
GISSER	46.6	**68.8**	36.1	36.2	D/—	—/L
INMCM	54.0	53.9	39.3	34.2	D/—	H/H
IPSL	63.0	**69.7**	50.4	44.4	D/—	—/L
MIROCH	46.2	64.7	48.2	47.2	M/M	L/L
MIROCM	47.7	**70.2**	37.8	43.7	M/—	—/L
MRI	**70.3**	**78.7**	47.1	45.2	D/—	L/L
NCAR	52.6	**71.8**	44.2	37.7	D/—	—/—

observations. Even where there is coverage the station density is generally lower and therefore fundamental data quality is more questionable.

4. How are surface specific humidity and temperature related processes represented by the CMIP3 multi-model archive with 'Climate of the 20th Century' forcings?

To date there has been no multi-model comparison of surface humidity with available observations. Here, an investigation into how well current models represent present day surface q and the T–q relationship is undertaken. The primary aim is to explore the models' ability to capture the mean state (climatological mean and interannual variability) and trend behaviour apparent in the observations. Additionally, this will test to some extent whether the surface humidity signal, which is very likely at least in part of anthropogenic origin (W07), is consistent across a wide spread of model physics, forcings and climate sensitivities.

4.1. Climatological means and standard deviations of seasonal specific humidity: HadCRUH verses CMIP3 models

The climate mean state is assessed over two seasons: DJF (December, January and February); and JJA (June, July and August). While forcings in the models will be aligned in real time with the observations, natural variability is not. It is assumed that the use of long-term averages (1974–1999) will ameliorate differences arising from this mismatched natural variability to a large extent. Observed climatological means range from 0.05 g kg^{-1} at the Poles to 21.95 g kg^{-1} in the Tropics (figures 3(a) and (b)). To broadly assess model-observed differences (without attributing

any formal statistical significance) the percentage of grid-boxes in each first ensemble member falling within observational natural variability (seasonal climatology $\pm 2\sigma$ of the observed interannual seasonal variability—figures 5(a) and (b) for comparison) is calculated (table 4, figures 3(c)–(h), 4(a)–(f)). The models are in general agreement with the observations: 11 (of 15) show greater than 50% agreement with the observations in both seasons; MRI and CSIRO3.0 in particular. The remaining grid-boxes are then apportioned to being either biased moist or biased dry and the model classified as moist/dry if greater than 2/3 of remaining grid-boxes fit into the moist/dry category. In this way, BCCR run 1 is relatively dry and CSIRO3.5 run 1 is relatively moist. To represent the dry to moist spread of models INMCM, MRI and CSIRO3.5 are shown respectively in figures 3(c)–(h) (figures 5(c)–(h) for variability comparison). INMCM and CSIRO3.5 have a relatively low proportion of grid-boxes in 'agreement' with the observations. There is a slight tendency towards drier JJA in the models. However, given the spread of moisture regimes we conclude that there is no overall bias in the models but that agreement is closest and bias least during DJF. As shown in figures 4(a)–(f) there are regions where models show common biases: most noticeably there is a propensity for overly dry conditions across the Amazon and Central Asia during JJA.

Observed interannual seasonal variability is smallest in the winter hemisphere, especially in DJF (minimum = 0.03 g kg^{-1}) (figures 5(a) and (b)). The extratropical summer hemisphere variability is higher, likewise especially in DJF (maximum = 2.28 g kg^{-1}). Given higher summertime temperatures, this should be expected if the mean and variability of RH are largely seasonally invariant. To crudely identify model-observed differences (again without attributing any formal statistical significance) the percentage of grid-boxes falling within observational natural variability (defined
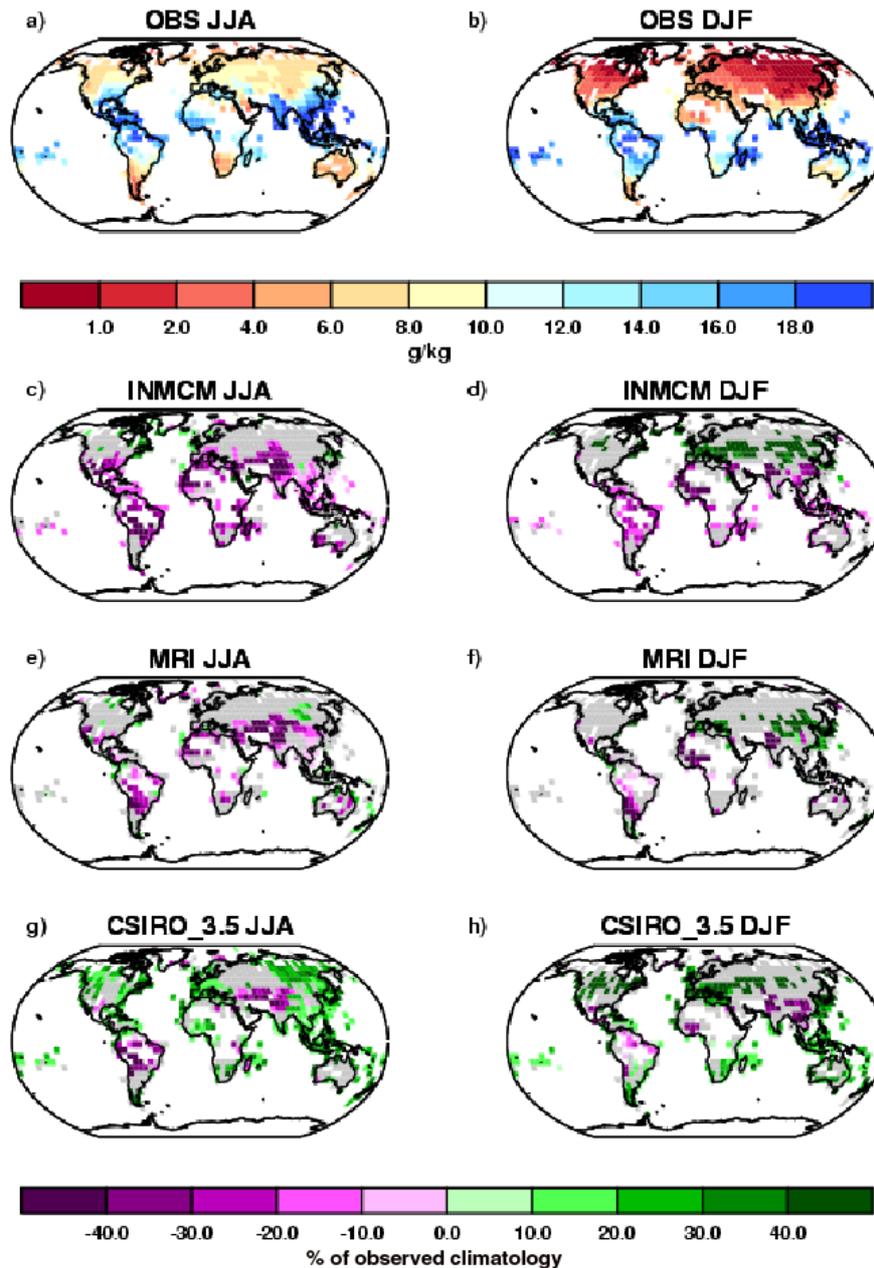
Figure 3. Comparing the seasonal 1974–1999 climatological means in HadCRUH observed surface specific humidity and a range of CMIP3 archive first ensemble members. Observations are shown at the top in g kg^{-1} with colour bar. Three ensemble members of the 15 CMIP3 models with surface specific humidity available are shown in the lower panels. INMCM is a particularly dry model, MRI is moderate within the spread of models and CSIRO3.5 is a moist model. These are shown as model—observed differences in terms of the percentage relative to the observed climatology. Greyed out areas identify regions where the model climatologies are within 2σ of the observed interannual seasonal variability. Statistics for all model first ensemble members are listed in table 4.

as ±20% of the standard deviation for each grid-box) are calculated for each first ensemble member by season (table 4, figures 5(c)–(h), 6(a)–(f)). Agreement assigned in this way is generally low with only one model exceeding 50% grid-box agreement for both seasons. All show more than 30% with the exception of CSIRO3.0 which interestingly agrees well in terms of climatology. Models are categorized as high or low variability relative to the observations by looking for a 2/3 proportion of remaining grid-boxes (as for moist/dry categorization above). For variability, unlike climatology, it is

apparent that the models are closer to the observed variability in the JJA season. There is a low variability tendency in many of the models relative to HadCRUH but CNRM, CSIRO3.5 and INMCM exhibit higher variability. Given that CSIRO3.5 is biased moist and INMCM is biased dry there is no clear link demonstrated between biases in the mean and variability relative to the observations. Variability over Southern and Eastern Asia is biased high in the majority of models over JJA especially. There is a common low bias over the Caribbean in both seasons. To what extent the models can
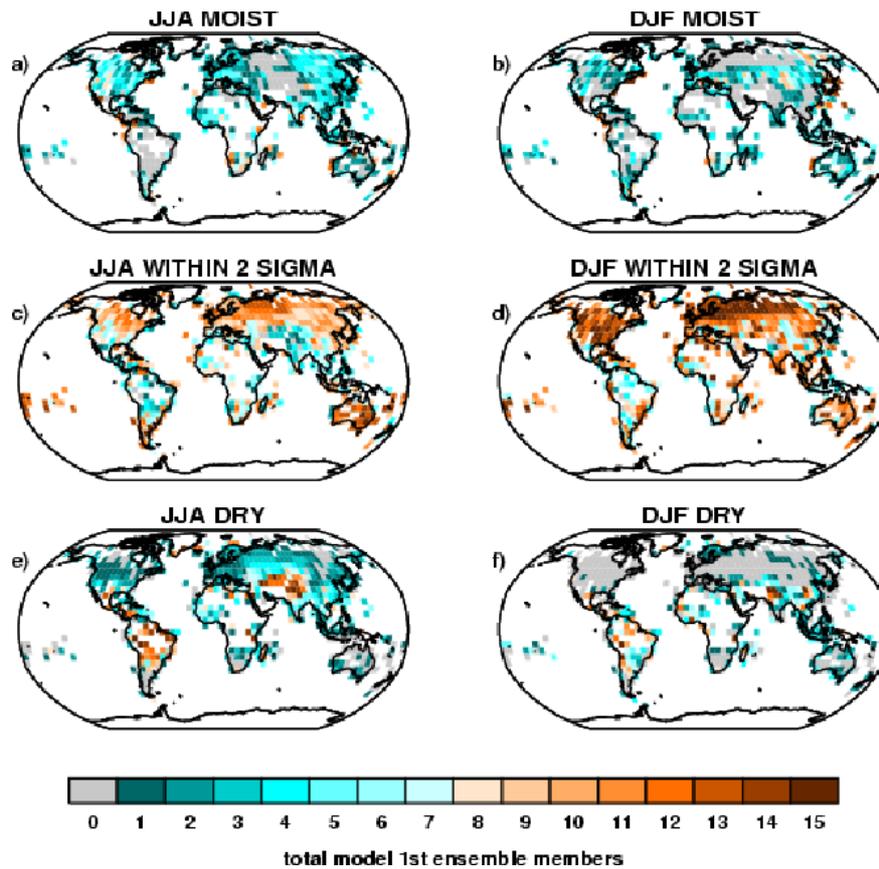
Figure 4. Summary of all model surface specific humidity seasonal climatologies (1974–1999) compared to the observations showing the number of model first ensemble members within each category: moister than the observations, within 2σ interannual variability of the observations and drier than the observations.

be expected to accurately reconstruct observational variance at the $5° \times 5°$ grid-box scale is debatable, especially in regions of inhomogeneous topography. Observations can originate from topographically heterogeneous stations and may not capture the entire variability of climate over the grid-box resulting in random sampling errors and inflated variance estimates. Conversely CMIP3 models resolve physical equations over a variety of different grid-box scales (often finer resolution than the observations) with a single topographical realization. Despite these considerations, we conclude that the models exhibit variability of the same order of magnitude as seen in the observations.

4.2. Specific humidity and temperature processes: HadCRUH verses CMIP3 models

For the Globe, extratropical Northern Hemisphere and Tropics, the model spread of trends of surface q encompasses the observed increasing trends (figures 7(a)–(d)). All model runs for these three regions show positive trends. For the extratropical Northern Hemisphere the observed trend is at the higher end of the model spread. Model spread over the Tropics is larger than for any other region. This could be due to proportionally larger effects from ENSO variability and higher interannual variance over this region. However, to conclude this it should follow that the spread of trends

within model ensembles should be greater here than for other regions but it is not (figures 7 and 8). This suggests that model uncertainty over the Tropics is a more likely reason. In the extratropical Southern Hemisphere there is essentially no trend in the observations but all model runs (except CSIRO3.5 run 3) show positive trends of a similar magnitude to those in the extratropical Northern Hemisphere. Given data sparsity and the larger observational uncertainty in this region (W08) it is conceivable that the models could lie closer to reality than the observations imply. Interannual variability in the large scale zonally averaged observed timeseries lies close to that of the multi-model mean and is encompassed by the spread of model variability. In analysing the timeseries it is important to note that not all models include volcanic forcings, only GISSER, GISSEH, MIROCM, MIROCH, MRI and NCAR (see table 1). In addition to the lack of ability to simulate ENSO event phasing and magnitudes precisely as they occurred in the real-world this leads to differences between the individual GCM runs and observed timeseries.

For the most part, large scale regional T–q relationships in the models agree well with the observations (figures 8(a)–(d)). Correlations between T and q timeseries are highest in the Tropics in both the models and observations, followed by the Globe and then the extratropical Northern Hemisphere. For all regions the majority of model runs show increasing q with
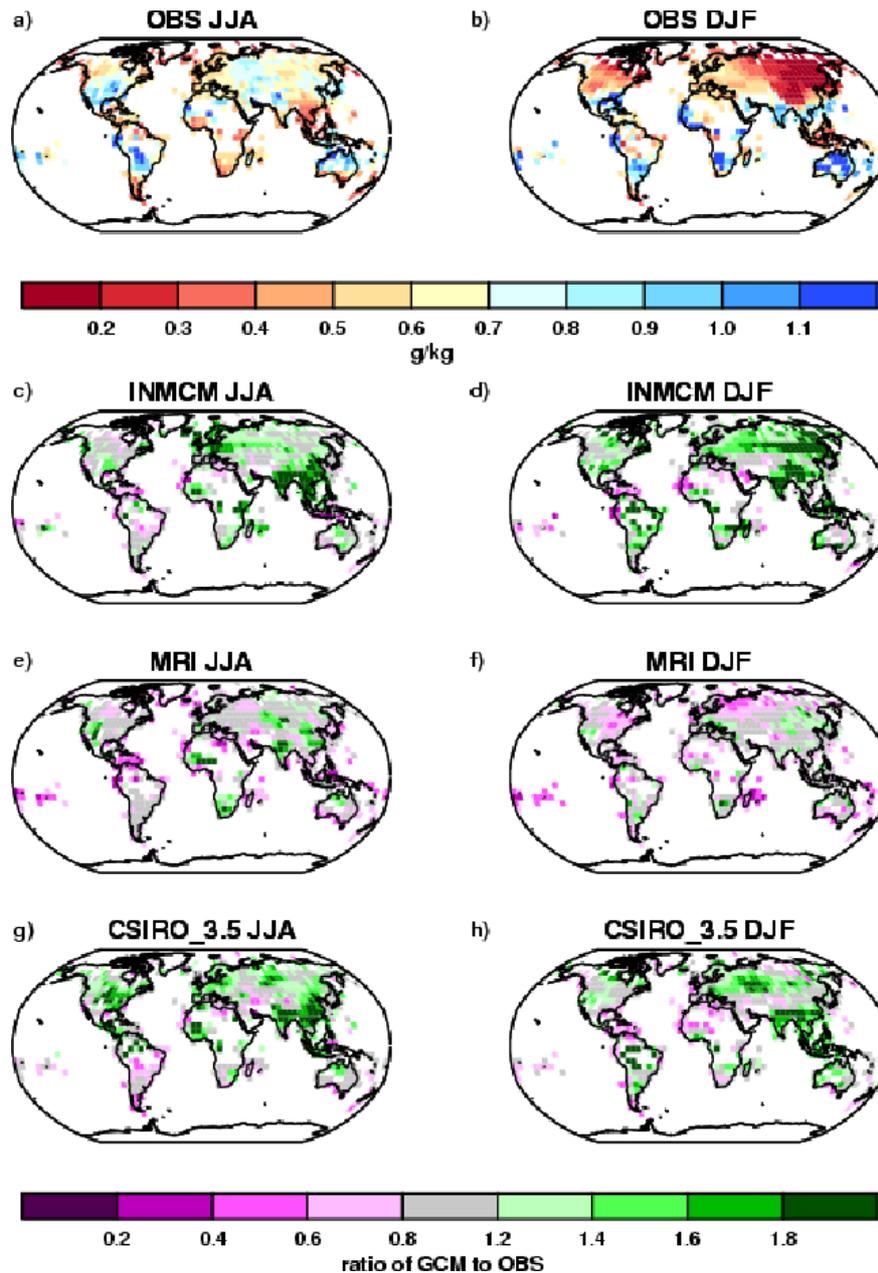
Figure 5. Comparing the seasonal 1974–1999 standard deviation in HadCRUH observed surface specific humidity and a range of CMIP3 archive first ensemble members. Observations are shown at the top in g kg^{-1} with colour bar. Three first ensemble members of the 15 CMIP3 models with surface specific humidity available are shown in the lower panels. INMCM is a particularly dry model, MRI is moderate within the spread of models and CSIRO3.5 is a moist model. These are shown as model—observed differences in terms of the ratio relative to the observed standard deviation. Greyed out areas identify regions where the models are within 20% (0.8–1.2) of the observed seasonal variability. Statistics for all first ensemble members are listed in table 4.

T close to Clausius–Clapeyron scaling of q_s with T with multi-model means consistently lying between 6 and 7% K^{-1}. There is much greater model spread in the extratropical Southern Hemisphere. For total change in q the spread is large in all regions. NCAR (run 7) and CSIRO3.5 (run 3) are strong outliers in the extratropical Southern Hemisphere. Models envelop the observed q changes proportionally and with T for the Tropics and Globe but show poorer agreement in the extratropical Northern Hemisphere and very poor agreement in the extratropical Southern Hemisphere (figures 8(a)–(d)).

For the Northern Hemisphere it is not possible to say whether the observations are showing too large an increase or the models too small—both have implications for changes in RH over large scales. All but CSIRO3.5 (run 3) show larger T–q scaling than the observations in the extratropical Southern Hemisphere. This could be due to: long-term biases remaining in the observations; the relative paucity of observational coverage; common model errors or a combination of these. The large model spread in this region could point to more model uncertainty in the extratropical Southern Hemisphere
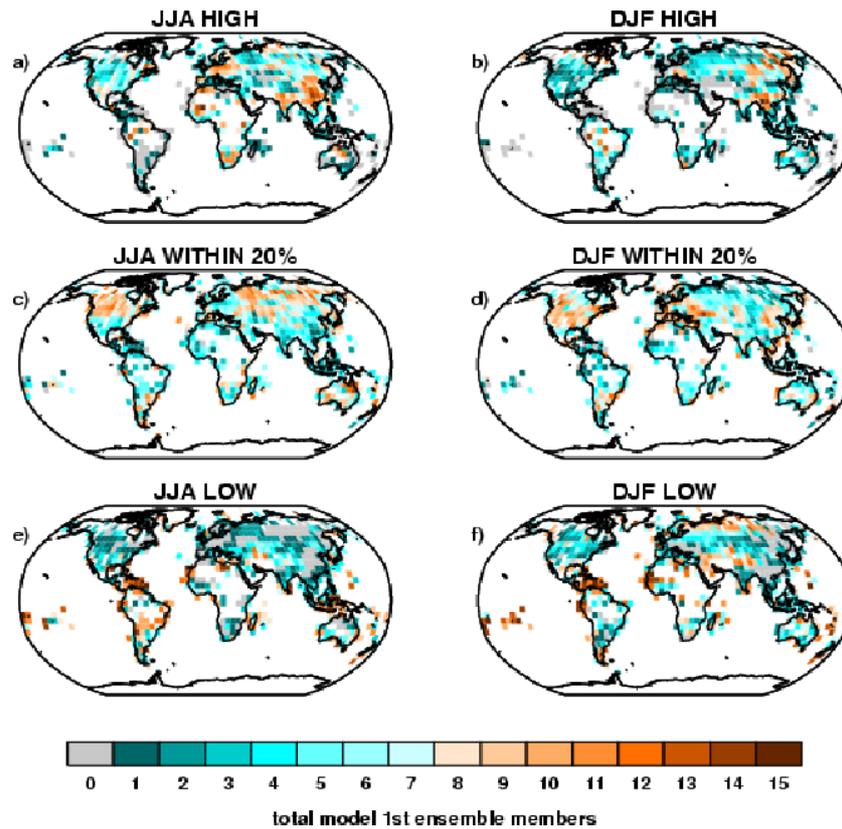
Figure 6. Summary of all model surface specific humidity interannual seasonal standard deviations (1974–1999) compared to the observations showing the number of model first ensemble members within each category: higher variability than the observations, within 20% of the variability of the observations and lower variability than the observations.

than in the other regions but could also be due to sparse data sampling (from matching spatially and temporally to the observed data coverage).

Incomplete spatial coverage of observational datasets is a common problem leading ultimately to uncertainty in terms of accurate representation of large scale averages. Indeed, Simmons *et al* (2010) found strong agreement between the ERA interim reanalyses and CRUTEM3 when spatially matched but larger trends in the spatially complete ERA interim. Notably there remains considerable uncertainty in humidity trends from models and reanalyses over regions where there are no observations with which to validate. To address this issue for HadCRUH, figure 8 has been repeated with full spatial coverage in the models over land. Changes in the rate of change of *q* with *T* are very small for all regions— from −0.03 in the Tropics to +0.14 for the Globe. All regions show small increases of ∼0.5% in the total change in *q* with the smallest change (0.41%) in the extratropical Southern Hemisphere and largest change (0.61%) in the Tropics. There is no systematic shift by ensemble members towards or away from the observations. If we accept that the models are making a satisfactory reconstruction of the observations and that the observations are reasonable this suggests that for *q* filling in the missing data, at least over the latitudinal extent studied here (70°N and 70°S) and period of study (1973–1999), will have little implication for the large scale features observed for the Globe, Tropics and Northern Hemisphere. As with Simmons

et al (2010), there will likely be some difference in the exact magnitude of changes.

5. Discussion

Increases in seasonal surface specific humidity at the largest scales are significant and broadly consistent with Clausius–Clapeyron scaling of temperature trends given constant relative humidity. Some smaller regions, however, have moistened at a greater rate or—south of the equator—become drier, though these latter results may be sensitive to local data coverage and quality issues.

The CMIP3 multi-model archive 'Climate of the 20th Century' mean surface specific humidities are broadly consistent with the observed climatology with no overall moist or dry bias. Interannual seasonal variability in the models tends to be biased slightly low relative to the observations except for CNRM, CSIRO3.5 and INMCM. In part this may be an expected result as it is unlikely that at the true observed grid-box scale variability will be adequately reproduced by either. The observations are likely to over-estimate grid-box variability through under-sampling, especially in poor station density regions where regression to the mean is inhibited and there is therefore more noise. Models may be more likely to under-estimate grid-box interannual variability in some regions due to coarsely resolved topography and land use/land cover (the latter is not included in all models). Regardless,
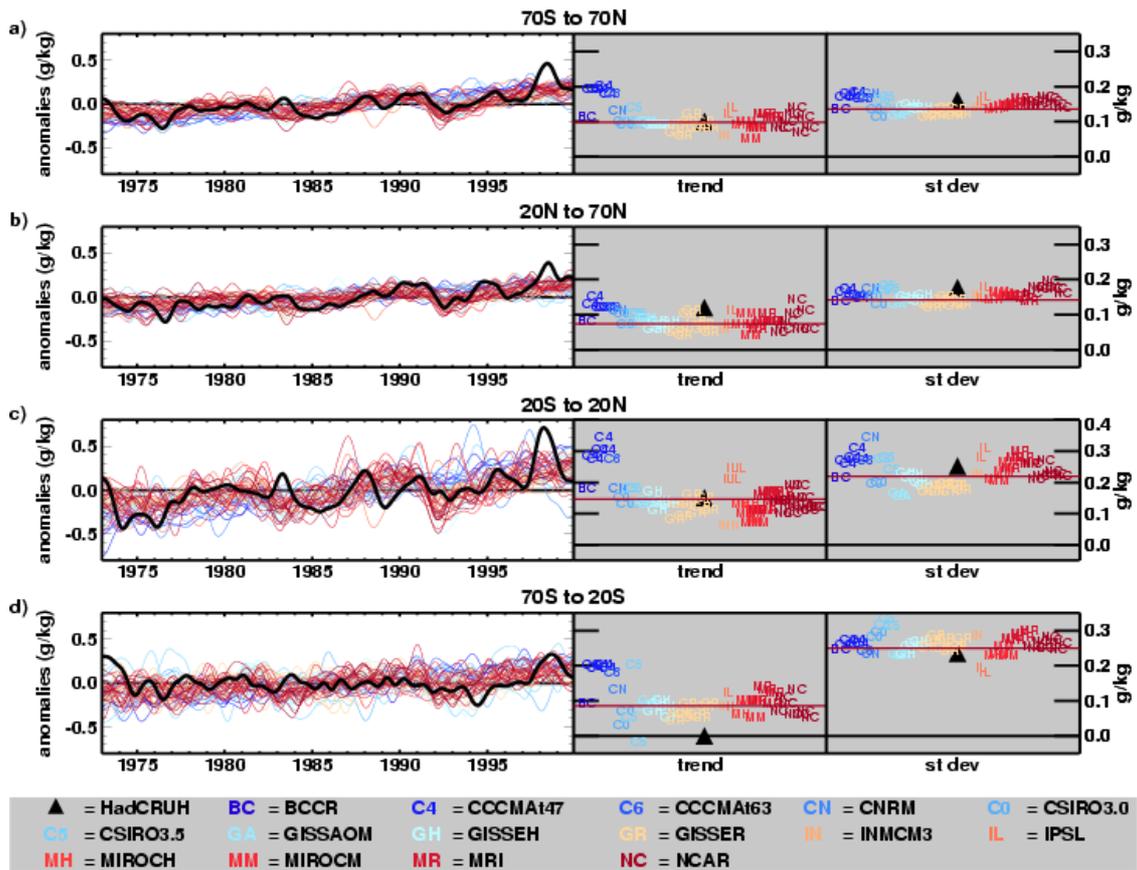
Figure 7. Comparing large scale HadCRUH observed surface specific humidity timeseries and trends with individual members of 15 CMIP3 models sub-sampled to HadCRUH coverage from 1973 to 1999 over land. Regional averages are created using the cosine of latitude weighting. Timeseries are monthly mean anomalies smoothed using a 21 point filter. Observations are in black and models are coloured in order from blue to red (see table 1). Decadal trends and standard deviation are shown on the right-hand side, again with observations in black. Trends are fitted using the median of pairwise slopes technique (Sen 1968, Lanzante 1996).

the models' interannual variability is of the right order of magnitude across a range of spatial scales.

When averaged over large scales, models capture the historical timeseries satisfactorily in terms of the variability for all regions and trend sign and magnitude for all except the extratropical Southern Hemisphere. Although models are sub-sampled to match observational data coverage, within grid-box station sampling of the observations is also sparse in the Southern Hemisphere and so to some extent may explain some of this inconsistency. All models concur on positive trends for all regions with the exception of one run in the Southern Hemisphere. Models show close to 7% K^{-1} scaling in all large regions with little difference between them. This captures the temperature–specific humidity relationship found in observations for the Globe and Tropics and is close to but smaller by $\sim$3% K^{-1} than that observed in the extratropical Northern Hemisphere. There is no agreement between all models and observed behaviour in the extratropical Southern Hemisphere.

Complete geographical sampling makes little difference to multi-model means overall. This suggests that where the observations and models concur (Globe, extratropical Northern Hemisphere and Tropics) the observational sampling is sufficient to capture the main features of recent changes.

In the extratropical Southern Hemisphere the observations remain inconsistent with the models whether sub-sampled or fully sampled. This region also shows larger model uncertainty in the temperature–specific humidity relationship. It is unfortunate that due to poor data coverage here the observational uncertainty is also high. Further work is needed to establish whether the shortcomings are really from the observations, the models or perhaps more likely, both.

We have shown herein that the climate models in the CMIP3 archive exhibit reasonable mean state (climatological means and interannual variability) specific humidity characteristics compared to the observations. The spread of their Climate of the 20th Century runs encapsulates the observations for recent changes in the global mean, tropical mean and Northern Hemisphere extratropical mean. This does not hold for the Southern Hemisphere extratropics. This discrepancy could relate to: residual observational error, spatial representivity, or common model errors. Although uncertainty in the observations is largest here there is insufficient evidence to formally discriminate between these factors. Notwithstanding the Southern Hemisphere extratropics this study indicates the strong agreement between anthropogenically forced GCMs and statistically significant changes observed in surface specific humidity. Hence, it supports findings from earlier single
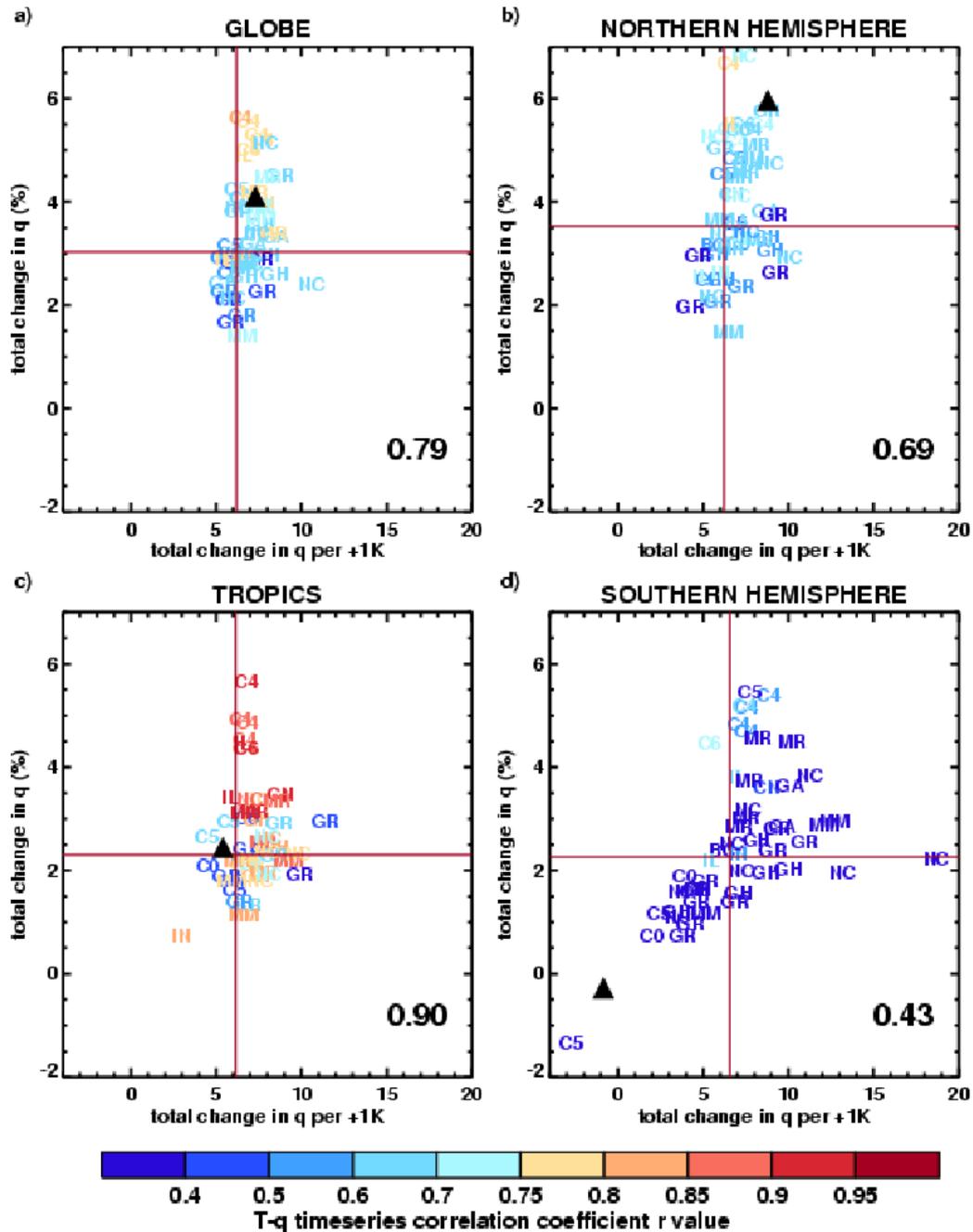
Figure 8. Comparing large scale changes in specific humidity relative to temperature over land from 1973 to 1999 from the HadCRUH and CRUTEM3 observations and 15 CMIP3 models sub-sampled to HadCRUH coverage. Regional averages are created using cosine of latitude weighting. Observations are shown as black triangles with the correlation *r* between observed specific humidity and temperature monthly mean anomaly timeseries shown in the bottom right corner. Model members are shown by name labels (described in figure 7) and colour coded by their specific humidity–temperature timeseries correlation *r*. Changes are calculated from trends fitted using the median of pairwise slopes technique (Sen 1968, Lanzante 1996).

model studies of the likely presence of an anthropogenic signal in specific humidity observations from the latter part of the 20th Century.

Acknowledgments

References

Allen M R and Ingram W J 2002 Constraints on future changes in climate and the hydrological cycle *Nature* **419** 224–32

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

Christensen J H *et al* 2007 Regional climate projections *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* ed S Solomon, D Qin, M Manning, Z Chen, M Marquis, K B Averyt, M Tignor and H L Miller (Cambridge: Cambridge University Press)

Dai A 2006 Recent climatology, variability, and trends in global surface humidity *J. Clim.* **19** 3589–606

Giorgi F and Francisco R 2000 Evaluating uncertainties in the prediction of regional climate change *Geophys. Res. Lett.* **27** 1295–8

Held I M and Soden B J 2000 Water vapor feedback and global warming *Annu. Rev. Energy Environ.* **25** 441–75

Held I M and Soden B J 2006 Robust responses of the hydrological cycle to global warming *J. Clim.* **19** 5686–99

Ingram W J 2002 On the robustness of the water vapor feedback: GCM vertical resolution and formulation *J. Clim.* **15** 917–21

Lanzante J R 1996 Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data *Int. J. Clim.* **16** 1197–226

Lott N, Baldwin R and Jones P 2001 The FCC integrated surface hourly database, a new resource of global climate data *National Climatic Data Center Technical Report No. 2001-01* NCDC Asheville, US, p 42

Manabe S and Wetherald R T 1967 Thermal equilibrium of the atmosphere with a given distribution of relative humidity *J. Atmos. Sci.* **24** 241–59

McCarthy M P and Willett K M 2006 Report on estimates of observational uncertainty in surface humidity and free-atmosphere temperature and humidity data *DEFRA Climate Prediction Program Annex III Milestone Report 03.02.06* Met Office Hadley Centre, Exeter, UK, p 17

Mears C A, Santer B D, Wentz F J, Taylor K E and Wehner M F 2007 Relationship between temperature and precipitable water changes over tropical oceans *Geophys. Res. Lett.* **34** L24709

Meehl G A, Covey C, Delworth T, Latif M, McAvaney B, Mitchell J F B, Stouffer R J and Taylor K E 2007 The WCRP CMIP3 multi-model dataset: a new era in climate change research *Bull. Am. Meteorol. Soc.* **88** 1383–94

New M, Hulme M and Jones P 1999 Representing twentieth-century space–time climate variability part 1: development of a 1961–90 mean monthly terrestrial climatology *J. Clim.* **12** 829–56

Santer B D, Penner J E and Thorne P W 2006a How well can the observed vertical temperature changes be reconciled with our understanding of the causes of these changes? *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences* ed T R Karl, S J Hassol, C D Miller and W L Murray (A report by the Climate Change Science Program and Subcommitee on Global Change Research, Washington, DC)

Santer B D *et al* 2006b Forced and unforced ocean temperature changes in Atlantic and Pacific tropical cyclogenesis regions *Proc. Natl Acad. Sci.* **103** 13905–10

Santer B D *et al* 2007 Identification of human-induced changes in atmospheric moisture content *Proc. Natl Acad. Sci. USA* **104** 15248–53

Santer B D *et al* 2008 Consistency of modelled and observed temperature trends in the tropical troposphere *Int. J. Clim.* **28** 1703–22

Sen P K 1968 Estimates of the regression coefficient based on Kendall's tau *J. Am. Stat. Assoc.* **63** 1379–89

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

Stone D A, Allen M R and Stott P A 2007 A multimodel update on the detection and attribution of global surface warming *J. Clim.* **20** 517–30

Stuart A and Ord J K 1987 *Kendall's Advanced Theory of Statistics* vol 1 *Distribution Theory* (London: Charles Griffin & Co) p 325

Sun D and Held I M 1996 A comparison of modeled and observed relationships between interannual variations of water vapor and temperature *J. Clim.* **9** 665–75

Taylor N A S 2006 Challenges to temperature regulation when working in hot environments *Indust. Health* **44** 331–44

Thorne P, Parker D, Christy J and Mears C 2005 Uncertainties in climate trends: lessons from upper-air temperature records *Bull. Am. Meteorol. Soc.* **86** 1437–42

Trenberth K 2007 Warmer oceans, stronger hurricanes *Sci. Am.* **297** 44–51

Trenberth K E, Fasullo J and Smith L 2005 Trends and variability in column integrated atmospheric water vapour *Clim. Dyn.* **24** 741–58

Wentz F J, Ricciardulli L, Hilburn K and Mears C 2007 How much more rain will global warming bring? *Science* **317** 233–5

Willett K W, Gillett N P, Jones P D and Thorne P W 2007 Attribution of observed humidity changes to human influence *Nature* **449** 710–2

Willett K W, Jones P D, Gillett N P and Thorne P W 2008 Recent changes in surface humidity: development of the HadCRUH dataset *J. Clim.* **21** 5364–83