Supplementary Material

1) Quality Control of the Hourly Data

The quality control process for HadISDH.2.0.0 is identical to that described for HadISDH.landq.1.0.0 in Willett et al. 2013 (Sect. 2 and 3.1) and fully described in Dunn et al. (2012). This includes three humidity specific tests for supersaturation, wick drying events and issues with static or removed values during extremes of temperature. Of the 3679 stations selected for HadISDH.2.0.0, for T_d, in total 79.7 % of stations have ≤ 1 % of hourly data removed and 98.8 % of stations have ≤ 5 % of hourly data removed. For T, 85.9 % of stations have ≤ 1 % of hourly data removed and 98.5 % of stations have ≤ 5 % of hourly data removed. Thus T_d removal percentages exceed those for T, and this is especially so in the tropics.

2) Climatology Selection Criteria

For each station, monthly-hour means are taken for each hour of the day where at least 15 days are present for that hour within the month (e.g., Jan-1976-00hrs, Jan-1976-01hrs etc.). For climatological monthly-hour means there must be at least 15 years of data within the 1976-2005 period for each month of the year and at least 1 year in each decade (1976-1985, 1986-1995, 1996-2005). Climatological monthly means are calculated from climatological monthly-hour means only when there are at least four climatological monthly-hour means across the diurnal cycle, with at least one in each eight hour tercile (00:00–08:00, 08:00–16:00, 16:00–24:00 UTC) of the day.

3) Monthly Mean Calculation

Monthly means are calculated from the hourly data by first taking the monthly-hour mean for each hour of the day where there are at least 15 days present within the month. Similarly to the climatology, for a monthly mean to be calculated there must be at least four monthly-hour means per day, with at least one in each eight hour tercile (00:00–08:00, 08:00–16:00, 16:00–24:00 UTC) of the day. This prevents biasing towards night or day, or biases arising from systemically changing observation times aliasing into the record.

4) Pairwise Homogenisation Algorithm (PHA) Overview

The PHA (Menne and Williams 2009) detects changepoints by comparing each station with every other station within its neighbour network. These networks contain up to 40 stations with the highest correlations of first differences of climate anomalies with the target station. Changepoints are assigned to a station and time point by identifying statistical breaks in paired series and then deconvolving the matrix of returned breaks to ascertain the station from which each breakpoint originates. The neighbours are then used to estimate the size of the identified inhomogeneities. In cases where the adjustment can be identified sufficiently confidently an adjustment is applied relative to the more recent homogeneous subperiod. Results of station removals during PHA (and ID PHA) processing, and summary adjustment statistics are listed for each variable in Table 1.

5) Application of PHA on T and DPD

First we homogenise T and DPD directly with PHA. DPD has been chosen over T_d for initial PHA homogenisation because applying PHA directly to DPD and then calculating T_d results in fewer cases of unphysical supersaturation than applying PHA to climate anomalies of T and T_d (3% of stations versus 5% of stations). There is little difference in the signal-to-noise ratio between DPD and T_d – changepoints are detected slightly more frequently in DPD than for T_d (2.53 versus 2.16 changepoints per station). Comparison of applying PHA directly to T_d as opposed to deriving T_d from homogenised T-DPD shows very little difference. While the largest individual gridbox trends occur when T_d is derived, the directly homogenised regional average trends (significant) are larger (by 0.01 °C decade^{-1}) for the globe, Northern Hemisphere and tropics. Over the poorer sampled Southern Hemisphere neither methods result in significant trends, varying from -0.03 °C decade^{-1} to 0.01 °C decade^{-1}. Of most interest is the change from positive trends in direct PHA T_d to negative trends in derived T_d over coastal southwestern USA. In general, the derived T_d shows a larger area of significant drying compared to using PHA directly.

Homogenisation statistics for T and DPD are shown in Table 1. During direct PHA, 3 and 12 stations are removed from T and DPD respectively because there are too few highly correlating neighbours to constitute a neighbour network. Further stations are removed for two reasons. Firstly, after direct PHA, which removes subperiods of data when an adjustment size cannot be robustly estimated, some stations no longer have sufficient data to calculate a climatology. Secondly, for all variables except T there is a possibility of adjustments resulting in physically unrealistic values in terms of saturation (i.e., RH should not go below 0 %rh or above 100 %rh). For DPD (and T_d as implicated by DPD), any stations where DPD goes below 0 ℃ are removed because this implies that $T_d > T$. Although this supersaturation is physically possible especially in very cold climates (Makkonen and Laakso, 2005), it is uncommon and in this case very likely an artefact of seasonally invariant adjustments. Any such occurrences lead to removal of the entire station.

6) Application of ID PHA on *T*, *Td*, *Tw*, *q*, *e* and RH

For each station all changepoints are obtained from the PHA logs for T and DPD for that station and sorted chronologically. PHA does not allow two changepoints to fall within 12 months of each other and so for ID PHA any changepoint lying within 12 months of an earlier changepoint is ignored. A neighbour network is established for each station from the 40 (maximum, with at least 7) highest correlating neighbours based on the first difference series of the anomalies. Stations with fewer than 7 neighbours are removed at this stage, in addition to the 12 stations removed during direct PHA on T and DPD, and the 6 stations removed for containing extremely large inhomogeneities (Supp. Mat. Table 1). A distribution of potential adjustments is obtained by calculating the difference in medians between the two homogeneous sub-periods surrounding the changepoint for each candidate minus neighbour pair within the network. The median of the distribution becomes the adjustment size and the 1.65σ uncertainty is obtained from the 5th to 95th percentiles, in keeping with direct PHA.

Although q and e are derived from hourly T_d only, changepoints from T as well as DPD must be used because if there is a simultaneous changepoint with an inhomogeneity of the same size in T and T_d then it would not appear in DPD. Also, given historical practices at many synoptic stations in reality the hourly T_d provided by ISD (and then HadISD) is mostly likely to have originated as T_w or even RH and then been converted which means that T, and any inhomogeneities in it, will have been incorporated at some stage in the processing chain. In practice, T also undergoes a second homogenisation using ID PHA with changepoints from DPD to apply suspected missing changepoints.

Homogenisation statistics are shown in Supp. Mat. Table 1. Although no data are removed during ID PHA, a few stations have insufficient monthly means to calculate a climatology (see Sect. 2) and so are now removed. Secondly, due to seasonally invariant adjustments there are cases where adjusted values are physically implausible. For T_w, each month is compared with homogenised T for the same month, and cases where $T_w >$ T are considered supersaturated. For RH, q and e, if the simultaneous homogenised RH value is greater than 100 %rh then all values are considered to be supersaturated. For q, e and RH, a value should never go below zero. Any supersaturated or sub-zero (in the case of q, e and RH) occurrences lead to removal of the entire station. There are a very high number of supersaturation occurrences for T_w resulting in a much smaller number of stations passing through to be part of the gridded product. It is clear that further work is needed to improve the method in future versions in terms of consistency between T and T_w.

7) A Model for Estimating Uncertainty at the Station Level

The uncertainty estimate for the station provides an individual value for each monthly mean anomaly based on uncertainty in the climatology calculation given missing data (clim), in any homogenisation adjustments made and missed (adj) and also from the initial hourly measurement (ob). All components are described fully in Willett et al. (2013) and so only the differences necessitated by the multivariable approach will be discussed here. To briefly recap, for any monthly mean anomaly our model is as follows (note Eq. 1 to 7 are in Table 1 of the Main Text):

$$u_{anom} = \sqrt{u_{clim}^2 + u_{adj}^2 + u_{ob}^2} \qquad (1)$$

7.1 Inhomogeneity Uncertainty

The inhomogeneity uncertainty is slightly more complex with the addition of ID PHA. In all cases it is a combination of both the uncertainty on any adjustments applied and the uncertainty for any inhomogeneities that likely remain within the data:

$$u_{adj} = \sqrt{u_{applied}^2 + u_{missed}^2} \qquad (2)$$

For direct PHA (and ID PHA, retaining consistency with PHA) a 1.65σ uncertainty range is provided for each adjustment and this is used here after transformation to 1σ standard uncertainty for consistency with the other components. For T and T_d the uncertainty estimates arising from the adjustments applied from ID PHA and direct PHA are combined chronologically. The uncertainty in missed adjustments is spatio-temporally static for each variable. It is obtained by assuming that the 'missing middle' should be filled (Figure 5, Main Text). This is done as described in Willett et al. (2013). A Gaussian curve is fitted to the middle of the actual distribution. Then a best fit distribution is created by selecting the higher of the Gaussian and the actual distribution. In practice, this takes in the wider tails of the actual distribution and the Gaussian middle which infills the 'missing middle' of the actual distribution. Differences are then taken between the merged best fit and the actual distribution (blue dotted lines in Figure 5, Main Text). The standard deviation of this difference becomes the missed adjustment uncertainty. Despite ID PHA resulting in a smaller 'missing middle', the estimated uncertainty from missed adjustments is similar. For T, T_w, T_d, e, q, RH and DPD these uncertainty estimates are 0.25 °C, 0.17 °C, 0.26 °C, 0.19 hPa, 0.14 g kg^{-1}, 1.00 %rh and 0.26 °C respectively.

7.2 Measurement uncertainty

Following the BIPM Guide to the Expression of Uncertainty in Measurement (BIPM, 2008), measurement uncertainties are assigned as belonging to one of two categories. Type A evaluation of uncertainties can be estimated from statistical analysis of multiple observations from the same population. Type B evaluation of uncertainties can be estimated from *a priori* knowledge of the measurement apparatus and the measuring conditions. As in Willett et al. (2013), only Type B uncertainties are used here.

Over a month, type B uncertainties can have randomly varying components, u_{rand}, whose effect can be reduced by averaging, and components which cause “systematic” errors, u_{sys} where the effect is not reduced by averaging. We do not provide a specific estimate for u_{sys} because this is accounted for by homogenisation and u_{adj} to some extent. Note that this assumption is valid when using the anomalies. However, for the absolute values, our estimated uncertainty is likely to be an underestimate. This is because the most recent homogeneous sub-period may not be truly representative of the ambient humidity or temperature depending on the instrument, shelter and observation quality. Similarly, a station with no changepoints may still be unrepresentative of the true ambient climate while retaining the accurate long-term trend.

To obtain u_{rand} (the measurement uncertainty estimate in the monthly mean) we have to assume that the point measurement (e.g., hourly) uncertainty is the same for every point measurement within the month. This is because it is not sensible to use the unhomogenised raw hourly values at this point, and the PHA method can only be applied to monthly data. For the point measurement uncertainty estimate u_i we use a standard uncertainty of 0.15 °C in the wet bulb depression above the ice point (assumed to be 0 °C in this case), extrapolating this below this point where necessary (Willett et al. 2013), using the assumption that all measurements are taken with an aspirated psychrometer. These are converted to the appropriate variable using Eq. 1 to 7 in Table 1. The resulting standard uncertainty in RH varies from 1 %rh to 3 %rh, decreasing with increasing T and increasing with decreasing RH (NPL/IMC, 1996). In addition, a 0.2 °C uncertainty in T (after Brohan et al. 2006) is used as the estimated u_i for T.

For T and T_w, u_i directly transfers to 0.2 °C and 0.15 °C respectively for all months. For all other variables u_i will vary depending on the temperature and humidity at each time step. To provide an example, Table 2 shows the estimated u_i for all variables at a range of dry bulb temperatures where we assume saturation (for ease of calculation). To obtain u_{rand}, these values need to be converted to an estimate over the monthly mean by dividing by the square root of the number of measurements made across a month. In the worst-case scenario this should be at least 4 measurements per day within at least 15 days resulting in N≥60:

$$u_{rand} = \frac{u_i}{\sqrt{N}} \qquad (3)$$

For conversion of the u_i in T_w to a u_i specific to the each of the other humidity variables, homogenised monthly RH and T are used in addition to the monthly value of that variable. For q and e, first the saturation equivalents q_s and e_s are calculated for each month by rearranging Eq. (7, Main Text) with RH set at 100 %rh. Then, using the actual monthly RH a combined RH value plus RH uncertainty RH+ΔRH is created by adding a change in RH dependent on the simultaneous monthly T as shown in Table 2. We then calculate the monthly value plus uncertainty q+Δq (or e+Δe) from the RH+ΔRH and the saturated value q_s (e_s) using a further rearrangement of Eq. (7, Main Text). Finally, the original q (or e) value is subtracted from the q+Δq (or e+Δe) value to obtain the u_i for that month. For T_d, the monthly value is converted to e using Eq. (3, Main Text) and e+Δe is calculated as described above. Then $T_d + \Delta T_d$ is calculated from e+Δe using a rearrangement of Eq. (3, Main Text). Subtracting T_d from $T_d + \Delta T_d$ provides the u_i value. In this case we have not used Eq (4, Main Text) (with respect to ice) when T is less than 0 °C which means that our uncertainty estimate in these cases will be slightly overestimated. For DPD, a simple addition of the u_i for T_d for that month and the 0.2 °C u_i for T is used. This is a conservative assumption. As the spatio-temporal coverage is not identical across all variables, in cases where there are no simultaneous monthly values for RH and/or T, values of 80 %rh and 0 °C are used respectively. For all variables, u_{rand} *is estimated from each* u_i *using* $N=60$ (Eq. 3).

Table 1. Homogenisation statistics for all variables. Note that a further 6 stations are removed for all variables due to very large adjustments being found during PHA for T (714963, 729595, 719410, 026720, 718260, 535880).

Variable	**station count (start, finish)**	**No neighbour stations (IDs)**	**Missing data stations**	**Subzero error stations**	**Supersaturation error stations**	**µ change-point frequency per station**	**Absolute IH size µ**	**Absolute IH size σ**	**IH size µ**	**IH size σ**	**5 largest IHs (ID, size)**
***T* (℃)**	3679, 3561	3+2 (+9 from DPD) (854690, 910660, 919250, 085010**, 896110**)	98	0	0	3.56 (PHA: 1.34)	0.38 (PHA: 0.73)	0.46 (PHA: 0.57)	-0.02 (PHA: -0.10)	0.60 (PHA: 0.94)	723783 -4.58, 718440 4.57, 023260 -4.44, 599480 -4.05, 766480 -3.96
DPD (℃)	3679, 3451	9 (689940, 847820, 854880, 889680, 890020, 890220, 916100, 916430, 919430)	213	0	149	2.52	0.98	0.69	-0.01	1.21	556640 7.17, 614970 7.11, 621030 7.05, 442840 6.90, 942380 -6.34
T_d (℃)	3667*, 3164	2 (085010**, 896110**)	348	0	147	3.62	0.77	0.65	-0.01	1.01	556640 -7.27, 614970 -6.76, 726548 6.72,

T_w (°C)	3667*, 2831	5 (085010**, 689060**, 689060**, 911650**, 911760**)	17	0	808	3.75	0.31	0.32	-0.01	0.45	621030 - 6.63, 442840 - 6.57 711820 4.80, 023260 - 4.34, 718440 3.60, 022860 - 3.34, 723783 - 3.31
q (g kg^{-1})	3667*, 3558	6 (085010**, 689060**, 895320**, 895710**, 896110**, 911650**)	17	52	28	3.76	0.26	0.29	-0.01	0.39	765770 3.64, 403400 - 3.31, 412400 - 3.30, 765770 3.06, 822810 - 2.89
RH (%rh)	3667*, 3607	9 (681040**, 787670**, 841170** 854700**, 859300**, 889630**, 895649**, 895710**, 939450**)	16	0	28	3.75	2.88	2.19	0.01	3.62	369820 - 27.30, 765770 21.58, 726548 21.33, 442770 - 20.78, 014060 - 18.46

***e* (hPa)**	3667*, 3558	6 (085010**, 689060**, 895320**, 895710**, 896110**, 911650**)	17	52	28	3.76	0.41	0.46	-0.01	0.61	403400 -5.33, 412400 -5.19, 822810 -4.58, 765770 4.57, 412680 -4.50

* 12 stations are removed from initial count because they failed PHA for T and DPD because they had no neighbour stations.

** Stations removed during indirect PHA because there were fewer than 7 neighbour stations.

Table 2. Estimates of standard uncertainty in humidity measurements calculated in terms of equivalent psychrometer uncertainty to represent a “worst case scenario”. All uncertainties are based on a 0.15 ℃ uncertainty in wet bulb depression. Estimates are made for each value assuming saturation at the given temperature and comparing with a value at RH equal to (100 %rh minus the associated uncertainty in RH) for that temperature band. For DPD, the 0.2 C uncertainty in dry bulb temperature is added linearly to the uncertainty in dew point temperature as in a worst case scenario the error in T and T_w would oppose. Calculations of specific humidity used equations from Table 1.

T (℃)	**Uncertainty in RH (%rh)**	**Uncertainty in q (g kg^{-1}) at saturation**	**Uncertainty in e (hPa) at saturation**	**Uncertainty in T_d (℃)**	**Uncertainty in DPD (℃)**
-50 and below	15	0.004	0.006	1.309	1.509
-40	15	0.012	0.019	1.428	1.628
-30	15	0.035	0.057	1.553	1.753
-20	10	0.064	0.104	1.094	1.294
-10	5	0.081	0.131	0.577	0.777
0	2.75	0.105	0.169	0.338	0.538
10	1.8	0.138	0.223	0.271	0.471
20	1.35	0.199	0.318	0.219	0.419
30	1.1	0.298	0.470	0.193	0.393
40	0.95	0.459	0.707	0.179	0.379
50+	0.8	0.672	1.000	0.162	0.362

Table 3 Description of data contained in the HadISDH CF-compliant netCDF file. Gridboxes are 5 ° by 5 ° beginning with centres of -177.5 °W, -87.5 °S. Times count from 1 on January 1973 to 492 on December 2013. Missing data are recorded as -1e30.

Field	**Description**	**Dimension**	**Maximum and Minimum values**						
			T (°C)	**T_w (°C)**	**T_d (°C)**	**q (g kg^{-1})**	**e (hPa)**	**RH (%rh)**	**DPD (°C)**
abs	Monthly mean	72,36,492	-53.18, 39.72	-47.93, 28.51	-54.41, 27.42	0.00, 23.57	0.00, 36.22	4.78, 98.80	0.02, 35.7
anoms	Monthly mean anomaly (seasonal cycle averaged over 1976-2005 climatology removed)	72,36,492	-16.18, 18.21	-16.02, 12.43	-17.37, 16.25	-6.28, 7.16	-10.04, 11.47	-33.62, 38.81	-14.40, 14.80
std	Standard deviation of all monthly mean anomalies within the gridbox for each month	72,36,492	0.00, 18.10 (100.00)	0.00, 15.72 (100.00)	0.00, 23.92 (100.00)	0.00, 9.36 (100.00)	0.00, 17.27 (100.00)	0.00, 40.72 (100.00)	0.00, 15.50 (100.00)
combinederr	Station uncertainty and sampling uncertainty combined in quadrature to give a 2σ uncertainty	72,36,492	0.07, 4.74	0.07, 4.02	0.08, 5.56	0.04, 3.51	0.06, 5.62	0.33, 27.83	0.07, 5.60

samplingerr	2σ spatial sampling error	72,36,492	0.01, 3.45	0.00, 3.14	0.01, 4.00	0.00, 0.75	0.01, 1.85	0.07, 23.63	0.00, 3.53
rbar	Average inter-site correlation	72,36	0.10, 0.92	0.10, 0.92	0.10, 0.91	0.10, 0.92	0.10, 0.92	0.10, 0.91	0.10, 0.89
sbarSQ	Estimate of mean variance of individual stations in the gridbox	72,36	0.08, 14.79	0.07, 13.43	0.08, 16.77	0.03, 10.00	0.07, 10.00	1.71, 73.90	0.08, 12.55
stationerr	Climatological, adjustment and measurement uncertainty combined in quadrature to give a 2σ uncertainty	72,36,492	0.04, 4.24	0.05, 3.47	0.04, 5.03	0.02, 3.50	0.03, 5.61	0.29, 20.71	0.04, 4.34
adjerr	2σ adjustment uncertainty	72,36,492	0.03, 4.03	0.04, 2.81	0.03, 4.83	0.02, 3.49	0.03, 5.59	0.20, 20.02	0.03, 3.77
obserr	2σ measurement uncertainty	72,36,492	0.00, 0.05	0.00, 0.04	0.00, 1.50	0.00, 0.13	0.00, 0.21	0.02, 3.87	0.01, 1.55
climerr	2σ climatological uncertainty	72,36,492	0.02, 2.45	0.02, 2.04	0.01, 2.85	0.00, 1.15	0.00, 1.83	0.08, 5.43	0.01, 2.83
clims	Monthly climatologies over the 1976-	72,36,12	-46.08, 37.72	-40.76, 27.39	-48.08, 25.91	0.04, 21.88	0.08, 33.66	16.15, 94.48	0.73, 31.08

	2005 period								
mean_n_stations	Total number of stations within the gridbox over the entire record (e.g., contributing to climatology)	72,36	0, 44	0, 36	0, 38	0, 44	0, 44	0, 44	0, 39
actual_n_stations	Actual number of stations in the gridbox for each month	72,36,492	0, 44	0, 36	0, 37	0, 44	0, 44	0, 44	0, 38