Supplemental Material for Historical Trends and Variability in Heat Waves in the United Kingdom

Michael G. Sanderson [1,*], Theo Economou [1,2], Kate H. Salmon [1] and Sarah E. O. Jones [1,3]

[1] Met Office, Exeter EX1 3PB, UK
[2] College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK
[3] JBA Consulting, Skipton BD23 1FJ, UK

S1 Recording types of weather station data

Three different observation reporting types were used for the stations in Table A1: DLY3208, AWSDLY and NCM. DLY3208 refers to data reported on Met Form 3208b on a monthly basis, which contains up to 31 daily observations. The reporting times were usually 09:00, although other times were reported for some stations. Observation stations within the AWSDLY (Automatic Weather Station Daily values) Station Network transmit their observations in the form of an AWSDLY message. AWSDLY messages carry daily and 12 hour information. National Climate Messages (NCM), produced from all Met Office and many auxiliary stations, contain data for 12-hour periods (usually 09:00-21:00 and 21:00-09:00).

For the majority (22 out of 29) of the weather stations in Table A1, the observations were reported under a single observation type (DLY3208 or NCM). For the remaining stations, data up to around 1999 were reported once per day on form 3208. Between 1999 and 2001 (depending on the station), the reporting type changed to 'AWSDLY' and data were reported either every 12 or 24 hours. This change usually indicated automation of the instrumentation; in a few cases it also implied a change in the instrumentation. The first stage of processing for all stations was to ensure that the daily maximum temperatures were assigned to the correct dates. For example, if daily maximum temperatures were noted on a given date, and the reporting period was the previous 24 hours, the daily maximum temperature would have occurred on the previous day.

Temperatures recorded under reporting types DLY3208 and AWSDLY overlapped at 4 stations, but significantly so at only 2 stations, Rothamstead and Wisley (Table S1). In these two cases, it appears that two sets of instruments were operated simultaneously. There was no overlap between changes in reporting type at any other stations in Table A1. The partial homogenization of the temperatures recorded at Rothamstead and Wisley is described in Section S2.

Table S1. Stations where different observation types overlapped by 1 or more days.

Station	Number of obs types[a]	Overlap (days)[b]	Dates of Overlap[c]	Offset / °C[d]
Durham	2	2	19.09.1999 – 23.09.1999	0.05[e]
London, St James Park	3	1	29.09.1999 – 30.09.1999	0.0[e]
Rothamstead	3	1490	29.09.1999 – 31.12.2003	0.1[f]
Wisley	2	238	27.06.2000 – 31.07.2001	0.4[f]

[a] The observation types were 'DLY3208', 'AWSDLY' and 'NCM', and appeared in that order.

[b] The number of days for which data were reported under both observation types 'DLY3208' and 'AWSDLY'. Temperature recorded under AWSDLY and NCM did not overlap.

[c] Dates over which the two observations types occurred together; data from both types were not available on every day within the range shown.

[d] A positive value indicates that temperatures reported as 'AWSDLY' are warmer than those reported as 'DLY3208'.

[e] The offset is the mean difference between the two reporting types on the days of overlap.

[f] The offset is the intercept from a linear regression analysis of the overlapping data, using a method which is robust to outliers; see text for details.

S2: Partial homogenisation of weather station data

This section describes the procedures used to partially homogenize the weather station data (where necessary) resulting from changes in reporting type, where such a change appeared to correspond to a change in instrumentation. All the processing was carried out using programs written in Python. The linear regression analyses were performed using the robust linear model (RLM) method (Venables and Ripley, 2002) which has been implemented in the 'statsmodels' package (http://statsmodels.sourceforge.net/documentation.html). It should be noted that the corrections to temperatures calculated below for the various weather stations are small, of the order of a few tenths of a degree. The data from the 29 stations in Table 1 have not been fully homogenized as the data records have different lengths (Table 1). It seems unlikely that temperatures from the stations would be altered by more than a few tenths of a degree if they were fully homogenized, and so little impact on the results of this study is expected.

S2.1 Rothamstead

Daily maximum temperatures at Rothamstead were reported under all three observation types (DLY3208, AWSDLY and NCM). Data were reported under type DLY3208 up to 31st December 2003, and from 29th September 1999 under type AWSDLY. Temperatures recorded under both observation types were available on 1490 days (Table S1). To check for any differences, data under AWSDLY were plotted as a function of DLY3208, after sorting them into ascending order (Figure S2.1, upper panel). A straight line was fitted to the data using the RLM. The intercept from the linear fit is the difference between the two datasets. The two datasets are highly correlated, and the mean offset is 0.1°C. Hence, on average, temperatures recorded under AWSDLY are 0.1°C warmer than those recorded under DLY3208.

Next, the differences (residuals) between the two datasets were plotted against their Julian day numbers (Figure S2.1, lower panel). A straight line was fitted to the residuals to check for any time-dependence, i.e., whether the offset between the two reporting types is constant or varying over time.

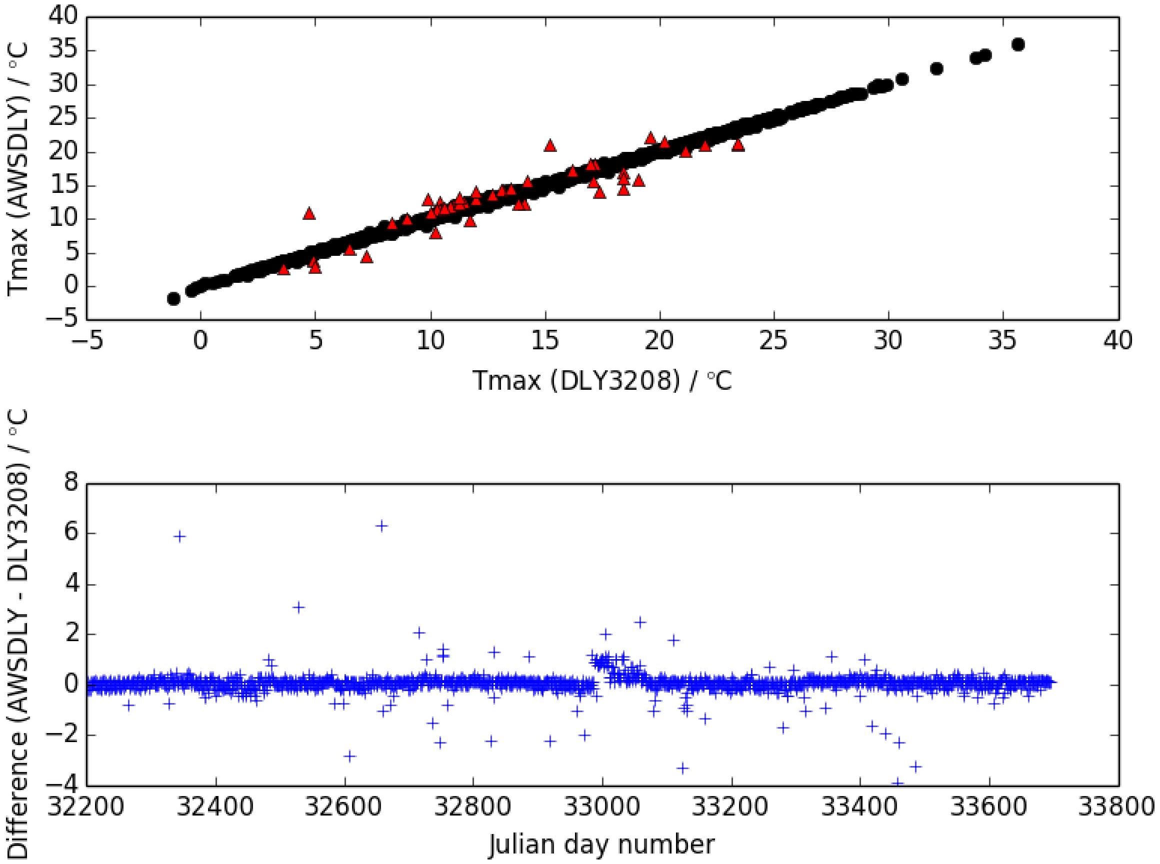

Figure S2.1. Analysis of daily maximum temperatures recorded at Rothamstead. Upper panel shows temperatures recorded under AWSDLY as a function of temperatures under DLY3208. Data shown using red triangles are where the differences relative to the offset are larger than ± 1°C. Lower panel shows the differences between the two datasets as a function of their Julian day numbers. The offset (intercept from the upper panel) is 0.1°C.

Differences in temperatures between the two reporting types were large (> 1°C) on a small number of dates (Figure S2.1, lower panel). Temperatures recorded on those dates and on 2 days either side were compared with temperatures from a nearby station (Oxford) on the same dates. It was found that temperatures recorded at Rothamstead under DLY3208 on these days were similar to those recorded at Oxford on the same days. In contrast, larger differences were noted for data at Rothamstead recorded under AWSDLY. The ‘spikes’ seen in the AWSDLY data were not seen at Oxford. The decision was made to use temperatures recorded under DLY3208 from Rothamstead in preference to AWSDLY. The offset (0.1°C) was subtracted from temperatures recorded under AWSDLY.

S2.2 Wisley

The procedure described in section S2.1 for Rothamstead was followed for temperatures recorded at Wisley. Daily maximum temperatures recorded under DLY3208 and AWSDLY were available on 238 days (Table A1), and are highly correlated, with a mean difference of 0.4°C (Figure S2.2, upper panel). The trend in the residuals with time was very close to zero and was not significant (Figure S2.2, lower panel).

As for Rothamstead, temperatures recorded under DLY3208 were used in preference to those recorded under AWSDLY. The offset (0.4°C) was subtracted from temperatures recorded under AWSDLY.

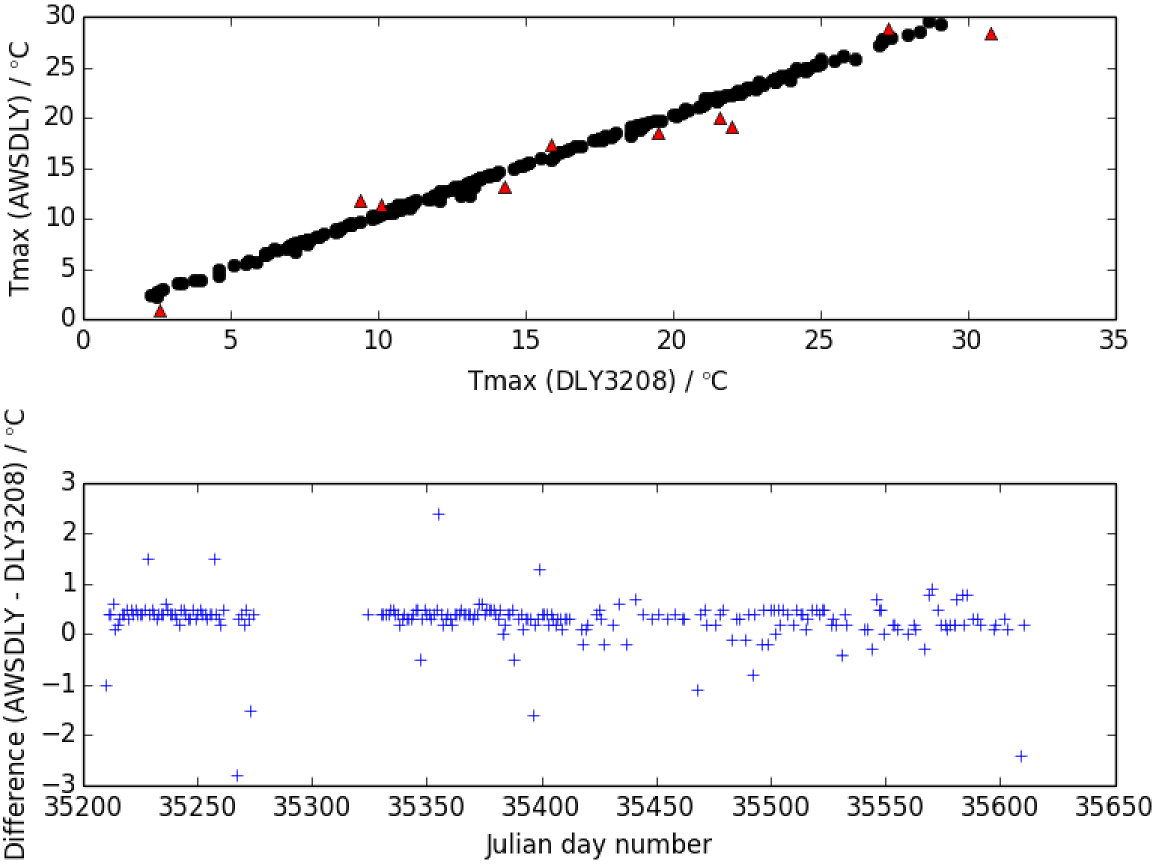

Figure S2.2. Partial homogenisation of daily maximum temperature data recorded at Wisley. Upper panel shows temperatures recorded under AWSDLY as a function of temperatures under DLY3208. Data shown using red triangles are where the differences relative to the offset are larger than ±1°C. Lower panel shows the difference between the two datasets as a function of their Julian day numbers. The median difference is 0.4°C.

S2.3 Sheffield

Daily weather observations have been made at a weather station in the grounds of Weston Park Museum, Sheffield, since 1882. Information about and photographs of this station are available at: http://lifeofdata.org.uk/node/weston-park-weather-station/.

Until 2000, the data were recorded manually by staff of the Weston Park museum, after which the museum's equipment was automated. During 2008, the Met Office installed its own equipment which transmits weather information every 10 minutes. This event is apparent in the raw data extracted from MIDAS, when the reporting type changed from DLY3208 to AWSDLY at the end of September 2008. The museum's and the Met Office's observing equipment are located within the same Stevenson's screen, and data from the two sets of equipment are compared regularly and any errors are noted and corrected. It is therefore considered unlikely that there would have been a large shift in the temperatures recorded when the reporting type changed during 2008.

S3 Combined Stations

In two cases, data from two or three weather stations were combined to make a single continuous series. This procedure was necessary where one station had closed and another opened in close proximity. The particular stations which were combined and the methods used are described below.

S3.1 Plymouth: Mountbatten, Mount Wise and Hoe

Daily maximum temperatures are available from the weather station at Plymouth Mountbatten (MB) from 30 November 1930 to the present day (Table S2). However, no data were recorded at MB between 1 Nov 1946 and 1 July 1948. A weather station at Mount Wise (MW), located on the opposite side of the Plym estuary about 3.5 km to the west, was operational during this period. Additionally, daily temperatures were recorded at Plymouth Hoe between 1874 and 1930, and 1959 to 1980. It is desirable to (a) infill the missing period of data from MB with the data from MW, (b) check for any clear shifts in temperatures recorded at MB pre-November 1946 and post-July 1948 (henceforth referred to as pre-1947 and post-1947), and (c) combine MB and Hoe to create a single series for the period 1874-2016 (Table S3.1).

The data from MB and MW do not overlap so a direct comparison is not possible. The closest station which was operational between 1946 and 1948 was Torquay Abbey Park (TAP), located approximately 43 km to the north-east and 450 m inland (Table S3.1). The only other station which was operational on the south west coast during this period was Falmouth. A comparison of Falmouth with MW and MB showed that the correlations were very similar. However, a small positive trend (significant at the 5% level) was found in the differences between Falmouth and Mount Wise, suggesting a possible small shift in climate between these two sites. The trend in the differences between MW and TAP was small and not significant at the 5% level. Daily maximum temperatures from TAP will therefore be used as a reference for combining temperatures from MB and MW. A weather station was located between MB and MW on Plymouth Hoe, but unfortunately no data are available from this station between 1946 and 1948 (Table S3.1).

Table S3.1. Stations used for partial homogenization of data for Plymouth Mountbatten

Station	Easting	Northing	Altitude / m	Data available
Plymouth Hoe	247800	53700	36	01.01.1874 - 30.11.1930
				01.01.1959 - 30.03.1980
Mount Wise	245900	54000	22	01.11.1946 - 01.06.1948
Plymouth Mountbatten	249216	52710	50	30.11.1930 - 31.12.2016
Torquay Abbey Park	290700	63900	15	01.01.1912 - 30.07.2005

The procedure used to partially homogenize daily maximum temperatures (Tmax) for Plymouth Mountbatten is described below. In all regressions, Tmax data from TAP were used as the independent variable, and Tmax data from MW or MB were the dependent variable.

Daily Tmax data from MW and TAP were found to be highly correlated, with a gradient of 0.987 and an intercept of 0.104. There was no trend in the differences between these two stations, indicating no drift in climate between them.

Next, data from MB pre- and post-1947 need to be compared with TAP. The period chosen over which to compare the data is arbitrary, but was a compromise between having sufficient data to establish a robust correlation against any drift caused by localised changes in climate at the two sites. An initial comparison using 10 years or more of Tmax data suggested that the relationship between Tmax at MB and TAP had changed slightly over time, possible due to increased urbanisation around MB and rebuilding following heavy damage incurred during the Second World War. The five year periods indicated below were a compromise. Periods from November to June were used to match the dates from MW.

Tmax from MB and TAP recorded between 1 November 1941 and 1 June 1946 were found to be correlated. Temperatures at MB were generally warmer than those at TAP. The slope and intercept are slope_mb_tap_pre and int_mb_tap_pre respectively.

Tmax from MB and TAP recorded between 1 November 1948 and 1 June 1953 were found to be correlated. Temperatures at MB were again generally warmer than those at TAP. The slope and intercept were slope_mb_tap_post and int_mb_tap_post respectively.

The gradients and intercepts for the pre-1947 and post-1947 periods were different, suggesting that either the position of the Mountbatten station was changed, or the characteristics of the surrounding areas at either TAP or MB (or both) had changed. The majority of the data from MB are available post-1947; hence, Tmax data from MW and MB pre-1947 were adjusted to match MB post-1947 as follows:

Tmax from MW were first adjusted to match the climate at TAP, and then scaled to match MB post-1947. Hence, for the period 1 November 1946 to 1 June 1948:

$$\text{Tmax(MB)} = (\text{Tmax(MW)} - \text{int\_mw\_tap}) \times (\text{slope\_mb\_tap\_post} / \text{slope\_mw\_tap}) + \text{int\_mb\_tap\_post}$$

These adjusted data were then inserted into the Tmax series from MB, to replace the period of missing data.

The Tmax data from MB pre-1947 were first adjusted to match the climate at TAP, and then scaled to match MB post-1947. If Tmax(PRE) are daily Tmax data from MB recorded before 1947, then:

$$\text{Tmax(MB)} = (\text{Tmax(PRE)} - \text{int\_mb\_tap\_pre}) \times (\text{slope\_mb\_tap\_post} / \text{slope\_mb\_tap\_pre}) + \text{int\_mb\_tap\_post}$$

This procedure produced a partially homogenized series of daily maximum temperatures for Plymouth Mountbatten for the period 1 November 1930 to 31 December 2016.

Daily maximum temperatures are available from Hoe between 1874 and 1930; given the close proximity of Hoe and MB, it was considered reasonable to join the temperatures from these two stations to effectively extend the record from MB back to 1874. Daily maximum temperatures from MB and Hoe were linearly correlated between 1959 and 1979. The slope and intercept for this period were assumed to be valid for the period 1874-1930, when temperatures from Hoe are also available (Table S3.1). The slope and intercept were used to scale the temperatures from Hoe for 1874-1930, to match the climate at MB. The scaled temperatures were then prepended to the series from MB. Overall, a scaled and partially homogenized data series for MB between 1874 and 2016 was created.

S3.2 Shoeburyness

A weather station was operated close to the coast at Shoeburyness (a coastal town in Essex located at the mouth of the Thames) between 1930 and 1981. A replacement station (Shoeburyness Landwick) was situated further along the coast and slightly further inland, from which data are available from May 1983 (Table S3.2). The nearest two coastal stations operation between 1978 and 1986 were Clacton-on-Sea (CS) and Margate (MAR) (Table S3.2). The methods used to (a) join the data from Shoeburyness (SH) and Shoeburyness Landwick (SLW), and (b) infill the gap (01.01.1982 - 30.04.1983) are described below.

Table S3.2. Stations used for partial homogenization of data for Shoeburyness

Station	Easting	Northing	Altitude / m	Data available
Shoeburyness	594800	185700	2	30.11.1930 - 30.12.1981
Shoeburyness Landwick	596079	187772	2	01.05.1983 - 31.12.2016
Clacton-on-Sea	617200	214300	16	01.01.1959 - 30.06.1995
Margate	636800	171400	16	31.12.1888 - 30.06.1988

Daily maximum temperatures from SH between 1977 and 1981 were closely correlated with both CS and MAR; similarly, daily maximum temperatures from SLW between 01 June 1983 and 31 May 1988 were correlated with CS and MAR. An analysis of the differences between SH and MAR, and SLW and MAR as a function of their Julian day numbers showed that there was a residual seasonal cycle; however, the differences between SH and CS and SLW and CS were more randomly distributed and no seasonal cycle was apparent (data not shown). Daily maximum temperatures from CS were therefore used to partly homogenize the temperatures from Shoeburyness and Shoeburyness Landwick and join them to create a single series. Temperatures from SLW were scaled to match those at SH.

The gap between SH and SLW (01.01.1982-30.030.1983) was infilled by scaling the temperatures from CS, using the linear coefficients:

$$\text{Tmax(SH)} = \text{int\_sh\_cs} + \text{slope\_sh\_cs} \times \text{CS}$$

The temperatures recorded at SLW were scaled to match the climate at SH by first adjusting to match the climate at CS, and then back to SH:

$$\text{Tmax(SH)} = (\text{Tmax(SLW)} - \text{int\_slw\_cs}) \times (\text{slope\_sh\_cs} / \text{slope\_slw\_cs}) + \text{int\_sh\_cs}$$

Overall, a partially homogenized series of daily maximum temperatures for Shoeburyness between 30 November 1930 and 31 December 2016 was produced.

S4 Sources of sea surface temperature and sea level pressure from which the AMO and summertime NAO series were derived.

Table S4: Sources of sea surface temperature (SST) and sea level pressure (SLP) data from which the AMO and summertime NAO indices were derived. The Reference column lists the publications which describe the underlying SST and SLP datasets.

Data source	Region	Data Period	Reference
SSTs			
HadSST2[a]	25°-60°N, 7°W-70°W	1874-2016	Rayner et al. (2006)
ERSST[a]	25°-60°N, 7°W-70°W	1880-2016	Smith et al. (2008)
Kaplan SST[b]	0°-70°N,	1856-2016	Kaplan et al. (1998)

HadSST3[c]	10°N-60°N, 10°W-70°W	1870-2016	Kennedy et al. (2011)
SLPs			
TP1980	40°N-70°N, 90°W-30°E	1899-2016	Trenberth and Paolino (1980)
20CR	40°N-70°N, 90°W-30°E	1871-2011	Compo et al. (2011)

[a] Calculation of the AMO as described in van Oldenborgh et al. (2009).

[b] Calculation of the AMO described in Enfield et al. (2001).

[c] AMO derived from HadSST3 by J.J. Kennedy (personal communication).

References for Table S4

Compo G.P. (2011). The Twentieth Century Reanalysis Project. Q.J.R. Meteorol. Soc., 137: 1–28. doi: 10.1002/qj.776

Enfield, D.B., A.M. Mestas-Nunez, and P.J. Trimble (2001). The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental U.S. Geophys. Res. Lett., 28: 2077-2080.

Kaplan, A., M. Cane, Y. Kushnir, A. Clement, M. Blumenthal, and B. Rajagopalan (1998) Analyses of global sea surface temperature 1856-1991, Journal of Geophysical Research, 103, 18,567-18,589.

Kennedy, J. J., N. A. Rayner, R. O. Smith, D. E. Parker, and M. Saunby (2011), Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties, J. Geophys. Res., 116, D14103, doi:10.1029/2010JD015218.

Rayner NA, Brohan P, Parker DE, Folland CK, et al. (2006) Improved analyses of changes and uncertainties in marine temperature measured in situ since the mid-nineteenth century: the HadSST2 dataset. J Climate 19:446-469.

Smith, T.M., R.W. Reynolds, T.C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean temperature analysis (1880–2006). Journal of Climate, 21, 2283–2296.

Trenberth, K. E., and D. A. Paolino Jr., 1980: The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. Mon. Wea. Rev., 108, 855-872.

van Oldenborgh, G. J., te Raa, L. A., Dijkstra, H. A., and Philip, S. Y.: Frequency- or amplitude-dependent effects of the Atlantic meridional overturning on the tropical Pacific Ocean, Ocean Sci., 5, 293-301

S5 Trends in day of year of last heat wave day and heat wave characteristics

Table S5.1. Trends in the day of year of the last heat wave day for 1961-2016 and three percentile-based thresholds. The trends have units of days per decade, and those shown in bold are significant at the 5% level. Positive values indicate that heat waves are occurring later in the year.

Station	93rd	95th	98th
Aldergrove	-0.386	-2.495	-1.206
Armagh	0.236	-1.409	-1.444
Balmoral	**6.062**	-0.304	-2.109
Bradford	0.562	-1.549	-0.658
Cambridge	**5.262**	**4.343**	-1.286
Craibstone	3.632	2.410	2.615
Cranwell	**4.291**	**5.969**	-1.554
Cromer	0.403	**5.992**	**11.603**
Douglas	**7.290**	4.962	1.874
Durham	2.441	-0.199	-0.392
Eastbourne	**6.773**	**5.495**	0.129
Lerwick	**3.096**	**5.546**	2.395
Leuchars	1.013	1.967	0.808
London St James Park	2.954	1.655	0.197
Morecambe	-0.369	2.955	-4.284
Morpeth	2.591	4.124	-0.594
Newton Rigg	2.527	0.858	1.741
Oxford	1.015	-2.566	-1.349
Plymouth Mountbatten	4.144	1.251	-2.622
Rothamstead	2.071	1.632	-1.906
Sheffield	-1.754	-2.588	-1.070
Shoeburyness	**6.591**	**4.918**	0.069
Skegness	**6.832**	**5.093**	4.077
Stornoway Airport	3.045	-0.048	-2.694
Teignmouth	**4.946**	**5.459**	0.000
Tiree	3.362	0.712	-3.603
Valley	-5.287	-3.221	-7.035
Wick Airport	2.942	0.645	0.852
Wisley	2.676	3.340	**-5.895**

Trends in heat wave numbers (HWN), longest heat waves (HWD) and sum of heat wave lengths (i.e., total number of heat wave days, HWF) are shown in Tables S5.2-S5.4 respectively. The trends were calculated over two time periods, 1931-2016 and 1961-2016. The heat waves were identified using three different thresholds derived using the 93rd, 95th and 98th percentiles of year-round temperatures recorded over 1971-2000. The trends were calculated using a method robust to outliers. Trends shown in bold are significant at the 5% level.

Table S5.2. Trends in numbers of heat waves (HWN) for two time periods and three percentile-based thresholds. The trends have units of numbers of heat waves per decade, and those shown in bold are significant at the 5% level. The thresholds were calculated using daily maximum temperatures for 1971-2000.

	1931 - 2016			1961 - 2016		
Station	93rd	95th	98th	93rd	95th	98th
Aldergrove	0.050	0.060	0.086	0.134	0.025	0.036
Armagh	0.060	0.021	0.008	0.047	-0.084	-0.025
Balmoral	0.039	-0.006	-0.023	0.144	-0.100	-0.098
Bradford	0.157	0.070	-0.020	**0.303**	0.044	-0.022
Cambridge	0.082	0.042	-0.003	0.302	0.196	0.006
Craibstone	0.024	-0.006	0.018	0.263	0.037	-0.037
Cranwell	0.165	0.010	-0.115	0.286	0.192	-0.103
Cromer	-0.023	-0.052	-0.048	0.288	0.150	0.000
Durham	-0.077	-0.037	-0.025	0.200	0.007	0.003
Eastbourne	**0.375**	0.140	0.055	**0.883**	**0.505**	0.101
Douglas	0.085	0.004	0.075	**0.488**	0.186	0.167
Lerwick	-0.057	0.068	0.029	**0.376**	**0.352**	0.129
Leuchars	0.099	0.016	0.000	0.239	0.064	0.000
London St James's Park	0.099	0.027	0.032	0.234	0.055	0.043
Morecambe	0.022	0.032	-0.041	0.000	0.098	-0.050
Morpeth	0.054	0.045	-0.067	**0.358**	0.100	0.025
Newton Rigg	0.056	-0.057	0.003	**0.359**	0.126	0.037
Oxford	0.142	0.010	-0.038	0.322	0.060	-0.030
Plymouth Mountbatten	0.112	0.031	0.070	**0.295**	0.155	-0.049
Rothamstead	0.106	-0.011	-0.005	0.144	0.053	-0.089
Sheffield	0.046	-0.011	-0.038	0.114	-0.058	-0.080
Shoeburyness	0.021	-0.009	-0.044	**0.358**	0.210	-0.005
Skegness	**0.300**	0.153	0.000	**0.678**	**0.268**	0.000
Stornoway Airport	0.100	-0.065	0.049	**0.460**	0.036	0.012
Teignmouth	**0.229**	0.062	0.045	**0.529**	0.201	0.142
Tiree	-0.101	0.097	-0.029	**0.333**	0.115	-0.085
Valley	0.097	0.073	-0.081	-0.078	0.015	-0.011
Wick_airport	**0.258**	0.120	0.027	**0.450**	0.095	0.073
Wisley	0.057	0.026	-0.011	0.236	0.115	-0.038

Table S5.3. Trends in longest heat waves (HWD) for two time periods and three percentile-based thresholds. The trends have units of days per decade, and those shown in bold are significant at the 5% level.

	1931 - 2016			1961 - 2016		
Station	93rd	95th	98th	93rd	95th	98th
Aldergrove	0.133	-0.012	0.137	0.399	-0.184	0.148
Armagh	0.168	0.027	-0.010	0.403	-0.217	-0.001
Balmoral	-0.041	0.066	-0.003	0.266	0.207	-0.047
Bradford	0.252	0.070	-0.149	0.444	0.175	-0.458
Cambridge	0.242	0.115	0.078	**0.673**	**0.421**	-0.107
Craibstone	0.098	0.030	0.044	0.311	0.173	0.036
Cranwell	0.250	0.011	-0.065	**0.896**	0.338	-0.018
Cromer	0.182	0.089	0.035	**0.572**	**0.362**	0.242
Durham	0.033	-0.026	0.144	0.353	0.199	0.070
Eastbourne	0.358	0.122	0.103	**1.135**	0.491	0.095
Douglas	0.252	-0.004	-0.037	0.559	0.134	-0.001
Lerwick	0.204	0.233	0.147	0.621	**0.588**	0.449
Leuchars	0.002	0.057	0.044	0.200	0.159	0.075
London_St James Park	0.231	0.219	0.076	**0.773**	**0.709**	-0.307
Morecambe	0.065	0.011	0.066	0.144	-0.147	-0.087
Morpeth	0.153	0.158	0.040	**0.722**	0.473	0.115
Newton Rigg	0.146	0.036	-0.041	0.543	0.140	-0.020
Oxford	0.192	0.157	0.114	**0.808**	**0.624**	-0.060
Plymouth_Mountbatten	0.255	-0.008	0.252	0.600	0.187	0.204
Rothamstead	0.203	0.078	0.144	**0.650**	**0.485**	-0.141
Sheffield	0.163	0.109	-0.096	**0.565**	0.295	-0.251
Shoeburyness	0.134	0.181	-0.013	**0.906**	**0.621**	0.007
Skegness	**0.299**	**0.210**	0.018	**0.498**	0.289	0.120
Stornoway Airport	0.195	0.158	**0.184**	**0.684**	0.503	0.393
Teignmouth	**0.417**	**0.381**	0.161	**1.083**	**0.804**	0.614
Tiree	0.313	0.340	0.102	1.157	**0.711**	0.344
Valley	0.190	0.119	-0.017	0.102	0.058	0.105
Wick Airport	**0.198**	0.104	0.059	0.228	-0.043	0.103
Wisley	0.012	0.091	0.042	**0.791**	**0.543**	-0.174

Table S5.4. Trends in sums of heat wave lengths (HWF) for two time periods and three percentile-based thresholds. The trends have units of days per decade, and those shown in bold are significant at the 5% level.

	1931 - 2016			1961 - 2016		
Station	93rd	95th	98th	93rd	95th	98th
Aldergrove	0.549	0.134	0.418	1.217	-0.148	0.144
Armagh	0.536	0.144	0.043	0.749	-0.681	-0.036
Balmoral	0.146	0.093	-0.050	0.790	-0.214	-0.223
Bradford	0.885	0.280	-0.343	1.545	0.399	-0.734
Cambridge	0.853	0.380	0.092	**2.297**	1.280	-0.556
Craibstone	0.363	0.079	0.070	1.356	0.295	-0.203
Cranwell	**1.060**	0.162	**-0.506**	**2.328**	1.160	-0.216
Cromer	0.282	-0.081	-0.080	1.917	0.966	0.344
Durham	-0.275	0.035	0.005	1.285	0.344	-0.063
Eastbourne	**2.080**	0.646	0.277	**5.157**	**2.431**	0.444
Douglas	0.508	0.026	0.215	**2.330**	1.003	0.631
Lerwick	-0.052	0.595	0.224	**2.460**	**2.151**	1.085
Leuchars	0.557	0.176	0.117	1.189	0.427	0.202
London_St James Park	0.777	0.494	0.181	**1.890**	1.200	-0.289
Morecambe	0.204	0.159	-0.074	0.127	0.191	-0.235
Morpeth	0.393	0.387	-0.141	**2.232**	0.989	0.265
Newton Rigg	0.505	-0.131	0.009	**2.304**	0.779	0.130
Oxford	0.758	0.325	-0.039	**2.155**	1.105	-0.245
Plymouth_Mountbatten	0.669	0.168	0.599	1.962	0.777	0.074
Rothamstead	0.767	0.198	0.174	1.463	0.887	-0.498
Sheffield	0.341	0.166	-0.241	1.101	0.148	-0.527
Shoeburyness	0.549	0.287	-0.137	**3.258**	**1.845**	-0.052
Skegness	**1.585**	**0.804**	0.046	**3.120**	**1.264**	0.196
Stornoway Airport	0.436	0.112	**0.451**	**2.378**	0.877	0.593
Teignmouth	**1.486**	0.697	0.460	**3.589**	**1.782**	0.901
Tiree	-0.278	0.719	0.042	2.492	1.138	0.105
Valley	0.542	0.381	-0.323	-0.155	0.122	0.014
Wick Airport	**1.305**	0.520	0.138	**2.176**	0.376	0.349
Wisley	0.344	0.268	0.002	**1.881**	1.232	-0.369

S6: Logistic modelling of AMO/NAO effects on heat wave lengths

The following figures show the mean effects (across eight weather stations with records beginning in the nineteenth century) of the AMO and summer NAO on the probabilities of 3/4 or more heat waves in a year, and the probabilities of the longest heat waves in each year being 8/10 days or longer. There are eight data points for each threshold, as there are four AMO datasets and two summer NAO datasets resulting in eight possible combinations.

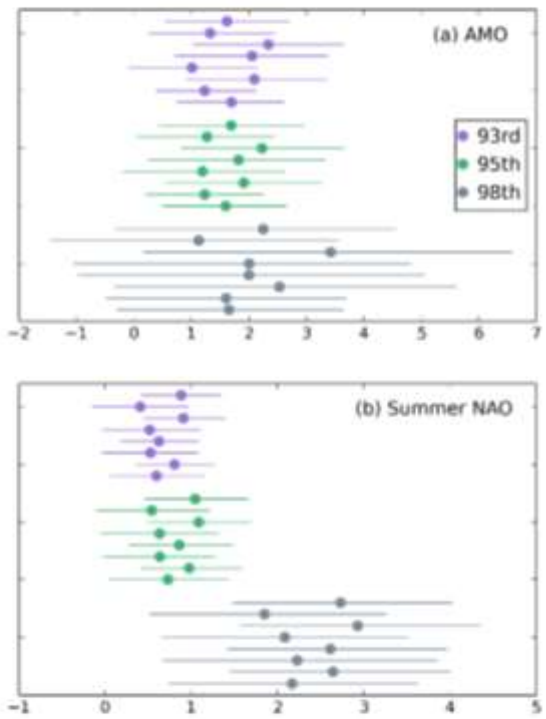

Figure S6.1 Effect of AMO and summertime NAO on probability of heat waves of 8 days or longer in a year. Data shown are mean effects across eight weather stations for three different percentile thresholds. (a) The AMO effects are fairly consistent across the percentile threshold choice and AMO-summer NAO dataset pairs. Some of the error ranges include zero, especially for the 98th percentile threshold. (b) The effects for the summer NAO are consistent for the 93rd and 95th percentile thresholds but those for 98th percentile are notably larger.

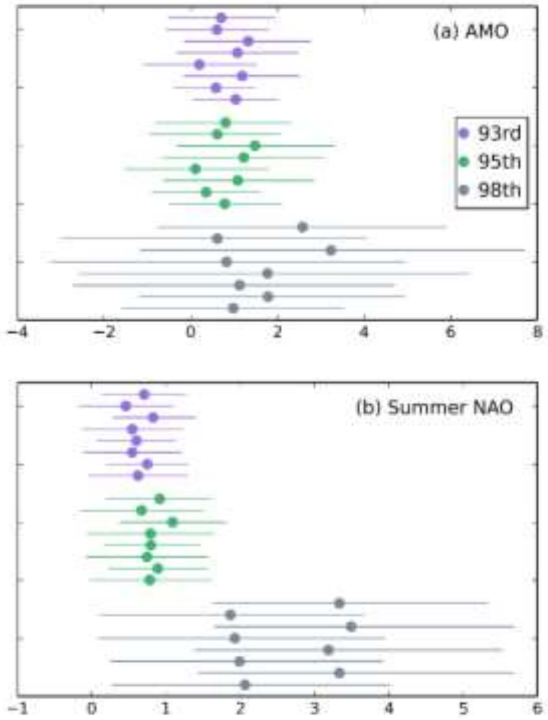

Figure S6.2 Effect of AMO and summertime NAO on probability of heat waves of 10 days or longer in a year. Data shown are mean effects across eight weather stations for three different percentile thresholds. (a) The AMO effects are fairly consistent across the percentile threshold choice and AMO-summer NAO dataset pairs. More of the error ranges include zero compared with Figure 6.1(a), but most of the masses of the distributions are above zero. (b) The effects for the summer NAO are consistent for the 93rd and 95th percentile thresholds but those for 98th percentile are notably larger.