

## An Analysis of Tropospheric Humidity Trends from Radiosondes

MARK P. MCCARTHY, P. W. THORNE, AND H. A. TITCHNER

*Hadley Centre for Climate Change, Met Office, Exeter, United Kingdom*

(Manuscript received 6 October 2008, in final form 27 May 2009)

### ABSTRACT

A new analysis of historical radiosonde humidity observations is described. An assessment of both known and unknown instrument and observing practice changes has been conducted to assess their impact on bias and uncertainty in long-term trends. The processing of the data includes interpolation of data to address known sampling bias from missing dry day and cold temperature events, a first-guess adjustment for known radiosonde model changes, and a more sophisticated ensemble of estimates based on 100 neighbor-based homogenizations. At each stage the impact and uncertainty of the process has been quantified. The adjustments remove an apparent drying over Europe and parts of Asia and introduce greater consistency between temperature and specific humidity trends from day and night observations. Interannual variability and trends at the surface are shown to be in good agreement with independent in situ datasets, although some steplike discrepancies are apparent between the time series of relative humidity at the surface.

Adjusted trends, accounting for documented and undocumented break points and their uncertainty, across the extratropical Northern Hemisphere lower and midtroposphere show warming of 0.1–0.4 K decade<sup>-1</sup> and moistening on the order of 1%–5% decade<sup>-1</sup> since 1970. There is little or no change in the observed relative humidity in the same period, consistent with climate model expectation of a positive water vapor feedback in the extratropics with near-constant relative humidity.

### 1. Introduction

The importance of water vapor in the study of climate change cannot be overstated. Water vapor through latent heat exchanges is the principal method of energy transport through the global atmosphere and it is a dominant greenhouse gas (e.g., Held and Soden 2000). Water vapor feedback represents the largest climate feedback mechanism simulated by general circulation models (GCMs) responding to a climate forcing such as increased CO<sub>2</sub> (Soden and Held 2006; Randall et al. 2007). However, if, as predicted by GCMs, relative humidity remains approximately constant under a changing climate, then the water vapor feedback is in part offset by the temperature lapse-rate feedback, and the net feedback uncertainty is much smaller than the individual components (Allan et al. 2002; Colman 2003) and little affected by model bias (John and Soden 2007). The radiative impact of water vapor relates to a fractional, not absolute, change in water vapor, resulting

in a disproportionate influence from the upper troposphere (e.g., Shine and Sinha 1991). These considerations provide a strong requirement for observational evidence with good vertical resolution for monitoring maintains the constant relative humidity in a changing climate.

The observation of water vapor in the atmosphere is an inherently difficult operation. Absolute concentrations decrease by orders of magnitude as one ascends from the surface to the stratosphere. This has limited our ability to detect emerging signals in global atmospheric water vapor, particularly in the radiatively important upper troposphere. Satellite observations of the 6.7- $\mu\text{m}$  water vapor emission band have been used to infer that trends in upper-tropospheric water vapor are consistent with warming and constant relative humidity (Soden et al. 2005). Regional trends in upper-tropospheric humidity may exist (Bates and Jackson 2001) but cannot be separated from instrumental biases and natural variability (McCarthy and Toumi 2004). More recently, an emerging picture of increases in surface and total column water vapor broadly consistent with constant relative humidity (Dai 2006; Trenberth et al. 2005) has been attributed to anthropogenically forced warming of

---

*Corresponding author address:* Mark McCarthy, Fitzroy Road, Met Office, Exeter, EX1 3PB, United Kingdom.  
E-mail: mark.mccarthy@metoffice.gov.uk

the climate system in recent decades (Willett et al. 2007; Santer et al. 2007).

The global radiosonde network provides a considerable resource of humidity observations through much of the twentieth century. However, the quality of the observations for climate applications is questionable (e.g., Elliott and Gaffen 1991; Elliott 1995; Guichard et al. 2000; Remsberg et al. 2000; Wang et al. 2003). Despite this, a number of studies have attempted to diagnose trends in tropospheric humidity from radiosondes for a number of regions (e.g., Gaffen et al. 1991; Gutzler 1992; Zhai and Eskridge 1997; Ross and Elliott 1996, 2001; Wang et al. 2001; Durre et al. 2009) and have tended to show increases in specific humidity with regional trends in relative humidity of mixed signs.

In this paper we present a new analysis of tropospheric humidity from radiosondes [the Hadley Centre gridded homogenized radiosonde humidity dataset (HadTH)], to estimate trends and uncertainty in both relative and specific humidity across the Northern Hemisphere over the last 30 yr. These data provide vertical resolution not available in many long-term surface or satellite-based measurements, and build upon earlier analyses of the radiosonde data, therefore complementing and building the existing evidence base. The purpose of HadTH is specifically to determine whether the radiosondes provide sufficient evidence to confidently reject the null hypothesis that relative humidity in the troposphere has remained at a constant level through recent climate change. Therefore, due care should be taken if HadTH is used for other applications because the historical biases affecting the radiosonde humidity data have not yet been as rigorously assessed as those for temperature.

## 2. Strategy

The network of routine radiosonde launching stations has undoubtedly been a major contribution to our understanding of the global circulation, hydrological cycle, and local dynamical processes. However, the network has not been explicitly designed to cater for the taxing requirements of climate monitoring, and it is heavily affected by technological and processing changes that act to mask any small climate signals that may be present in the data (e.g., Elliott and Gaffen 1991; Elliott 1995; Soden and Lanzante 1996; Sapucci et al. 2005). Individual station humidity records will contain changes in the measured humidity over time that result from changes in measurement technology (e.g., resistance, capacitive, or skin hygrometers), instrument design (e.g., housing of the instruments or inclusion of multiple sensors), instrument response times, and on-site or other data processing (e.g., Wade 1994; Elliott et al. 1998).

These considerations are further complicated by changes to the sampling and sensitivity across a variety of atmospheric states [e.g., the accuracy and precision in the cold, dry upper troposphere has improved considerably over time (Remsberg et al. 2000)].

The limited availability of metadata to describe many of these changes makes a physically based approach for absolute calibration of the myriad radiosonde instruments unfeasible. Therefore, we must rely on the statistically based adjustment of the long-term climate record. However, humidity biases are a function of the atmospheric state (e.g., Wade 1994; Elliott and Gaffen 1991), complicating the application of corrections to daily data using a simple statistical method. The non-linearity in humidity conversions makes it inappropriate to undertake bias correction on monthly mean temperature and dewpoints and then convert to relative and specific humidity (Gaffen et al. 1991), or to estimate monthly mean corrections for relative and specific humidity from monthly mean corrections applied to temperature and dewpoint. Therefore, we undertake the humidity conversion before the homogenization, and then homogenize the temperature, relative humidity, and specific humidity records independently. Because we have not forced consistency between the temperature and humidity variables, any discrepancies in the large-scale mean trends would indicate a serious failing of the method. The cost of this method is increased error and uncertainty in trends and variability at individual stations.

Homogenization of radiosonde temperature records for climate has tended to be manually intensive (e.g., Lanzante et al. 2003; Thorne et al. 2005a), although automated methods have more recently been developed (Haimberger 2007; Sherwood 2007; McCarthy et al. 2008). There are advantages and pitfalls with any method, and the uncertainty in long-term trends is poorly quantified in most, if not all, such methods (Thorne et al. 2005b). For the homogenization of the humidity data we use the method of McCarthy et al. (2008), which provides a reproducible homogenization with uncertainty resulting from both the detection and adjustment of the break points. The limitations of the system have been documented in McCarthy et al. (2008) and Titchner et al. (2009). The method appears to adequately estimate the random uncertainty in trend estimates resulting from the homogenization, but it struggles to fully account for any systematic bias that might pervade the network. In this analysis we introduce an additional first-guess correction based upon available metadata to offset this.

McCarthy (2008) shows that the spatial sampling uncertainty in large-scale mean humidity resulting from

the sparse radiosonde distribution is likely to be large. This will be particularly true in the Southern Hemisphere and the tropics. Stations in the 0°–70°N region were initially included in the homogenization system, but because of the large uncertainties in the tropics, from both the homogenization and limited spatial sampling, the subsequent analysis of the large-scale average trends has been limited to the 20°–70°N region.

The rest of this paper describes the impact of each stage of the data homogenization and resultant dataset. In section 3 we discuss the generation of monthly mean humidity data (section 3a), which is in-filled to avoid missing cold temperature and dry day sampling biases (section 3b) and then converted to anomalies (section 3c) for homogenization. The homogenization is then undertaken in two stages: a first guess of known changes (section 4a), followed by the neighbor-based homogenization (section 4b). Sections 5 and 6 then provide a discussion and conclusions from the resultant homogenized trends.

### 3. Data processing

We use daily temperature and dewpoint data from the surface and five fixed pressure levels—850, 700, 500, 400, and 300 hPa—from the Integrated Global Radiosonde Archive (IGRA; Durre et al. 2006). Data at 1000 hPa were included initially but were rejected from the extensive analysis because of limited availability. The IGRA quality control does not allow for observations where multiple, differing sources of the same observations exist, because there is no a priori way of preferentially selecting one data source over another (I. Durre 2005, personal communication). This results in some significant gaps in the IGRA humidity record at some stations. For this analysis data completeness was considered a high priority; therefore, the dataset was supplemented with additional data from the earlier Comprehensive Aerological Reference Dataset (CARDS; Eskridge et al. 1995) at stations where CARDS was found to be significantly more complete than IGRA. Metadata from IGRA (Durre et al. 2006) have been used for the identification of break points.

#### a. Calculation of monthly means

Monthly means for temperature, relative humidity, and specific humidity, at each pressure level, and separately for 0000 and 1200 UTC launch times, were calculated for the 34-yr period from December 1969 through to November 2003. The method for converting data into relative and specific humidity is described in appendix A. Biweight means (Lanzante 1996) were used, in order to reduce the influence of outliers, and

monthly means were only calculated where at least 15 days of temperature observations were present within a month. For humidity some of the “observations” had to be interpolated to remove sampling biases (see section 3b) so constraints on the number of observations were only applied to the temperature data. Only data with a launch time within 3 h of either 0000 or 1200 UTC were retained.

#### b. Known sampling bias

In this section we address two important sampling biases in the raw humidity data, from missing dry observations [section 3b(1)] and missing cold observations [section 3b(2)]. We then conduct a quantitative assessment of the interpolation method employed [section 3b(3)].

##### 1) MISSING DRY OBSERVATIONS

Between 1973 and 1993, at many stations operated by the United States and a few other countries, low-humidity observations were considered unreliable, and where the relative humidity fell below 20% it was routinely recorded as 19% relative humidity or reported as a dewpoint depression of 30°C (Wade 1994). This procedure results in a bias in monthly mean averages during this period (Elliott et al. 1998), and a spurious trend in the long-term record because of the termination of this operating practice in 1993. Ross and Elliott (1996) replaced these “dry” observations with a value of 16% relative humidity, estimated from Canadian stations that did not undertake this practice. At most stations there are now several years of low-humidity observations post-1993. We have estimated replacement values for individual stations and levels as the median relative or specific humidity from observations made on or after 1 January 1995, and where the relative humidity is less than 20%. While this may allow us to recover trends that result from changes in the frequency of dry events, we will not recover any trend resulting from changes in the magnitude of dry events.

##### 2) MISSING COLD OBSERVATIONS

A second sampling issue in the database relates to humidity observations in cold conditions. It was standard practice until 1993 for U.S. stations to report humidity measurements as missing when the ambient air temperature was below  $-40^{\circ}\text{C}$ . Most early radiosonde hygrometers are considered unreliable at and below these temperatures (Remsberg et al. 2000). We account for this by rejecting all levels for a particular season and station where the air temperature is below  $-40^{\circ}\text{C}$  in more than 5% of available observations. Despite this, a warm temperature humidity sampling bias still exists

in the early part of the IGRA database. This is apparent by comparing temperature records that are sampled using all temperature observations with those using only temperature data where an associated humidity observation also exists. This suggests that humidity observations are missing or have been preferentially rejected by quality control under colder conditions early in the record, and therefore dry events are missing in the early part of the record. This is undesirable for a climate record and results in a spurious drying trend.

Most stations have had near-complete humidity sampling (relative to temperature) since the mid-1990s. Therefore, for each station we determined the least squares linear regression between monthly mean temperature and relative humidity, and between monthly mean temperature and the natural logarithm of specific humidity using data from 1995 to 2003. Examples for the 850-hPa level at Lindenberg, Germany (52.22°N, 14.12°E), and Minamitorishima, Japan (24.30°N, 153.97°E), are shown in Fig. 1. The regression coefficients vary considerably from place to place and with pressure level, but in general specific humidity is more strongly tied to temperature [regression coefficients of  $0.076 \log(\text{g kg}^{-1}) \text{K}^{-1}$  and  $0.063 \log(\text{g kg}^{-1}) \text{K}^{-1}$  for Lindenberg and Minamitorishima, respectively] than relative humidity, which is generally very weak or weakly negatively correlated to temperature (regression coefficients of  $0.2\% \text{K}^{-1}$  and  $-0.6\% \text{K}^{-1}$  for Lindenberg and Minamitorishima, respectively). In other words, for many locations month-to-month variations in specific humidity largely, but not completely, compensate for temperature changes to maintain the relative humidity (or relative saturation deficit), consistent with previous studies (e.g., Ross et al. 2002).

The regression coefficients are used to estimate the mean of the missing humidity observations from the mean of the associated temperature measurements that do exist. Where the regressions are not statistically significant the interpolation is still performed, rather than developing an alternative method in these instances. We must also assume that the regression coefficients for the modern era are representative of the earlier period. The suitability of this approach has therefore been quantitatively assessed and summarized in the next section.

### 3) ESTIMATE OF SAMPLING BIASES

To investigate the impact and suitability of the data interpolation and the underlying assumptions, we selected 24 case study stations (see appendix B) from a cross section of locations. By manual inspection it was confirmed that these stations were free from the specific warm temperature or dry condition sampling biases of

the type described above. For each station we conducted the following tests:

- 1) Dry days = 19% RH: Replace all incidences of daily  $\text{RH} < 20\%$  with  $\text{RH} = 19\%$  to test the impact of the low-humidity reporting practice.
- 2) Dry days = 16% RH: Replace all incidences of daily  $\text{RH} < 20\%$  with  $\text{RH} = 16\%$  to test the impact of the Ross and Elliott (1996) amendment to the above. There is no trivial way to interpret either the 19% or 16% RH replacement in terms of specific humidity  $q$ , so experiments 1 and 2 were only applied to relative humidity.
- 3) Dry days interpolated: Replace all relative and specific humidity where  $\text{RH} < 20\%$  with the median value of relative or specific humidity from all instances of  $\text{RH} < 20\%$  in 1995–2003 data from the same station [see section 3b(1)]; this tests the approach adopted in this work to correct for the low-humidity sampling bias.
- 4) Cold days missing: Successively remove the coldest  $x$  days of observations before calculating monthly means, where  $x$  ranges from 1 to 25. This tests the impact of the temperature sampling bias described in section 3b(2) for varying amounts of missing data.
- 5) Cold days interpolated: Replace the relative and specific humidity observations for the coldest  $x$  days with values estimated from the temperature observations and linear regression coefficients for 1995–2003 [see section 2b(2)], to test the correction for the cold sampling bias.

The largest potential sampling bias results from the missing cold days as shown in Fig. 2. The interpolation of humidity significantly reduces, but does not completely remove, this bias. For example, a situation in which 15 humidity observations are available and 15 are missing the sampling bias can exceed 5% relative humidity and 13% specific humidity at 850 hPa. The interpolation reduces relative humidity bias to 2% and specific humidity bias to 5% in this test for the 850-hPa level, averaged across all of the test-case stations. At 500 hPa the cold-day sampling results in a smaller bias in relative humidity, and the data interpolation has no significant impact on this bias. This reflects a weak correlation between temperature and relative humidity at this level in the test-case stations. However, the specific humidity bias is larger at 500 hPa and the interpolation removes virtually all of this bias. The root-mean-square error is not improved by the interpolation, except for the cases with a small number of observations (and therefore the greatest interpolation), but this is less important than the bias for diagnosing long-term climate change. The root-mean-square error at 850 hPa is representative of

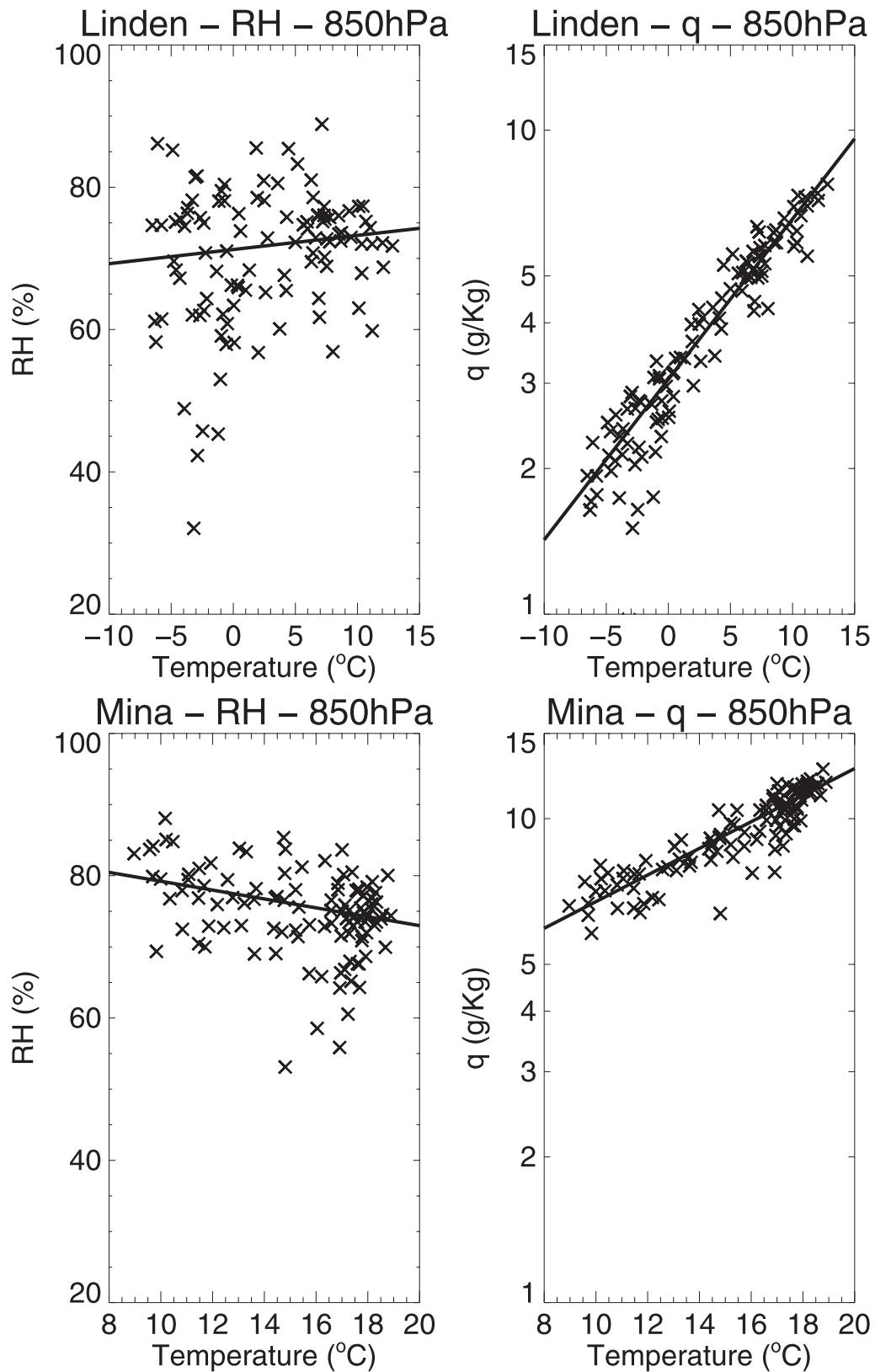


FIG. 1. Monthly mean RH and natural log of specific humidity against monthly mean temperature at 850 hPa for (top) station 10393, Lindenberg, and (bottom) station 47991, Minamitorishima.

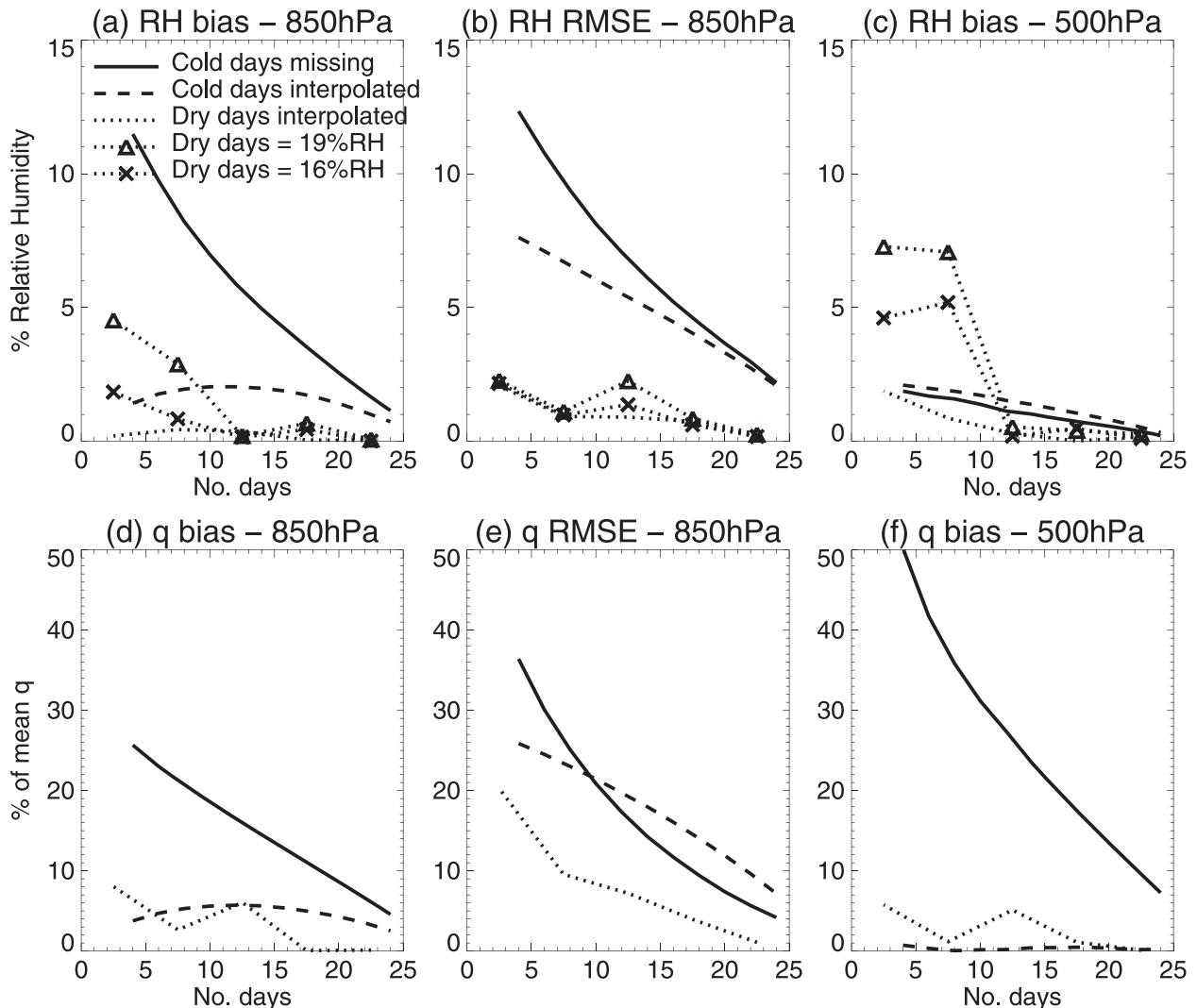


FIG. 2. (a) Absolute bias and (b) root-mean-square error in estimates of monthly mean RH at 850 hPa for varying number of available days of observation. (c) As in (a), but for 500 hPa. (d)–(f) As in (a)–(c), but for specific humidity. The  $x$  axis is the number of actual observations available for the monthly mean calculation, and all of the other days for the month are either missing or have been interpolated. Missing days are the coldest  $x$  days in the month (solid) and the humidity data for the coldest  $x$  days have been interpolated from the temperature data (dashed). All  $\text{RH} < 20\%$  have been replaced with the median of  $\text{RH} < 20\%$  for data between 1995 and 2003 (dotted). All  $\text{RH} < 20\%$  have been replaced with  $\text{RH} = 19\%$  (triangles) and all  $\text{RH} < 20\%$  have been replaced with  $\text{RH} = 16\%$  (crosses).

behavior at other pressure levels and so only the example of 850 hPa is shown in Figs. 2b,e. The method of low-RH replacement used here is an improvement over the previously used 16% RH value (Ross and Elliott 1996), but only for those situations where the relative humidity is persistently low (i.e., more than 20 days month<sup>-1</sup> that are below 20% RH; see Figs. 2a,c).

We used the results from the test-case evaluations to estimate the potential sampling bias in the humidity dataset after interpolation of the missing data. We assume a linear relationship between the bias and the number of interpolated observations in a month. Figures 2a,c,d,f suggest that this is not perfect; in particular, the

cold-day interpolation at 850 hPa (dashed curve in Figs. 2a,c) suggests a peak bias when approximately 50% of the month is interpolated. However, it should be a reasonable approximation for the purpose of estimating the potential bias. Cold-day sampling results in potential bias in relative humidity of  $0.08\% \text{ day}^{-1}$  at both 850 and 500 hPa,  $0.2\% \text{ day}^{-1}$  for specific humidity at 850 hPa, and  $0.015\% \text{ day}^{-1}$  for specific humidity at 500 hPa. Dry-day sampling results in estimated potential bias in relative humidity of  $0.02\% \text{ day}^{-1}$  and  $0.03\% \text{ day}^{-1}$  at 850 and 500 hPa, respectively,  $0.18\% \text{ day}^{-1}$  for specific humidity at 850 hPa, and  $0.13\% \text{ day}^{-1}$  for specific humidity at 500 hPa.

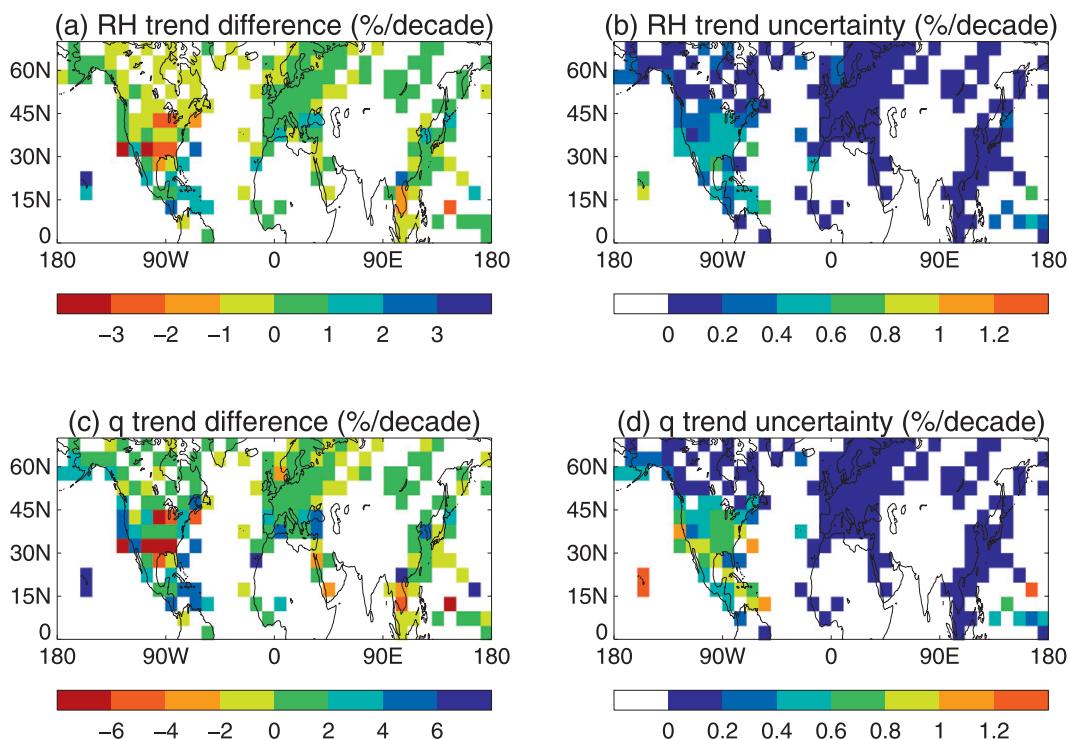


FIG. 3. The difference in linear trends between the data both (a) with and (c) without interpolation of missing humidity observations. (b),(d) The estimated trend uncertainty resulting from the interpolation process based on the amount of data that have been interpolated: (a),(b) RH and (c),(d) specific humidity.

For each of the stations in the complete dataset we determined the time series of potential systematic bias resulting from the actual number of interpolated days in each month resulting from both the cold-day and dry sampling biases described. The resulting trend in the time series of this estimated bias is shown for both relative and specific humidity at 500 hPa in Figs. 3b,d on a  $5^\circ$  latitude  $\times$   $10^\circ$  longitude grid. Figure 3 also shows the trend difference between the interpolated and noninterpolated datasets (Figs. 3a,c). The potential trend bias estimated in this way is significantly smaller than the difference between the interpolated and noninterpolated datasets, providing evidence that the application of the data interpolation described above is appropriate for reducing trend bias. However, we must accept that systematic trend biases of the order of  $1\%$  decade $^{-1}$  may still occur at specific locations across North and Central America and the Pacific (Fig. 3d). It should also be noted that the sign of this bias may not be consistent everywhere because stations have come from different source datasets, with some CARDS records archiving dry days as 19% RH, yielding a positive bias in monthly means, and with IGRA records archiving dry days as a dewpoint depression of  $30^\circ\text{C}$ , yielding a negative bias.

### c. Calculation of climatology, anomalies, and gridded means

Monthly climatologies were determined for the period December 1980 through to November 2000. Climatologies were calculated separately for day and night, and for each month, variable, and pressure level. Climatologies were calculated where at least 2 months of data were available in at least three seasons of the year in at least 7 yr of each decade of the climatology period. This nominally limits HadTH to a top level of 300 hPa at tropical latitudes, and 500 hPa at high latitudes. Where climatologies exist the monthly mean data were converted to anomalies with respect to 1981–2000 and set to missing elsewhere. Gridded fields, where they are presented, have been calculated as a simple average of all radiosonde stations falling within a  $5^\circ$  latitude  $\times$   $10^\circ$  longitude grid box, as in Thorne et al. (2005a).

## 4. Homogenization

### a. First guess of known radiosonde changes

Technological advancements in humidity hygrometers have resulted in large step changes in the radiosonde humidity record at many stations. Improved sensitivity

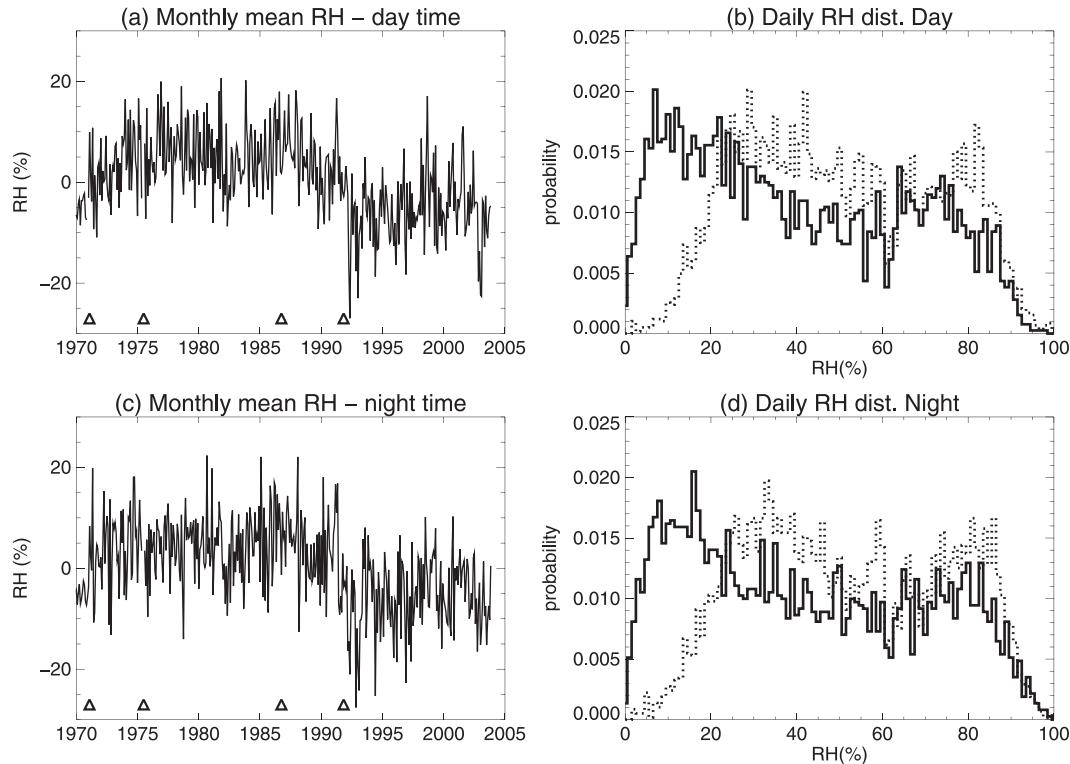


FIG. 4. Time series of monthly mean relative humidity at 500 hPa for Lindenberg for (a) day and (c) night soundings. (b),(d) Frequency distributions of point measurements of relative humidity at 500 hPa for the periods of 1970–91 (dotted) and 1992–2003 (solid). Triangles denote the location of recorded metadata events. 1971, 1986, and 1991 are all radiosonde instrument changes; 1975 is a recorded station identification change.

to dry conditions and improvements to instrument response times result in negative humidity trends in the raw radiosonde archive at many stations across Europe, Russia, and Japan. An example for Lindenberg is shown in Fig. 4. A large step change in relative humidity is apparent in 1989, associated with a change from MARS (Russian sonde) to Vaisala radiosondes, which corresponds to a switch from a goldbeater's skin to a capacitive hygristor. This particular change occurred at many stations across the former East Germany, and there are other examples of similar large step changes in humidity.

Because of the known limitations of the automated homogenization method of McCarthy et al. (2008), the first step of the homogenization was to provide objective first-guess adjustments for the most common radiosonde instrument changes, many of which result in large changes in the recorded mean humidity (e.g., Fig. 4). From the available metadata record we identified all specific radiosonde instrument changes (e.g., from MARS to Vaisala RS80) that were recorded at four or more radiosonde stations. For each such event we then calculated the difference in the median temperature and relative and specific humidity, for the 2-yr period before and after the recorded event. In previous studies (e.g.,

Thorne et al. 2005a; McCarthy et al. 2008), much longer periods have been used for the estimation of breaks. In this case a shorter window is used because we are using the station time series in isolation, and therefore a long window would act to remove the trend we are attempting to detect. For example, in the presence of a linear specific humidity trend of  $2\% \text{ decade}^{-1}$  (or  $0.016 \text{ g kg}^{-1} \text{ decade}^{-1}$  at 500 hPa), every such break correction calculated from the station series has the potential to reduce the trend by  $0.2\% \text{ decade}^{-1}$  using a 2-yr window, or  $0.5\% \text{ decade}^{-1}$  using a 5-yr window, if the break is near the central point of the time series. The cost of reducing the averaging window is to increase the uncertainty from the natural variability in the time series. Tests on random samples from the case study stations (appendix B) at 500 hPa gave a standard deviation of the difference in medians for 2-yr windows of the order of  $0.08 \text{ g kg}^{-1}$ , and from a 5-yr window  $0.04 \text{ g kg}^{-1}$ . This component of uncertainty will be reduced by averaging over multiple stations. Two-year windows result in a standard deviation in median differences of  $0.5 \text{ K}$  for temperature at 500 hPa, on a  $0.2 \text{ K decade}^{-1}$  trend. This is the same as the adjustment uncertainty of  $0.5 \text{ K}$  estimated more comprehensively by McCarthy et al. (2008) for the neighbor-based

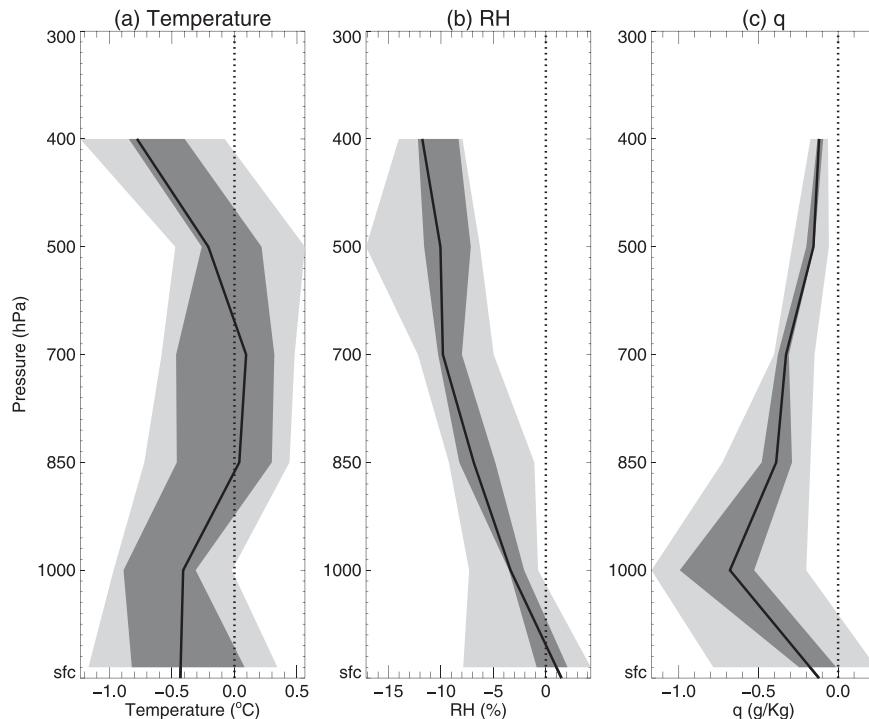


FIG. 5. Profile of the first guess of a break resulting from a switch from MARS to Vaisala RS80. The median of 14 individual station estimates (solid line), the interquartile range (dark shading), and the full spread of 14 station estimates (light shading) are denoted.

homogenization system. In other words, the random error of this first-guess adjustment is of a similar magnitude to that from the neighbor-based homogenization, but each adjustment introduces a systematic error by acting to reduce trend present in the time series. It is this systematic error we wish to avoid, so the first-guess adjustment is limited to known radiosonde model changes only.

The first-guess adjustment for the MARS to Vaisala RS80 instrument change is shown in Fig. 5. At 700 hPa and above the estimated relative humidity correction is close to 10%, and all 14 contributing stations have a change of greater than 5% (i.e., MARS recorded between 5% and 15% higher relative humidity than Vaisala RS80). A full list of first-guess break points is provided in appendix C (Table C1). These are not expected to represent a comprehensive assessment of radiosonde bias related to these particular radiosonde model changes; rather, they are a first-guess estimate of the change points that are then modified by the subsequent, more rigorous, neighbor-based homogenization. The values presented (in Table C1) are averages across all of the radiosonde pressure levels and are provided to illustrate the order of magnitude of the estimated step changes associated with the metadata events. The first-guess adjustments are calculated and applied separately for each pressure level.

### b. Homogenization

The method of McCarthy et al. (2008) is employed to

- detect and adjust for both known and unknown break points,
- refine the first-guess break points from section 4a using a neighbor-based method, and
- quantify the trend uncertainty resulting from the homogenization method.

This is an automated, iterative, neighbor-based homogenization scheme. Break-point identification applies a Kolmogorov–Smirnov test to time series of differences between individual station series and a weighted composite of near neighbors. The contributing neighbors are taken from those locations in the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis where the correlation with the target stations time series of seasonal means is greater than  $1/e$ . An example for Valentia, Ireland, at 850 hPa is shown in Fig. 6. The extent of the humidity neighbor region is smaller than that for temperature because of the more heterogeneous distribution of water vapor within the troposphere. However, over Europe the correlation scale for both relative humidity and specific humidity is large enough to allow

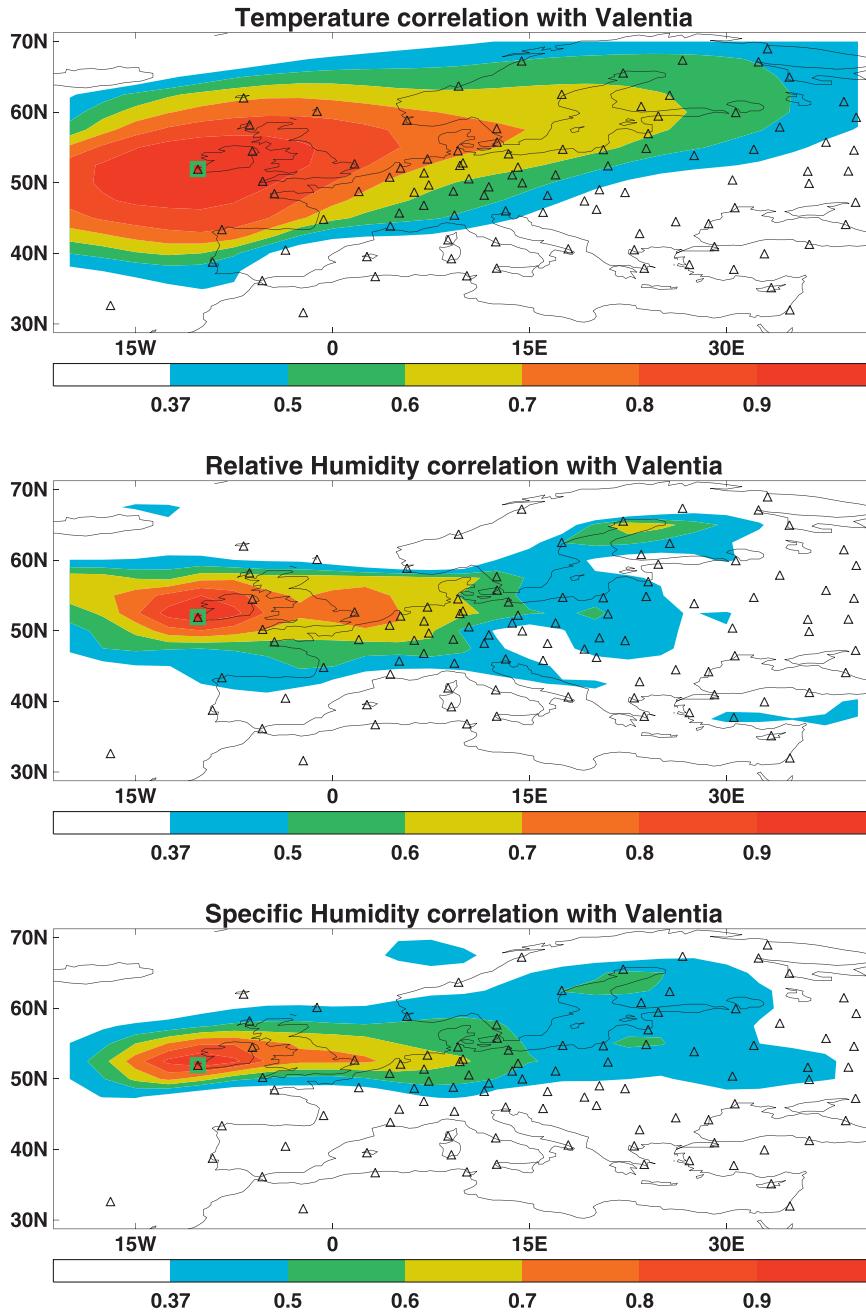


FIG. 6. Fields of 850-hPa correlation coefficients from NCEP–NCAR reanalysis of seasonal mean time series against the grid box containing the location of the radiosonde station at Valencia (green square). The location of radiosonde stations that have been included in HadTH are marked (triangles). All stations that fall within the contiguous area of correlation greater than  $1/e$  are used as neighbor stations for Valencia.

for over 40 contributing neighbor stations at 850 hPa for this station on the western edge of Europe. The system uses the actual radiosonde station data to derive the neighbor reference selection, and the NCEP–NCAR reanalysis plays no further role in the homogenization. At least three neighbors are required for the homogenization.

The homogenization system is fully automated, with a set of 14 tunable parameters that control the subjective decision processes relating to the identification and adjustment of break points at individual levels. The homogenization has been run 100 times with a range of parameter settings, and the details of these are discussed

in McCarthy et al. (2008). The homogenization was conducted independently on temperature, relative humidity, and specific humidity, for both day and night. The 100 homogenizations provide a first-order estimate of the uncertainty that results from the homogenization because each run will identify a different set of possible break points and their adjustments. We have shown previously (McCarthy et al. 2008) that this method provides a reasonable estimate of the random uncertainty in large-scale trend estimates.

We also wish to construct the most representative homogenization from this ensemble to create a single coherent dataset, which will be referred to as the “best guess” version. To do that we identify break points that consistently appear in at least 50 of the 100 homogenization members [see also Guo et al. (2008), who used a slightly more liberal 25-member criteria in an analysis of break points in Chinese radiosonde data]. To do this we first cluster the break points because the same break is not located at exactly the same time in different members of the homogenization ensemble. First, we take the count of breaks occurring in each year from the 100 members. Starting from the largest concentration we then attribute breaks located within two calendar years (the default minimum break-point separation in the break-point detection algorithm) to that particular break point, and then repeat for the next largest concentration of breaks. An example of the distribution of breaks identified in the 100-member homogenization is shown in Fig. 7. It is immediately apparent that there is a high degree of consistency in the location and sign of the breaks, although the adjustment magnitude can differ considerably between ensemble members (see also Guo et al. 2008; Titchner et al. 2009).

The identified sign and location of break points in Fig. 7 are also broadly consistent for temperature and the two humidity variables. Even with a small number of break points the resultant uncertainty on the trend in an individual station time series is large. For the example shown in Fig. 7, the adjusted best-guess trends are  $0.25 \text{ K decade}^{-1}$ ,  $0.53\% \text{ decade}^{-1}$ , and  $0.16 \text{ g kg}^{-1} \text{ decade}^{-1}$  for temperature and relative and specific humidity, respectively. The 95% range of trend estimates for the 100 homogenization experiments at 850 hPa is  $0.3 \text{ K decade}^{-1}$  for temperature,  $1.7\% \text{ decade}^{-1}$  for relative humidity, and  $0.1 \text{ g kg}^{-1} \text{ decade}^{-1}$  for specific humidity.

First-guess adjustments for Lindenberg were implemented for a change in 1986 associated with a switch in radiosondes from RKZ5 to MARS, and a subsequent change from MARS to Vaisala in 1991 (see also Fig. 5). The neighbor-based homogenization modifies

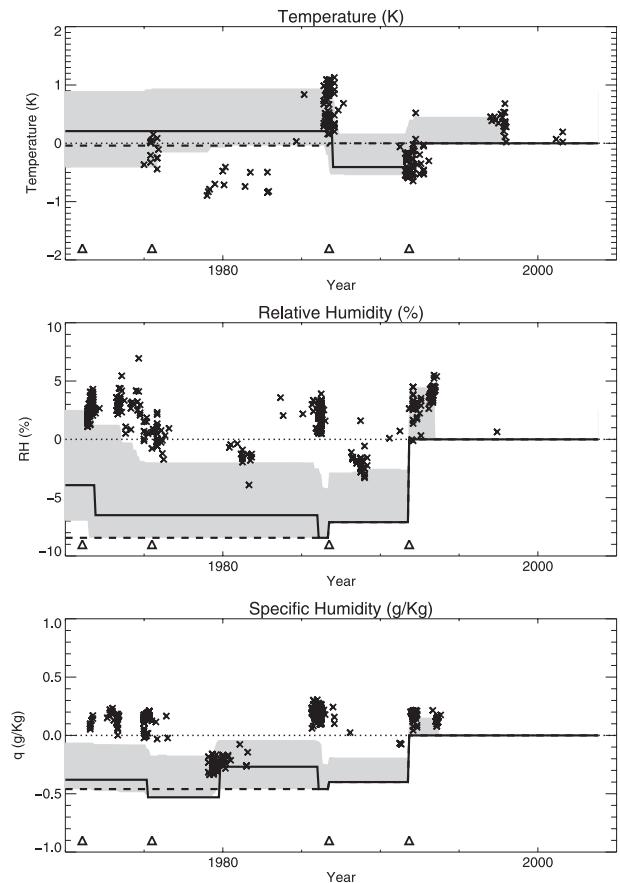


FIG. 7. Time series of break points identified for Lindenberg at 850 hPa. Each cross represents the location and estimated magnitude of a break identified by a single member of the ensemble of 100 homogenizations applied to the data. The adjustment time series for the best guess of the ensemble (solid line; see text), the adjustment time series from the first-guess corrections (dashed line), the spread of adjustment time series for the 100 homogenizations (gray shading), and the location of recorded metadata events (triangles) are shown.

both of these adjustments, introducing much larger temperature adjustments at both and offsetting the humidity change in 1986. Some of the homogenizations also reduce the magnitude of the 1991 change. A humidity correction in the 1970s has been implemented at different times for relative and specific humidity, and an additional correction to specific humidity is applied in 1980. Inconsistencies such as these are inevitable in the homogenization scheme utilized here, but are encapsulated in the uncertainty from the 100-member ensemble of experiments. This approach is highly unlikely to be suitable for the analysis of trends in individual station records. If such local or regional detail is required, then a much more rigorous assessment would be required. Over large scales the noise resulting from localized natural and synthetic

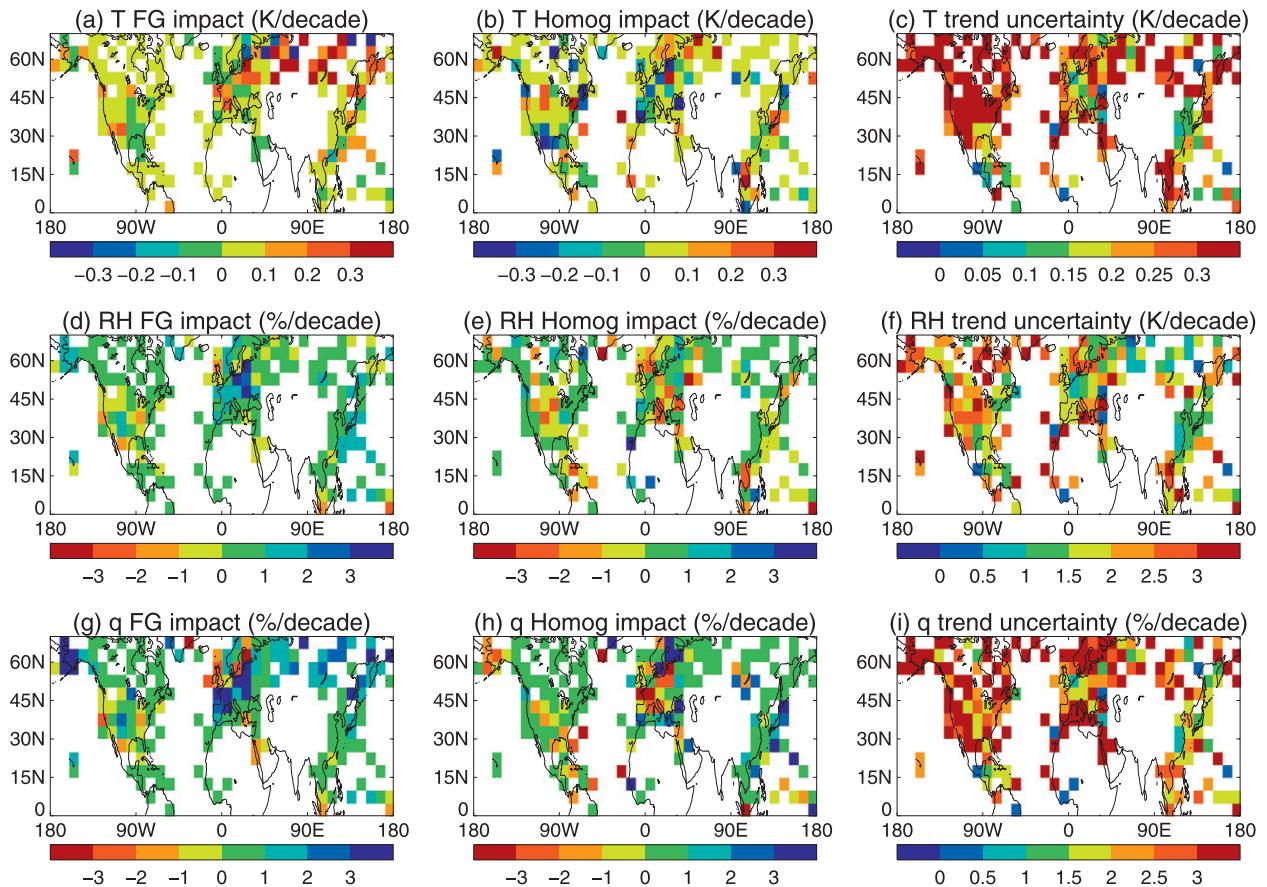


FIG. 8. Difference in trends between first-guess adjusted data and the unadjusted data for (a)  $T$ , (d) RH, and (g)  $q$ . (b), (e), (h) As in (a), (d), and (g), but for the difference between the best-guess homogenized dataset and the first-guess adjusted data. (c), (f), (i) The uncertainty in gridbox mean trends from the 100-member ensemble of homogenizations.

events will be reduced, making the detection and quantification of uncertainty of small climate signals more likely.

The impact of homogenization on trends at 850 hPa is shown in Fig. 8. The first-guess corrections resulting from a number of Russian radiosonde model changes increase the daytime warming trends, consistent with a reduction in the day–night bias resulting from improvements to the instrument housing and solar radiation corrections (e.g., Luers and Eskridge 1998). The relatively large adjustment to the trend results from the cumulative impact of the adjustments (presented in Table C1). Both relative and specific humidity trends are increased by the first-guess adjustments over parts of Europe, Asia, and Alaska. The subsequent homogenization shows a mixed signal, but tends to reduce temperature and humidity trends over parts of Europe. The estimated gridbox uncertainties are large, and generally larger than the magnitude of the trends, demonstrating the limited scope for diagnosing trends in small regions or for individual stations.

## 5. Trends and uncertainty

The net impact of the interpolation and homogenization of the dataset is shown in Fig. 9. The maps illustrate the spatial impact of the homogenization and are not meant to suggest that we have faith in trends at individual grid cells. Most notably, an apparent drying over Europe and Japan is removed. Increased temperature trends at high latitudes are a result of the inclusion of the “missing” humidity temperatures (i.e., the unadjusted data include only temperature data where humidity observations also exist, while the adjusted data include all temperature data and interpolated humidity) and the adjustment to Russian temperatures (Figs. 8a and 9c). There is still a degree of spatial heterogeneity, with a few large individual outliers, reflecting the large uncertainty at the gridbox level shown in Fig. 9. Despite this, the homogenized data show more widespread warming, smaller trends in relative humidity, and larger trends in specific humidity than the unadjusted data.

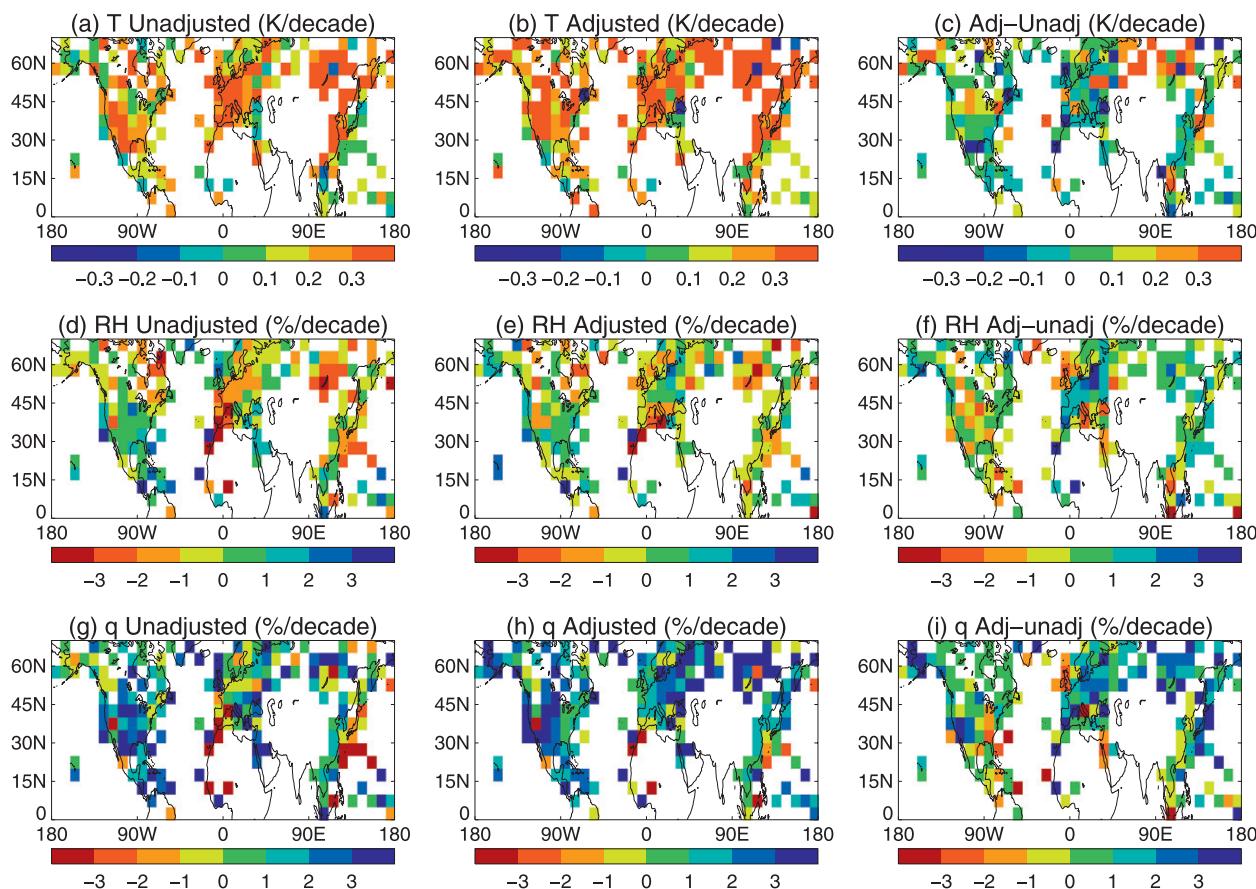


FIG. 9. Linear trends at 850 hPa in (a),(d),(g) unadjusted, (b),(e),(h) adjusted, and (c),(f),(j) adjusted minus unadjusted HadTH (a)–(c) temperature, (d)–(f) relative humidity, and (g)–(i) specific humidity.

In Fig. 10 we show trend profiles for the area-weighted average of the Northern Hemisphere extratropics (north of 20°N). The uncertainties represent the total homogenization uncertainty, but do not account for other sources, such as spatial sampling. The homogenization results in increased daytime warming trends. The improved agreement between the daytime and nighttime data for both temperature and specific humidity is encouraging and would be expected given the known solar radiation biases that afflict the daytime radiosonde launches (Sherwood et al. 2005; Randel and Wu 2006). The profile of extratropical temperature trends are broadly consistent with previously published observational and model results (Santer et al. 2006). The largest warming of the surface relative to the troposphere occurs at high latitudes. These results cannot help to address issues relating to tropical upper-tropospheric warming rates. Relative humidity trends are reduced to close to zero at most levels. Specific humidity increases, accounting for the associated uncertainty, at levels above the surface are within the expectation of moistening at

a rate that is consistent with maintaining relative humidity. Small negative trends in relative humidity do persist at the lowest levels, suggesting the possibility that moistening near the surface does not balance the warming at this level.

Idealized GCM experiments have shown that limits to local evaporation and moisture convergence over land occurring at levels colder than the surface act to reduce relative humidity in the boundary layer over land, and enhance the land–sea temperature contrast under warming scenarios (Joshi et al. 2008). The HadTH data are predominantly sampled over land; therefore, a small reduction in near-surface relative humidity in HadTH could be consistent with GCM simulations of local feedbacks of the hydrological cycle over land under recent warming, but the uncertainties associated with the observations are still too large to reject the null hypothesis of zero change in relative humidity.

The radiosonde records compare well with equivalently sampled surface data records, shown in Fig. 11, using the HadTH radiosonde surface level, the Hadley

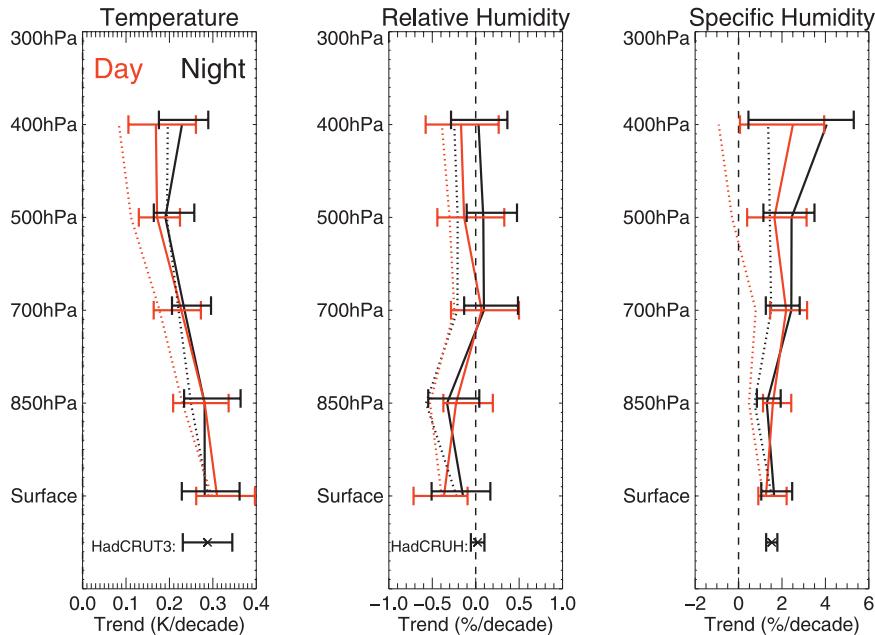


FIG. 10. Vertical profile of trends for  $T$ , RH, and  $q$  for day (red) and night (black) data for area-weighted means of data between  $20^{\circ}$  and  $70^{\circ}$ N. Solid lines are the best-guess adjusted data, with error bars denoting the spread of estimates from 100 homogenizations. The dotted lines are the unadjusted trends. Trends for HadCRUT3 and HadCRUH (see text) are also shown with their statistical fit uncertainty.

Centre and Climate Research Unit gridded near-surface temperature dataset (HadCRUT; Brohan et al. 2006), and the Hadley Centre and Climate Research Unit global surface humidity dataset (HadCRUH; Willett et al. 2008). The HadCRUT and HadCRUH have been sampled only at locations and times that exist within the radiosonde HadTH dataset. A discrepancy exists between the estimates of near-surface relative humidity trends. Closer analysis of the time series suggests that the difference in relative humidity trends is a result of a steplike change in the difference between HadTH and HadCRUH around 1992, and agreement is better for the preceding and following periods. It is beyond the scope of this investigation to suggest a resolution, but HadTH uncertainty does encapsulate the HadCRUH trend magnitude.

## 6. Conclusions

We have conducted an analysis of the long-term record of humidity from Northern Hemisphere radiosonde stations. The unadjusted observations are plagued by large, spurious step changes resulting from improvements to the measurement, observing, and archiving practices over time. These activities have tended to result in similar impacts on station humidity trends resulting

in an apparent reduction in relative humidity at many stations in the network. These coherent biases are difficult for some traditional homogenization methods to identify and appropriately correct. We use a combination of interpolation to remove temporally varying sampling bias, a crude first-guess adjustment, and a more sophisticated neighbor-based homogenization in an attempt to investigate the robustness of the apparent trends in the radiosonde humidity record.

Following the adjustment process the daytime temperature and specific humidity trends are increased and more consistent with the nighttime trends and are in the range of  $0.1\text{--}0.4\text{ K decade}^{-1}$ . Relative humidity trends are reduced from a negative value in the unadjusted data to near zero. Trends at the surface agree well with alternative surface-based estimates of temperature and humidity, although a discrepancy in the near-surface relative humidity is apparent.

Based on the available evidence we cannot reject the underlying hypothesis that under recent climate change the average tropospheric relative humidity in the Northern Hemisphere remains unchanged while temperature and specific humidity increase. Specific humidity increases since 1970 are on the order of  $1\%\text{--}5\%\text{ decade}^{-1}$ . Further rigorous analysis is required to better understand the uncertainty and more confidently

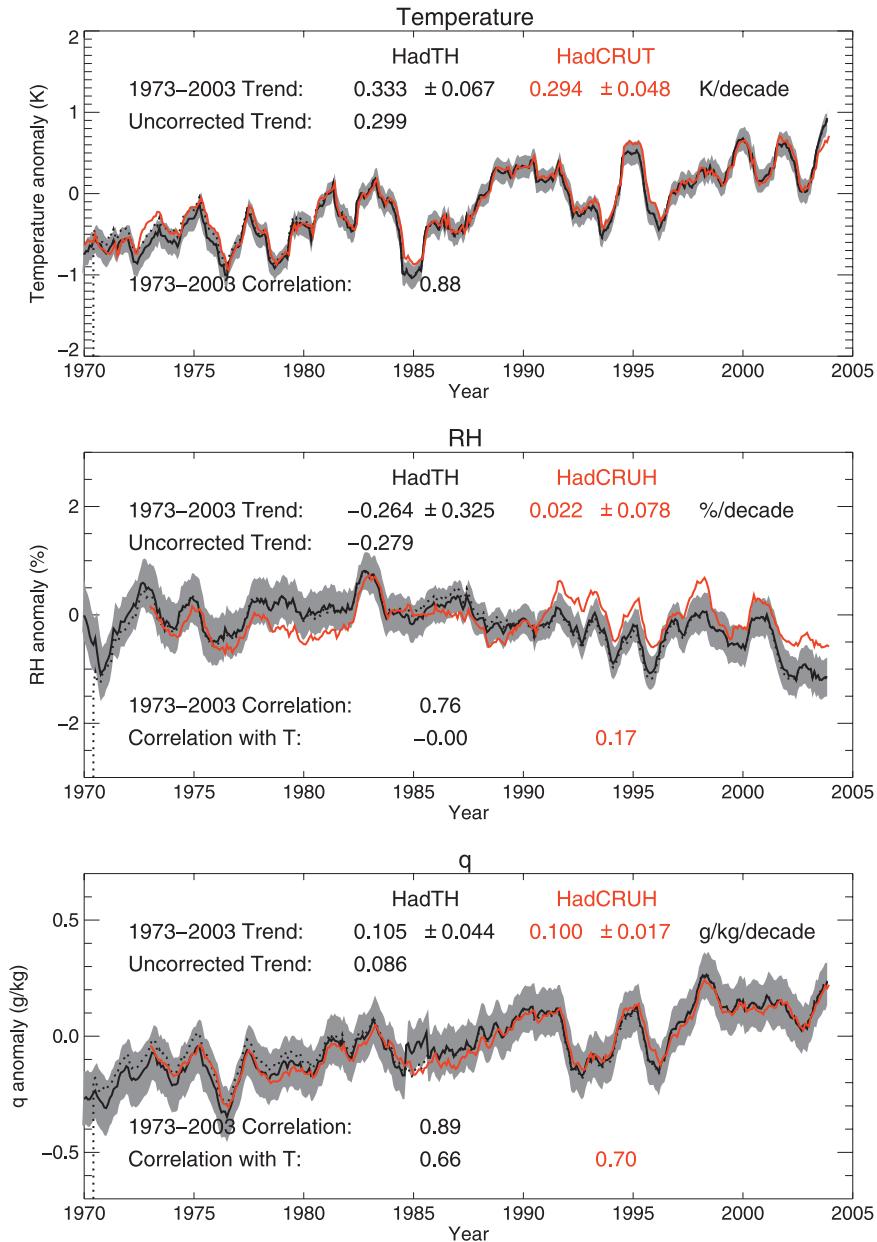


FIG. 11. Time series of (top)  $T$ , (middle) RH, and (bottom)  $q$  for adjusted surface records from HadTH (black), HadCRUT3 (red), and HadCRUH (red), and unadjusted HadTH (dotted). Gray shading represents the HadTH homogenization uncertainty estimates. The trend magnitudes for each dataset are also provided for the period of 1973–2003, along with the uncorrected HadTH trend. The correlation coefficient between the monthly mean HadTH and HadCRUT and HadTH and HadCRUH are also given. In addition the correlation to the surface temperature is provided for the humidity data.

assess the magnitude of tropospheric temperature and humidity trends. Multivariable analyses of the radiosonde record, including humidity and winds [e.g., see Allen and Sherwood (2008) for a novel use of wind observations] may also provide crucial additional evidence to aid the detection and adjustment of the previously

well-studied, but still highly uncertain, temperature records.

*Acknowledgments.* This work was supported by the Joint DECC, Defra, and MoD Integrated Climate

TABLE B1. Test-case stations used in the determination of residual bias and uncertainty following the interpolation of missing data.

Station number, location	Country	Latitude (°)	Longitude (°)
02963, Jokioinen	Finland	60.82	23.50
03808, Camborne	United Kingdom	50.22	-5.32
06610, Payerne	Switzerland	46.82	6.95
07510, Bordeaux	France	44.83	-0.68
08495, Gibraltar	Gibraltar	36.15	-5.35
10393, Lindenberg	Germany	52.22	14.12
10410, Essen	Germany	51.40	6.97
11520, Praha/Libus	Czech Republic	50.00	14.45
11952, Poprad/Ganovce	Slovakia	49.03	20.32
12843, Budapest/Lorinc	Hungary	47.43	19.18
16320, Brindisi	Italy	40.65	17.95
22550, Arkhangelsk	Russia	64.58	40.50
24266, Verkhoyansk	Russia	67.55	133.38
27459, Niznij Novgorod	Russia	56.27	44.00
31736, Khabarovsk	Russia	48.53	135.23
40179, Bet Dagan	Israel	32.00	34.82
40375, Tabouk	Saudi Arabia	28.37	36.58
47991, Minamitorishima	Japan	24.30	153.97
54342, Shenyang	China	41.82	123.55
57494, Wuhan	China	30.63	114.07
58238, Nanjing	China	32.00	118.80
60390, Dar-el-beida	Algeria	36.72	3.25
61641, Dakar/Yoff	Senegal	14.73	-17.5
71600, Sable Island	Canada	43.93	-60.02

Programme DECC/Defra (GA01101), MoD (CBC/2B/0417 Annex C5).

## APPENDIX A

### Calculating Humidity Variables

Radiosondes measure relative humidity, but standard practice for reporting radiosonde data across the global telecommunications system has been for dewpoint measurements, so it is this variable that dominates the historical archives. Daily temperature and dewpoint [ $^{\circ}\text{C}$ ;  $T$  in Eq. (1) below] were used respectively to calculate saturated and actual vapor pressure  $e$  from a modified version of the Magnus formula for vapor pressure over liquid water (Buck 1981),

$$e = 6.1121 \times f \times \text{EXP} \left[ \frac{\left(18.729 - \frac{T}{227.3}\right)T}{257.87} \right]. \quad (\text{A1})$$

The enhancement factor ( $f$ ) requires pressure  $p$  (hPa) and is defined in Eq. (A2),

$$f = 1.0007 + 3.46 \times 10^{-6}p. \quad (\text{A2})$$

The relative humidity is defined as the ratio of the vapor pressure to the saturated vapor pressure. Vapor pressure calculated from dewpoint temperature was used to determine the specific humidity  $q$  ( $\text{g kg}^{-1}$ ) from Eq. (A3),

$$q = \frac{622e}{p - 0.378e}. \quad (\text{A3})$$

## APPENDIX B

### Case-Study Stations

To test the methods for interpolating missing data to reduce sampling bias, a set of 24 stations listed in Table B1 were used.

## APPENDIX C

### First-Guess Homogenization

Before applying the full homogenization, a first guess of break points resulting from known instrument changes was conducted. Table C1 lists the instrument changes that were identified as occurring in at least four stations within the radiosonde archive. For ease of reference

TABLE C1. Radiosonde instrument change events, approximate year(s) of implementation, the total number of stations affected, and the adjustment averaged across all available stations and pressure levels. Events have been separated by the principal country of origin. Changes of  $|T| \geq 0.5$  K,  $|\text{RH}| \geq 1\%$ , or  $|q| \geq 0.1$  g kg<sup>-1</sup> have been highlighted in bold.

Event	Year	No.	Temperature (K)		RH (%)		$q$ (g kg <sup>-1</sup> )		
			Day	Night	Day	Night	Day	Night	
United States									
Hypsometer to Viz Accu-Lok	1980–81	17	-0.3	-0.3	<b>+2.0</b>	<b>+2.1</b>	+0.06	+0.02	
Transponder type to Viz Accu-Lok	1981	11	-0.4	-0.2	<b>+2.7</b>	<b>+3.4</b>	<b>+0.12</b>	<b>+0.14</b>	
Viz Accu-Lok to Viz B	1988	62	+0.2	+0.2	-0.1	-0.3	+0.02	+0.01	
Viz (generic) to Viz B	1988	6	-0.1		+0.5		-0.01		
Viz B to Vaisala RS80–56	1995	24	-0.2	-0.0	-0.8	<b>-1.1</b>	-0.09	-0.06	
Viz B to Viz B2	1997–99	10	+0.2	+0.2	<b>+1.5</b>	<b>+1.4</b>	<b>+0.16</b>	<b>+0.25</b>	
United Kingdom									
Kew MkII to Kew MkIII	1978	5	-0.1	-0.2	<b>+2.6</b>	<b>+3.2</b>	-0.07	+0.01	
Kew MkIII to Vaisala RS80	1990	5	<b>+0.5</b>	+0.2	<b>-3.5</b>	<b>-4.3</b>	+0.00	+0.02	
Scandinavia and northern Europe									
Vaisala (generic) to Vaisala RS18	1975–77	7	+0.1	+0.1	<b>-4.6</b>	<b>-4.5</b>	<b>-0.19</b>	<b>-0.23</b>	
Vaisala RS18 to Vaisala RS21	1977–81	15	<b>+0.5</b>	-0.0	-0.7	<b>+1.9</b>	<b>+0.13</b>	+0.01	
Vaisala RS21 to Vaisala RS80	1981–88	35	+0.0	+0.2	<b>+2.7</b>	<b>+4.5</b>	+0.09	<b>+0.22</b>	
France									
Mesural FMO 1944C to 1944C, 1945A, or 1945B	1974	12	+0.0	+0.1	<b>+1.0</b>	<b>+1.0</b>	-0.06	+0.01	
1944C, 1945A, or 1945B to Mesural FMO 1950	1979	13	+0.3	+0.2	<b>+1.1</b>	+0.0	+0.09	+0.06	
Mesural FMO 1950 to Mesural FMO 1955	1983	13	-0.4	-0.4	<b>-2.5</b>	<b>-3.3</b>	<b>-0.23</b>	<b>-0.27</b>	
Mesural FMO 1955 to Vaisala RS80	1990	10	-0.3	+0.1	<b>-4.3</b>	<b>-2.7</b>	<b>-0.28</b>	<b>-0.17</b>	
Germany (excluding former East Germany)									
Graw M60 to Vaisala RS80	1988–90	6	+0.0	+0.2	<b>-6.9</b>	<b>-5.8</b>	<b>-0.16</b>	<b>-0.17</b>	
Eastern Europe and Russia (including former East Germany)									
A22 series to RKZ5	1973–79	24	-0.4	-0.2	+0.0	+0.9	-0.01	+0.00	
A22 series to MRZ	1987–90	18	<b>+0.9</b>	-0.1	-0.4	-0.5	+0.06	+0.07	
RKZ2 to RKZ5	1973–81	7	<b>-0.8</b>	+0.0	<b>+1.1</b>	+0.2	<b>-0.13</b>	+0.04	
RKZ2 to MARS	1983–86	28	-0.2	+0.2	<b>-1.4</b>	<b>-2.3</b>	-0.07	-0.06	
RKZ5 to MARS	1984–91	48	+0.1	+0.0	-0.9	+0.2	-0.03	+0.01	
MARS to MRZ	1988–91	39	<b>-0.7</b>	<b>-0.5</b>	-0.3	-0.2	-0.08	+0.00	
MARS to Vaisala RS80	1991	14	-0.3	-0.1	<b>-6.7</b>	<b>-5.8</b>	<b>-0.30</b>	<b>-0.24</b>	
Japan									
Meisei RSII-56 to Meisei RSII-80	1981	18	<b>-0.5</b>	-0.1	<b>-5.8</b>	<b>-5.1</b>	<b>-0.27</b>	<b>-0.22</b>	
Meisei RSII80 to Meisei RSII91	1994	17	+0.0	+0.2	<b>+1.4</b>	<b>+1.1</b>	+0.09	<b>+0.10</b>	
RSII_KC79 to RSII-KC96	1997	4	<b>+0.9</b>	<b>+0.9</b>	<b>+2.5</b>	<b>+1.3</b>	<b>+0.37</b>	<b>+0.36</b>	

the changes have been separated broadly into countries and the approximate year or years of the change are supplied. The values represent the average adjustment applied across the available pressure levels. For example, a value of +2.0% in relative humidity means that the monthly mean relative humidity anomalies prior to the break point are increased by an average value of 2.0%, but the actual adjustments applied at each level will differ from this value. The figures in Table C1 are meant to be illustrative of the sign and magnitude of the adjustments being made, rather than fully comprehensive.

#### REFERENCES

- Allan, R. P., V. Ramaswamy, and A. Slingo, 2002: A diagnostic analysis of atmospheric moisture and clear-sky radiative feedback in the Hadley Centre and Geophysical Fluid Dynamics Laboratory (GFDL) climate models. *J. Geophys. Res.*, **107**, 4329, doi:10.1029/2001JD001131.
- Allen, R. J., and S. C. Sherwood, 2008: Warming maximum in the tropical upper troposphere deduced from thermal winds. *Nature Geosci.*, **1**, 399–403.
- Bates, J. J., and D. L. Jackson, 2001: Trends in upper tropospheric humidity. *Geophys. Res. Lett.*, **28**, 1695–1698.
- Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, **110**, D12106, doi:10.1029/2005JD006548.
- Buck, A. L., 1981: New equations for computing vapor pressure and enhancement factor. *J. Appl. Meteor.*, **20**, 1527–1532.
- Colman, R., 2003: A comparison of climate feedbacks in general circulation models. *Climate Dyn.*, **20**, 865–873.
- Dai, A. G., 2006: Recent climatology, variability, and trends in global surface humidity. *J. Climate*, **19**, 3589–3606.
- Durre, I., R. S. Vose, and D. B. Wueertz, 2006: Overview of the Integrated Global Radiosonde Archive. *J. Climate*, **19**, 53–68.
- , C. N. Williams Jr., X. Yin, and R. S. Vose, 2009: Radiosonde-based trends in precipitable water over the Northern Hemisphere: An update. *J. Geophys. Res.*, **114**, D05112, doi:10.1029/2008JD010989.

- Elliott, W. P., 1995: On detecting long-term changes in atmospheric moisture. *Climatic Change*, **31**, 349–367.
- , and D. J. Gaffen, 1991: On the utility of radiosonde humidity archives for climate studies. *Bull. Amer. Meteor. Soc.*, **72**, 1507–1520.
- , R. J. Ross, and B. Schwartz, 1998: Effects on climate records of changes in National Weather Service humidity processing procedures. *J. Climate*, **11**, 2424–2436.
- Eskridge, R. E., O. A. Alduchov, I. V. Chernykh, P. Zhai, A. C. Polansky, and S. R. Doty, 1995: A Comprehensive Aerological Reference Data Set (CARDS): Rough and systematic errors. *Bull. Amer. Meteor. Soc.*, **76**, 1759–1775.
- Gaffen, D. J., T. P. Barnett, and W. P. Elliott, 1991: Space and time scales of global tropospheric moisture. *J. Climate*, **4**, 989–1008.
- Guichard, F., D. Parsons, and E. Miller, 2000: Thermodynamic and radiative impact of the correction of sounding humidity bias in the tropics. *J. Climate*, **13**, 3611–3624.
- Guo, Y., P. W. Thorne, M. P. McCarthy, H. A. Titchner, B. Huang, Z. Panmao, and Y. Ding, 2008: Radiosonde temperature trends and their uncertainties over eastern China. *Int. J. Climatol.*, **28**, 1269–1281.
- Gutzler, D. S., 1992: Climatic variability of temperature and humidity over the tropical western Pacific. *Geophys. Res. Lett.*, **19**, 1595–1598.
- Haimberger, L., 2007: Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate*, **20**, 1377–1403.
- Held, I. M., and B. J. Soden, 2000: Water vapour feedback and global warming. *Annu. Rev. Energy Environ.*, **45**, 441–475.
- John, V. O., and B. J. Soden, 2007: Temperature and humidity biases in global climate models and their impact on climate feedbacks. *Geophys. Res. Lett.*, **34**, L18704, doi:10.1029/2007GL030429.
- Joshi, M. M., J. M. Gregory, M. J. Webb, D. M. H. Sexton, and T. C. Johns, 2008: Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. *Climate Dyn.*, **30**, 455–465.
- 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. Climatol.*, **16**, 1197–1226.
- , S. A. Klein, and D. J. Seidel, 2003: Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology. *J. Climate*, **16**, 224–240.
- Luers, J. K., and R. E. Eskridge, 1998: Use of radiosonde temperature data in climate studies. *J. Climate*, **11**, 1002–1019.
- McCarthy, M. P., 2008: Spatial sampling requirements for monitoring upper-air climate change with radiosondes. *Int. J. Climatol.*, **28**, 985–993.
- , and R. Toumi, 2004: Observed interannual variability of tropical troposphere relative humidity. *J. Climate*, **17**, 3181–3191.
- , H. A. Titchner, P. W. Thorne, S. F. B. Tett, L. Haimberger, and D. E. Parker, 2008: Assessing bias and uncertainty in the HadAT adjusted radiosonde climate record. *J. Climate*, **21**, 817–832.
- Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 630–635.
- Randel, W. J., and F. Wu, 2006: Biases in stratospheric and tropospheric temperature trends derived from historical radiosonde data. *J. Climate*, **19**, 2094–2104.
- Remsburg, E. E., and Coauthors, 2000: Instrumentation and data sets. SPARC Assessment of Upper Tropospheric and Stratospheric Water Vapour, SPARC Rep. WCRP-113, WMO/TD No. 1043, 33–36.
- Ross, R. J., and W. P. Elliott, 1996: Tropospheric water vapor climatology and trends over North America: 1973–93. *J. Climate*, **9**, 3561–3574.
- , and —, 2001: Radiosonde-based Northern Hemisphere tropospheric water vapor trends. *J. Climate*, **14**, 1602–1612.
- , —, and D. J. Seidel, 2002: Lower-tropospheric humidity–temperature relationships in radiosonde observations and atmospheric general circulation models. *J. Hydrometeorol.*, **3**, 26–38.
- Santer, B. D., J. E. Penner, and P. W. Thorne, 2006: 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*, T. R. Karl et al., Eds., Climate Change Science Program and the Subcommittee on Global Change Research, 89–118.
- , and Coauthors, 2007: Identification of human-induced changes in atmospheric moisture content. *Proc. Natl. Acad. Sci. USA*, **104**, 15 248–15 253.
- Sapucci, L. F., L. T. Machado, R. B. Da Silveira, G. Fisch, and J. F. G. Monico, 2005: Analysis of relative humidity sensors at the WMO radiosonde intercomparison experiment in Brazil. *J. Atmos. Oceanic Technol.*, **22**, 664–678.
- Sherwood, S. C., 2007: Simultaneous detection of climate change and observing biases in a network with incomplete sampling. *J. Climate*, **20**, 4047–4062.
- , J. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late-20th-century warming. *Science*, **309**, 1556–1559, doi:10.1126/science.1115640.
- Shine, K., and A. Sinha, 1991: Sensitivity of the Earth’s climate to height-dependent changes in the water vapour mixing ratio. *Nature*, **354**, 382–384.
- Soden, B. J., and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean–atmosphere models. *J. Climate*, **19**, 3354–3360.
- , and J. R. Lanzante, 1996: An assessment of satellite and radiosonde climatologies of upper-tropospheric water vapor. *J. Climate*, **9**, 1235–1250.
- , D. L. Jackson, V. Ramaswamy, M. D. Schwarzkopf, and X. L. Huang, 2005: The radiative signature of upper tropospheric moistening. *Science*, **310**, 841–844.
- Thorne, P. W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P. Brohan, 2005a: Revisiting radiosonde upper air temperatures from 1958 to 2002. *J. Geophys. Res.*, **110**, D18105, doi:10.1029/2004JD005753.
- , —, J. R. Christy, and C. A. Mears, 2005b: Uncertainties in climate trends: Lessons from upper-air temperature records. *Bull. Amer. Meteor. Soc.*, **86**, 1437–1442.
- Titchner, H. A., P. W. Thorne, M. P. McCarthy, S. F. B. Tett, L. Haimberger, and D. E. Parker, 2009: Critically re-assessing tropospheric temperature trends from radiosondes using realistic validation experiments. *J. Climate*, **22**, 465–485.
- Trenberth, K. E., J. Fasullo, and L. Smith, 2005: Trends and variability in column-integrated atmospheric water vapour. *Climate Dyn.*, **24**, 741–758.
- Wade, C. G., 1994: An evaluation of problems affecting the measurement of low relative humidity on the United States radiosonde. *J. Atmos. Oceanic Technol.*, **11**, 687–700.

- Wang, J., H. L. Cole, and D. J. Carlson, 2001: Water vapour variability in the tropical western Pacific from 20-year radiosonde data. *Adv. Atmos. Sci.*, **18**, 752–766.
- , D. J. Carlson, D. B. Parsons, T. F. Hock, D. Lauritsen, H. L. Cole, K. Beierle, and E. Chamberlain, 2003: Performance of operational radiosonde humidity sensors in direct comparison with a chilled mirror dew-point hygrometer and its climate implication. *Geophys. Res. Lett.*, **30**, 1860, doi:10.1029/2003GL016985.
- Willett, K. M., N. P. Gillett, P. D. Jones, and P. W. Thorne, 2007: Attribution of observed surface humidity changes to human influence. *Nature*, **449**, 710–712, doi:10.1038/nature06207.
- , P. D. Jones, N. P. Gillett, and P. W. Thorne, 2008: Recent changes in surface humidity: Development of the HadCRUH dataset. *J. Climate*, **21**, 5364–5383.
- Zhai, P., and R. E. Eskridge, 1997: Atmospheric water vapor over China. *J. Climate*, **10**, 2643–2652.