

Detection of anthropogenic influence on a summertime heat stress index

Thomas R. Knutson¹ · Jeffrey J. Ploshay¹

Received: 5 August 2015 / Accepted: 20 May 2016 / Published online: 2 June 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract One of the most consequential impacts of anthropogenic warming on humans may be increased heat stress, combining temperature and humidity effects. Here we examine whether there are now detectable changes in summertime heat stress over land regions. As a heat stress metric we use a simplified wet bulb globe temperature (WBGT) index. Observed trends in WBGT (1973–2012) are compared to trends from CMIP5 historical simulations (eight-model ensemble) using either anthropogenic and natural forcing agents combined or natural forcings alone. Our analysis suggests that there has been a detectable anthropogenic increase in mean summertime heat stress since 1973, both globally and in most land regions analyzed. A detectable increase is found over a larger fraction of land for WBGT than for temperature, as WBGT summertime means have lower interannual variability than surface temperature at gridbox scales. Notably, summertime WBGT over land has continued increasing in recent years—consistent with climate models—despite the apparent ‘hiatus’ in global warming and despite a decreasing tendency in observed relative humidity over land since the late 1990s.

1 Introduction

One of the most consequential future impacts of long-term climate warming could be the impact of heat stress on humans. Various approximate measures of heat stress have been developed in the medical and human health communities (Epstein and Moran 2006). Here we analyze a simplified form of the wet bulb globe temperature (WBGT), which is a linear combination of air temperature (T_a) and wet bulb temperature (T_w): $WBGT = 0.7 T_w + 0.3 T_a$. This form is appropriate for indoor conditions and neglects, for example, the effects of direct

Electronic supplementary material The online version of this article (doi:10.1007/s10584-016-1708-z) contains supplementary material, which is available to authorized users.

✉ Thomas R. Knutson
Tom.Knutson@noaa.gov

¹ Geophysical Fluid Dynamics Laboratory / NOAA, Princeton, NJ 08540, USA

sunlight on heat stress (Epstein and Moran). T_w alone has also been used as a simpler index of human heat stress in previous studies. For climate warming, a human health concern is that the core body temperature of humans is a fixed threshold (37 °C or 98.6 °F), whereas climate warming tends to elevate measures of heat stress, such as WBGT, over time. For WBGT, there is a resulting reduced differential between the ambient WBGT and the human skin temperature. During periods of heat stress, the smaller this differential, the more difficult it becomes for the human body to cool itself during periods of exertion, or in extreme cases even while the body is at rest (Yaglou and Minard 1957).

As practical illustrations of this issue, Epstein and Moran (2006; Appendix) present a review of health guidelines for working or exercising under various WBGT conditions. For example, U.S. military studies with heat-acclimated soldiers indicate that for a WBGT range of 25.6–27.7 °C, sustained hard work performance for at least 4 h can be maintained using a cycle of 40 min work, 20 min of rest. However, if WBGT is ~6 °C higher (>32.2 °C) the guidelines are much more stringent: cycles of 10 min work, 50 min rest. The American College of Sports Medicine (1984) recommends delaying/rescheduling distance running events, if WBGT exceeds 28 °C. Physical activity and health for members of the general population, including the elderly, infirm, etc. would presumably be affected to an even greater degree.

The WBGT, wet bulb temperature, and dewpoint temperature all include both temperature and moisture influences on heat stress, similar to the heat index or apparent temperature indices (Steadman 1984). The WBGT, wet bulb temperature, and dewpoint temperature are indices which, as they approach the human body skin temperature, signify increasing difficulty for the body to cool itself down. The apparent temperature or heat index is, in contrast, a “feels-like” index, where the temperature index is elevated above the regular air temperature to reflect the effect that moisture has in making the temperature feel hotter than it actually is.

A number of studies point to a growing concern over increasing heat stress in the 21st century as a result of human-caused global warming, particularly when moisture, as well as temperature, effects are considered. Delworth et al. (1999) found that the greenhouse gas-induced increases in heat index, or apparent temperature, substantially exceed the increases in temperature alone, particularly in humid regions of the tropics and subtropics. Sherwood and Huber (2010) proposed that heat stress imposes an upper bound on how much global warming humans can adapt to, since with large global warming, T_w in some regions can approach levels (35 °C) that should induce hyperthermia in humans. Willett and Sherwood (2012), in a study of past and future-projected WBGT trends in 15 regions, found positive historical trends (1973–2003) in almost all regions studied. Their statistical model suggested that, assuming a uniform 2 °C warming above present levels, a 35 °C WBGT threshold would be exceeded during at least some extreme events in almost all 15 regions. In a study of future heat stress projections for WBGT, Dunne et al. (2013) estimated that under the Representative Concentration Pathway (RCP) 8.5 scenario, global labor capacity could be reduced by 60 % by the year 2200 in peak heat stress months. Pal and Eltahir (2015) identified parts of the Arabian Gulf region as an area where a T_w threshold of 35 °C could be reached soon this century under business-as-usual emission scenarios. Fischer and Knutti (2013) found that for future climate change projections, uncertainties in heat stress metrics that include both temperature and humidity jointly are typically smaller than the uncertainties in the two variables analyzed independently. Though not including moisture effects on heat stress, future projections of increasing heat wave occurrence and intensity have also been reported associated with rising temperatures (Meehl and Tebaldi 2004; Fischer and Schär 2010, for Europe; Lau and Nath 2012, for North America; and Cowan et al. 2014, for Australia).

Despite their likely importance for future health impacts, relatively little analysis has been reported on the detection/attribution of increases in heat stress metrics that include humidity effects. Gaffen and Ross (1998) and Grundstein and Dowd (2011) documented significant increasing trends in apparent temperature extremes over much of the continental US since 1949. Though not an attribution study, Schär (2016) notes an occurrence of T_w of 34.6 °C in Bandar Mahshahr (Iran) on 31 July 2015.

There is much more previous work on detection/attribution of anthropogenic contributions to heat stress increases due to temperature alone. Examples include observed increases in high temperature extremes (Fischer and Knutti 2015), or to the probability of occurrence of extreme heat wave events such as the 2003 European heat wave (Stott et al. 2004), the Russian heat wave of 2010 (Rahmstorf and Coumou 2011; but see also Dole et al. 2011 and Otto et al. 2012), the Australian heat waves during summer of 2012/2013 (Perkins et al. 2014), or extreme seasonal or annual Australian temperatures (e.g., Lewis and Karoly 2013; Knutson et al. 2014).

A detectable increase in observed surface specific humidity due to anthropogenic influences on climate has been reported (Willett et al. 2007) based on a land/ocean combined observational dataset. Recently, a decrease in surface relative humidity over land regions has been reported (Simmons et al. 2010; Willett et al. 2014a), which, together with the ongoing global warming ‘hiatus’ (e.g. Fyfe et al. 2013), raises questions on whether there has been any detectable anthropogenic influence on summertime heat stress.

In this study, we expand on previous climate change detection studies, based on temperature, heat waves, or specific humidity, to ask whether there is a detectable anthropogenic influence on a WBGT, which includes both temperature and humidity influences. We focus on land regions during summer months—June–August in the northern hemisphere and December–February in the southern hemisphere—typically the season of maximum climatological heat stress.

2 Methodology

Heat stress indices have been reviewed in Epstein and Moran (2006). Aside from behavioral factors (clothing and activity levels), the key environmental factors affecting human heat stress include ambient temperature, the amount of radiation (e.g., direct sunlight adds to heat stress), environmental humidity, and windspeed. The WBGT formulation from the classic heat casualties study of Yaglou and Minard (1957) was given as $WBGT = 0.7T_w + 0.3T_g$, where T_w is the wet bulb temperature as measured by a sling psychrometer, and T_g is the globe temperature measured by a thermometer placed inside a 6-inch (15 cm) diameter copper globe with a black matte painted exterior. WBGT is sometimes defined as $WBGT = 0.7T_w + 0.1T_a + 0.2T_g$, where T_a is the dry bulb temperature measured by standard thermometer (e.g. Epstein and Moran 2006, Eq. 3). As they noted, for indoor conditions, $T_g \sim T_a$, and the even simpler approximation $WBGT = 0.7T_w + 0.3T_a$ can be used, regardless of which of the above original expressions for WBGT is used. Here we use the simplified form ($WBGT = 0.7T_w + 0.3T_a$), and therefore our analysis applies to mean heat stress levels for fully shaded/sheltered conditions, neglecting solar insolation or wind effects, and does not address peak heat stress levels which are amplified by the diurnal cycle of temperature, solar insolation, etc. This simplified index approach reflects our focus on changes in heat stress changes associated with well-established anthropogenically induced changes in large-scale environmental factors (surface temperature

and specific humidity). We do not consider potential anthropogenically induced changes in solar radiation, windspeed, or diurnal cycles of temperature and specific humidity. Further, we do not assess several localized heat stress influences caused by humans such as urban heat island effects, black pavement, or other land surface modifications that can locally affect heat stress. Regarding the diurnal cycle of temperature, Hartmann et al. (2013) assess only medium confidence in reported decreases in the diurnal temperature range, with potential for biases to affect previously reported results. They note that the reported changes in diurnal temperature range are much smaller than mean temperature changes, and from this we infer that diurnal cycle changes would likely have only a secondary impact on heat stress trends. Nonetheless, an extension of our analysis to the case of daily maximum heat stress levels would be a worthwhile future study.

We use surface specific humidity from the HadISDH observational data set version HadISDH.1.0.0.2012p (Willett et al. 2014b) covering 1973–2012 (<http://www.metoffice.gov.uk/hadobs/hadisdh/>) primarily over land regions. Willett et al. discuss a number of sampling issues with the humidity data. As we will show, trends in WBGT appear to be dominated by temperature (warming) trends as opposed to changes in relative humidity, from which we propose that humidity data quality issues are unlikely to have an important influence on our main findings. Using such a short record (1973–2012) is a limitation for detection/attribution studies. For example, in an earlier similar study for surface temperature trends alone, Knutson et al. (2013) find that the percent of analyzed global area with a detectable trend (ending in 2010) is about twice as great for trends beginning around 1900 as for trends beginning around 1970. However, the short record limitation is necessary in our study due to the limitations of long-term surface humidity data over land regions.

We analyze a simplified WBGT index derived from monthly mean specific humidity and temperature data from either HadISDH or climate models. Before creating WBGT anomalies, we combine the observed anomalies of temperature and specific humidity with climatological values of specific humidity and temperature from HadISDH to create WBGT absolute values. We also use surface pressure climatological (1973–2012) and monthly values from the NCEP/NCAR Reanalysis (Kalnay et al. 1996) for our calculations. The same procedures are followed for the climate models: combining monthly mean temperatures, specific humidity, and surface pressure to create monthly mean WBGT time series.

The T_w component of the WBGT requires further method description. The specific formulas used for computing T_w , and sample T_w calculations, are documented in [Supplemental Material](#). Since T_w is a nonlinear function of temperature and moisture, using monthly mean values of temperature and specific humidity to compute a monthly mean T_w is an approximation compared to averaging high frequency (sub-monthly-scale) T_w data directly. In [supplemental material](#), we show that the effects of our approximations on T_w trends are relatively minor compared to the long-term trends, justifying the use of the simplified WBGT index derived from monthly mean temperature and moisture data.

Our trend assessment methodology follows that used for surface temperature by Knutson et al. (2013). We use a subset of available Coupled Model Intercomparison Project 5 (CMIP5) models (Taylor et al. 2012). We selected the eight CMIP5 models which had both surface humidity data and historical forcing experiments available through 2012 for combined anthropogenic and natural forcings (i.e., “All-Forcing”) and natural forcings only (i.e., “Natural-Forcing”). We use long “Control runs”, with forcings held constant at pre-industrial levels, to simulate internal climate variability. Forced responses of the models were estimated from the ensemble mean across models of the available ensembles of the

forced simulations. Figure S1 (Supplemental Material) indicates the eight CMIP5 models used, the control run lengths, and the number of All-Forcing and Natural-Forcing ensemble members. While we selected the eight models based on the availability of natural forcing runs to 2012, a broad assessment of the performance of the models based on Fig. 9.7 of IPCC AR5 (Flato et al. 2013) suggests that these eight models have a mixture of above- and below-average performance in simulating the observed seasonal cycle climatology of a representative set of large-scale atmospheric variables.

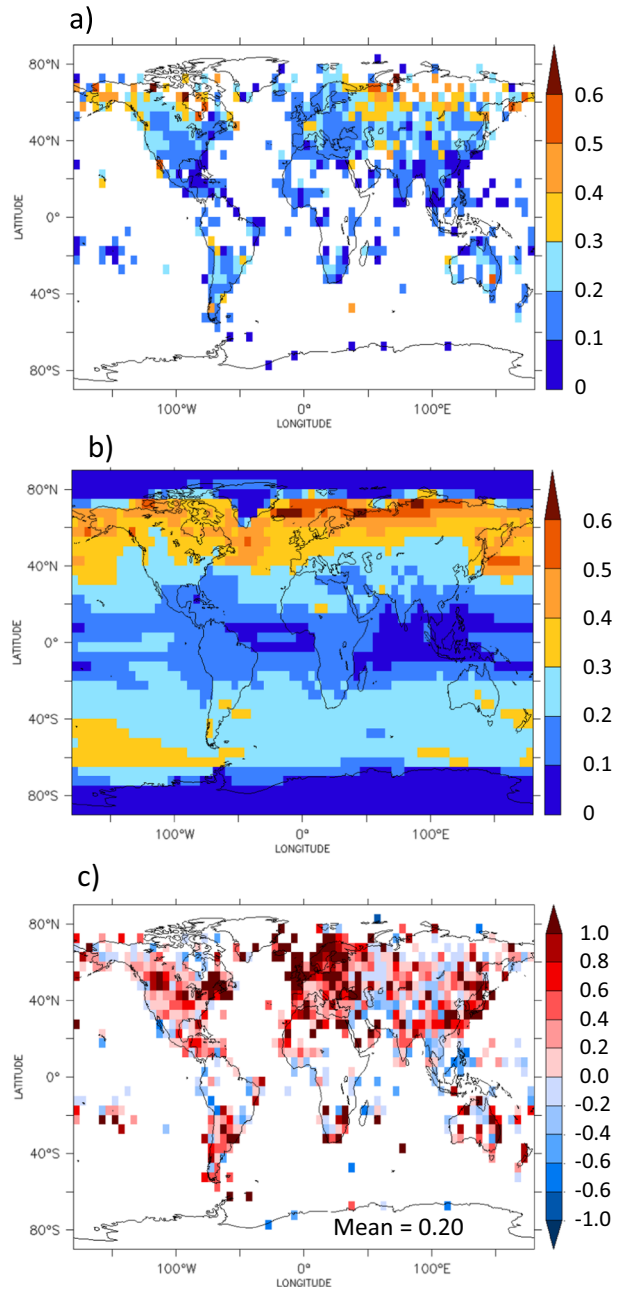
In our study, we attempted to use the same analysis methods for the model and observed data as far as practical, including use of monthly mean temperature, specific humidity and surface pressure to compute monthly mean T_w . For example, we subsampled the monthly mean model data according to the observed (monthly mean) missing data mask (which varies in space and time), after regridding the model data onto the observed grid. Trends were based on standard least-squares regression. Further details on methodology are given in the results section.

3 Simulated variability of summertime WBG T

The interannual variability of modeled and observed WBG T is compared in Fig. 1, as variability is particularly important for trend detection tests. Similar maps of summertime *mean* WBG T are contained in Supplemental Material. In Fig. 1, we compare an estimate of the intrinsic or internal variability from observations with variability from unforced control runs. We first adjust the observations by subtracting an estimate of the forced response (the CMIP5 multimodel All-Forcing ensemble anomalies over the 1973–2012 base period), and then compute the standard deviation of the residuals. The standard deviations of WBG T for each of the eight CMIP5 model control runs (Supplemental Fig. S4) are averaged to create a multimodel mean standard deviation (Fig. 1b). The fractional difference between observed and modeled internal variability, [(model – obs)/obs], shown in Fig. 1c, has a global mean value of +0.20, suggesting a modest overestimation of internal variability by the models on average. All eight individual models have a positive bias in the global mean of this metric (Fig. S4). The comparison maps in Fig. S4 show a mix of over- and underestimated internal variability regionally by the models, but with more regions where the models apparently overestimate internal variability. As a sensitivity test, we recomputed Fig. 1 but without removing any forced response from observations (Supplemental Material, Fig. S5). This test again shows a positive bias in interannual variability for the models, suggesting the positive bias in variability is relatively robust to the method of estimating the forced response in observations. In addition to simulation deficiencies, differences between observed and simulated internal variability could be due to errors in observations or to the limited length (40 years) of observational record for sampling internal climate variability (e.g., Fig. S1).

Estimated biases in simulated climatological means and internal variability (Fig. 1; Supplemental Figs. S1–S4) bear on the reliability of trend assessments in our report. For example, model-simulated internal variability that is too large (small) in a region will tend to make our climate change detection results overly conservative (not conservative enough). Conversely, simulated internal variability that is too large (small) makes it too easy (difficult) for the model All-Forcing runs to be consistent with observations. While there is room for improvement in the models and forcing estimates, the current simulations seem appropriate to use, with these caveats, for our trend assessment.

Fig. 1 Interannual standard deviation of summertime mean WBGT (°C) from: **a** observations (1973–2012) with CMIP5 All-Forcing signal removed; **b** average standard deviation from eight CMIP5 model preindustrial control runs. Months included are June–August (Northern Hemisphere) and December–February (Southern Hemisphere). **c** Fractional difference between observations and models of the standard deviation: $(\text{model} - \text{observed})/\text{observed}$. White regions in (a, b) indicate sparse data



4 Analysis of global trends

Global (land) summertime WBGT and related temperature and humidity anomaly time series are compared between observations and models in Fig. 2. The observed and All-Forcing ensemble anomalies are generally in good agreement throughout the analysis period. Observed WBGT

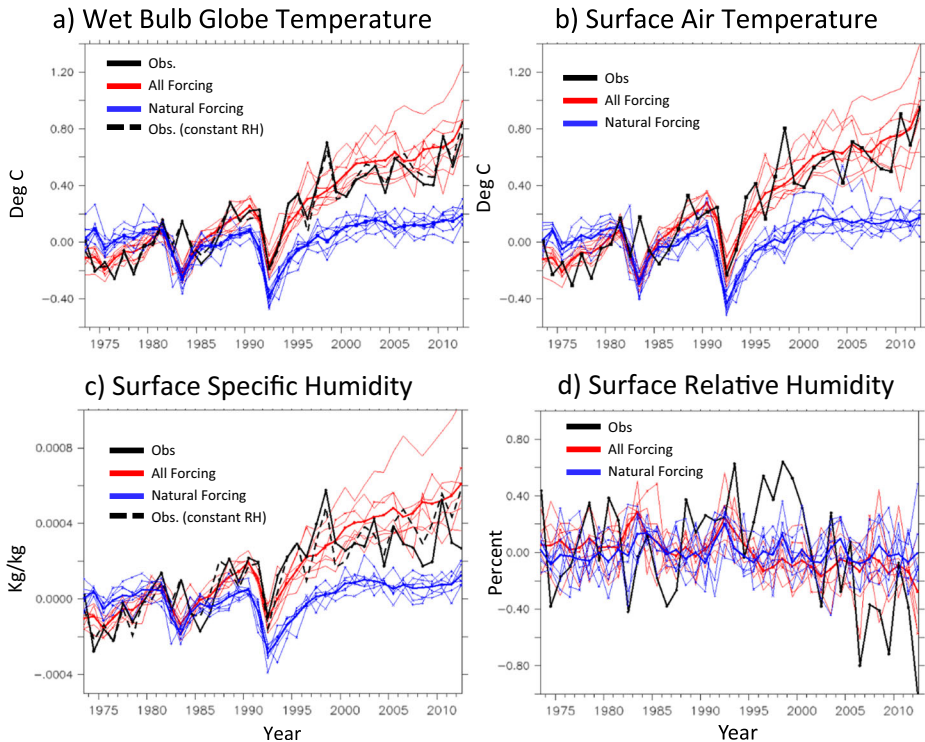


Fig. 2 a Global average summertime surface WBGT anomalies ($^{\circ}\text{C}$), referenced to 1973–1992 means. Black curves: observed anomalies; red lines: CMIP5 All-Forcing runs eight-model ensemble mean (*thick*) and individual model ensemble means (*thin*); *blues lines*: same as red lines except for the Natural Forcing runs. (b–d) as in (a) but for surface air temperature **b**, specific humidity **c**, and relative humidity **d**. Black dashed curves in (a, b): WBGT or surface specific humidity anomalies assuming relative humidity held constant at the summertime climatological 1973–1992 value. For the time series, the December–February mean combines December from the current calendar year with January and February of the subsequent year

(Fig. 2a) has increased since 1973, and the All-Forcing ensemble shows a similar rise not seen in the Natural Forcing runs. Observations indicate a temporary global land summer cooling following the Mt. Pinatubo eruption (1991). Simulated cooling responses to prominent volcanic forcing events are apparent for both the Mt. Pinatubo and El Chichon (1982) eruptions in the All-Forcing and Natural-Forcing series. Observations do not show a pronounced cooling event for El Chichon, perhaps due to the concurrent 1982–83 El Niño event. Of note, for this land-centric, summertime heat-stress metric, there is little evidence in the observations for a global ‘hiatus’ in the observed increase since 2000—a period during which global temperatures, averaged over relatively well-observed regions, show a pronounced slowing of the warming compared to models (Fyfe et al. 2013) and compared to the warming rate over the previous several decades.

Similar results are found for global land summertime surface air temperature (Fig. 2b). Multidecadal changes in the All Forcing ensemble are similar to observed, with global land summer temperatures rising since 2000. Previous analyses have found that observed hot temperature extremes (Seneviratne et al. 2014) and summertime mean temperatures (Ying et al. 2015) over land regions in recent decades do not show the ‘hiatus’ seen in global mean temperature, consistent with our findings. Global mean surface specific humidity during

summer (Fig. 2c, averaged over these same regions) shows more of a hiatus-like behavior, with little increase since 2000, contrasting with the increase in the All-Forcing ensemble over that time. This different behavior for observed specific humidity partly reflects a decrease in average summertime relative humidity over land since the late 1990s (Fig. 2d). This can be seen by comparing the distinct hiatus in observed specific humidity with the more distinct increase in a hypothetical specific humidity over the same period in which we assume a constant relative humidity at the 1973–1992 summer climatological value (Fig. 2c, black solid vs. black dashed). The relative humidity decrease in the observed climate data over land in recent decades is not well-captured in the All-Forcing ensemble (Fig. 2d). This observed decrease has been previously documented (Simmons et al. 2010; Willett et al. 2014a); its cause remains unclear, although one possibility is that the reduction arises from a limited moisture supply over land as ocean warming has not kept pace with land warming over the period. To illustrate the limited effect of the decreasing relative humidity on the WBGT trends since the 1990s, the dashed curve in Fig. 2a shows variations in WBGT assuming relative humidity remained fixed over time. This indicates that to first order the WBGT increases have been driven by the surface temperature increase alone, assuming a fixed relative humidity, with the observed decrease in relative humidity having limited impact. Figure 2 indicates that a typical increase in summer-mean surface air temperature was about 1 °C (1973–2012) whereas a typical decrease in relative humidity was about 0.6 %. A 1 °C warming at constant relative humidity produces about 0.8 °C increase in WBGT, while a relative humidity decrease of 0.6 % at constant temperature decreases WBGT by only about 0.08 °C, confirming that the direct temperature effect is dominant over the relative humidity change influence.

We use a sliding trend analysis (Knutson et al. 2013) to assess causes of observed trends in WBGT and surface air temperature (Fig. 3). The black curve in Fig. 3a depicts observed linear trends in the global summertime surface WBGT series (Fig. 2a) for various starting years, all trends ending in 2012. The color-shaded envelopes in Fig. 3 depict CMIP5 modeled trend distributions (5th to 95th percentiles ranges of the multi-model distributions) for the All-Forcing (red) and Natural Forcing (blue) runs. These illustrate two alternative hypotheses (All-Forcing vs. Natural Forcing) about the climate system—the latter excluding any anthropogenic forcing. Where the black curve (observed trend) lies outside of the blue (Natural Forcing) envelope, we interpret as a detectable trend compared to Natural Forcing only. Where the black curve lies within the red region, we interpret as an observed trend consistent with the All-Forcing ensemble. Cases where the observations are both inconsistent with Natural Forcing and consistent with (or above) the All-Forcing distribution, we interpret as model-based evidence of a detectable anthropogenic influence.

The red-shaded region in Fig. 3a is constructed by pooling the different ensemble mean responses of different models, and samples of internal variability trends of different models, together to create a “grand multi-model distribution” encompassing the All-Forcing models (Knutson et al. 2013). This 5th to 95th percentile range then contains some spread due to internal variability and some spread due to differences in forced responses among the eight models. Alternative forms of a modeled distribution could be used instead, such as a multi-model grand ensemble mean bounded by a spread based on the average of the 5th and 95th percentiles obtained from the individual model control runs. Though not used here, this latter approach provides a distribution of trends from a hypothetical single model having the average characteristics of the eight individual models.

According to our multi-model tests, the increasing trends in global WBGT are detectable for start years ranging from 1973 to about 1987 (all trends ending in 2012). For all of these

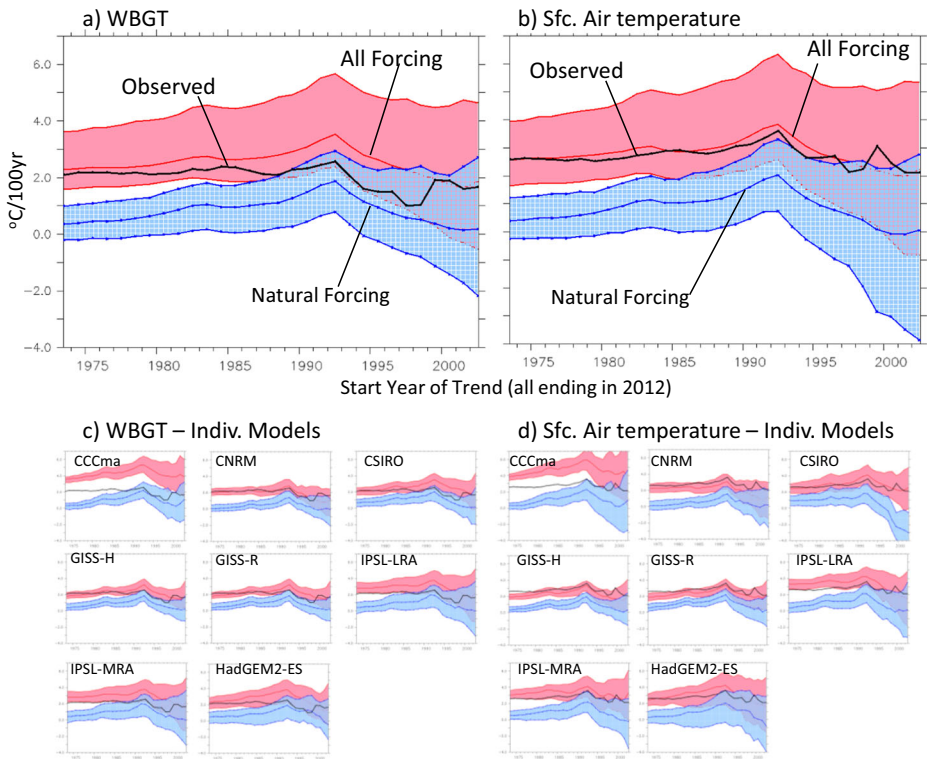


Fig. 3 Trends ($^{\circ}\text{C}/100\text{ yr}$) in global area-averaged summertime surface **a** WBGT and **b** air temperature series from Fig. 2 as a function of starting year, with all trends ending in 2012. The observed trend values (*black curves*) are compared to the 5th–95th percentile ranges of trends from the CMIP5 eight-model ensembles of All-Forcing (*red shading*) and Natural-Forcing (*blue region*) experiments. Overlap of red and blue regions has darker blue shading. Red and blue lines depict 5th, 50th (median) and 95th percentiles of the model distributions. Results for individual models shown in (**c**, **d**)

cases the observed global trends are attributable at least in part to anthropogenic forcing, since they are detectable (i.e., inconsistent with Natural-Forcing), but also consistent with All-Forcing experiments (within the red envelope). There is not strong evidence for a significant inconsistency of global mean land trends for summer WBGT over land in these results: the observed trends are consistent with the All-Forcing trends (though not necessarily detectable) for all start years through at least 2002. The lack of general detectability of the observed trends for start years later than 1987 is not unexpected, as longer records are advantageous for separating trend signals from noise in surface temperature observations (e.g. Knutson et al. 2013), and surface temperature trends beginning in 1988 and later are relatively short for climate change detection purposes. Since the observations also have errors due to station and sampling uncertainties (discussed in detail in Willett et al. 2014b), some of the discrepancy between modeled and observed WBGT trends could also be due to observational errors, discussed further below.

The global WBGT trend assessment results for the eight individual CMIP5 models are shown in Fig. 3c. For the individual model case there is no ambiguity about how to construct modeled trend distribution (ensemble mean of forced run bounded by the 5th to 95th percentile range of that model’s control run). The results are fairly similar to the multi-model ensemble

results in Fig. 3a, with the notable exception of the CCCma model, where the simulated All-Forcing trend shows a significantly stronger warming than observed, and thus the CCCma model is inconsistent with observations (too much warming) for the WBGT 1973–2012 trend.

Figure 3b, d shows the same analysis as (a, c) but for global land summertime mean surface air temperature rather than WBGT. The temperature trends are slightly more consistently detectable across various start years than the WBGT trends, but otherwise the global assessment results are similar. The excessive increasing trend for WBGT in the CCCma assessment is also apparent in the CCCma temperature assessment.

5 Analysis of geographical distribution of trends

Maps of trends (1973–2012) in surface summertime WBGT are compared for observations and the CMIP5 ensemble All-Forcing and Natural-Forcing trends in Fig. 4. Increasing trends are found in nearly in all locations with sufficient data coverage, except for a few outlier gridboxes with negative trends (a). The observed trend map features are well-represented in the All-Forcing ensemble (b). The Natural-Forcing ensemble (c) gives a poor representation of observed trend behavior, as expected from the global analysis (Fig. 3). Figure 4d shows the difference between the All-Forcing and observed trends maps. The observed increasing WBGT trend is stronger than the All-Forcing ensemble mean over Europe and parts of Asia, and less than the All-Forcing trend over parts of eastern Asia, North America, South America, and Australia.

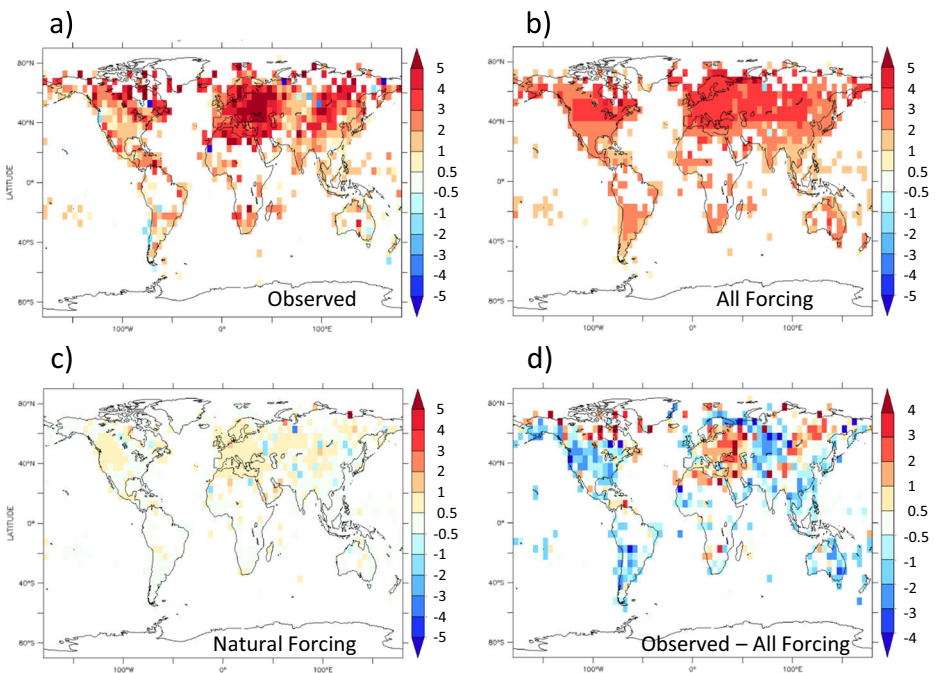


Fig. 4 Geographical distribution of trends (1973–2012) in summertime WBGT (unit: $^{\circ}\text{C}/100\text{ yr}$) for: **a** observations; or CMIP5 All-Forcing **b** or Natural Forcing **c** experiment eight-model ensembles; and **d** observed minus All-Forcing trend (a – b)

The trends assessment discussed for Fig. 3 can be adapted into a map-based regional assessment (Fig. 5). For each gridpoint we compare the observed trend in summer WBGT (1973–2012) with the All-Forcing and Natural-Forcing aggregate distributions of trends from the eight CMIP5 models. For example, for the All-Forcing comparison, an aggregate model distribution of 1973–2012 trends is created by combining the eight individual model trend distributions. Each individual model distribution is created from that model's ensemble mean trend from its All-Forcing ensemble members combined with randomly sampled 40-yr trends from the model's control run. Based on this, we assess whether the observed trend is "detectable" (outside the 5th to 95th percentile range of the Natural Forcing distribution), or "consistent with modeled" (meaning within the All-Forcing 5th to 95th percentiles). Cases of particular interest are where the observed trend is both inconsistent with Natural Forcing and either consistent with, or greater than, the All-Forcing 5th to 95th percentile range: this we interpret as having a detectable anthropogenic increase. According to our model-based analysis and these criteria, there is a detectable increase in summer WBGT over 72 % of the analyzed area, no detectable change over 27 % and a detectable decrease for 2 %. There is detectable anthropogenic increase over 69 % of the global analyzed area. The observed increase in WBGT exceeds the 95th percentile of the All-Forcing distribution (i.e., the models, as a group, underestimate the magnitude of the increase) in 2 % of the analyzed area. For 83 % of the analyzed area, observed trends are consistent with the All-Forcing runs.

A similar analysis for surface air temperature (Fig. 5b) indicates no detectable change over 51 % of the analyzed area, compared to 27 % for WBGT, indicating a greater level of detectability for summertime WBGT than for air temperature alone, at least at the gridpoint scale. The smaller fraction of area with detectable increases for temperature, compared to WBGT, is consistent with the relatively higher variability of summer surface temperature the grid point scale (Supplemental Fig. S6), which makes detection of observed trends, of given magnitude, less likely for temperature than for WBGT. Areas where non-detection occurs for temperature are generally also areas with non-detection for WBGT, as 23 % of the analyzed area did not have a detectable change for both temperature and WBGT, which is close to the percent area with non-detection for temperature alone. Attributable anthropogenic increases were identified over 47 % of the analyzed area for temperature, or somewhat less than for WBGT (69 %). The All-Forcing historical runs show about the same percent of analyzed area consistent with observations for temperature (84 %) as for WBGT (83 %).

We analyzed the 1973–2012 trends of the eight individual CMIP5 models separately, to explore the robustness of our aggregated multimodel ensemble results. The results, summarized in Fig. 5c–f, depict the number of models out of eight at each gridpoint that either have detectable and consistent anthropogenic increases, or are consistent with observations for that gridpoint. For WBGT (c, d) the detectable or consistent finding typically occurs for half or more of the models, though rarely for all eight models. The fewest number of individual models showing a detectable or consistent trend occurs for temperature for the case of detection of an anthropogenic increase (Fig. 5e). Over much of the globe fewer than half of the models support the finding of a detectable anthropogenic increase in surface temperature since 1973 when assessed at the gridpoint scale (Fig. 5e). This is not unexpected given the relatively short record being analyzed, and the expected reduction of signal-to-noise ratios when examining individual gridpoints as opposed to larger regions where effects of internal variability tend to be partially averaged out. Versions of Fig. 5 a, b for individual models are shown in Supplementary Material (Figs. S9, S10); these indicate that regional detection and attribution results for some of the individual models can differ from the multi-model ensemble results.

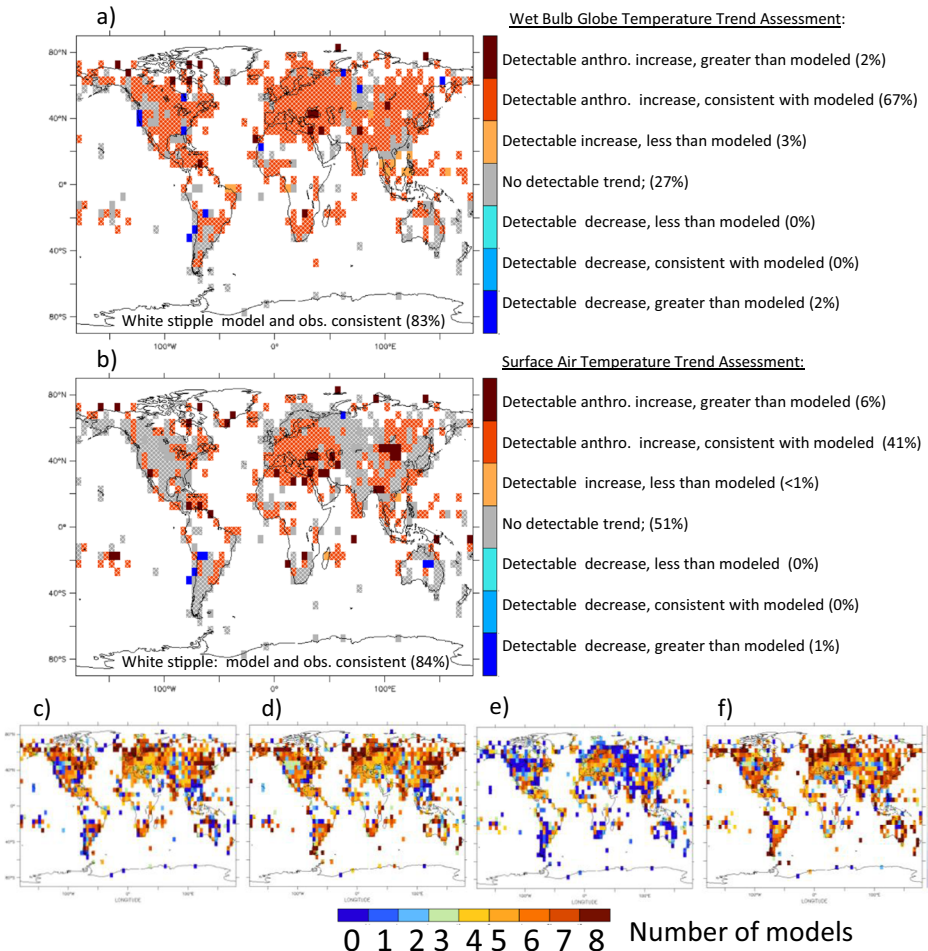


Fig. 5 Assessment of summertime trends (1973–2012) in surface: **a** WBGT and **b** air temperatures. The assessment compares observed trends with eight-model aggregate distributions of trends from CMIP5 All-Forcing and Natural-Forcing experiments (Fig. 4) (see text). The colors in (a, b) indicate different categories of assessment result; the categories are defined in the legends, along with percent of analyzed area for each category. Grid boxes where the observations and multi-model ensemble of trends are consistent (see text) are white stippled. Panels **c** and **d** show the number of CMIP5 models out of eight where: **c**) the WBGT trend is categorized as a detectable and consistent anthropogenic increase, or **d**) model All-Forcing runs are consistent with observations. Panels **e** and **f** are same as **c** and **d** but for surface temperature trends

6 Concluding remarks

Overall, our results suggest that a detectable anthropogenic increase in summertime WBGT has emerged over many land regions since 1973. This increase is consistent with, and has been primarily driven by, the increase in surface air temperature, and has occurred despite a decrease in average relative humidity over land regions in recent decades. Owing to smaller levels of interannual summertime variability for WBGT, compared to air temperature, the observed WBGT increases since 1973 are relatively more detectable than temperature increases, particularly at the gridpoint scale. Our findings suggest that mean levels of heat stress during

summer have been elevated as well, as a consequence of the narrowing, on average, of the differential between summertime environmental WBGT and the human body temperature (which is a critical fixed threshold in the problem).

Our analysis provides a quantitative assessment of this possible anthropogenic influence on summertime heat stress. Our findings cannot be used as absolute conclusive evidence of an anthropogenically driven increase in heat stress, since there are a number of important caveats that apply to this study. For example, we have used a simplified heat stress index and do not consider possible changes in solar insolation, wind, urbanization, or diurnal cycle effects, though some of these neglected effects (e.g., urbanization) are apparently further increasing heat stress levels in some areas. An important remaining question is whether the CMIP5 models provide an adequate estimate of trend possibilities due to natural variability. For example, if internal climate variability were underestimated substantially by the models (i.e., by more than a factor of two for the global mean) or if the response to natural forcings were strongly underestimated, our global mean detection result could be compromised. Further, the attribution of the observed increase in WBGT at least partly to anthropogenic forcing assumes that the modeled response to anthropogenic forcing, and the specifications of the forcing agents, are realistic enough to be adequate for this conclusion. Gaining further confidence in model simulations of internal climate variability, or better constraining models and forcings with observations, are challenging research problems for which innovative future detection/attribution techniques and expanded use and availability of paleoclimate proxy data may be promising approaches. Among further assumptions are that the trends in the observational data represent real climate trends rather than artifacts (e.g. data inhomogeneities). In that regard, the spatial coherence of features in the observed trend maps (Fig. 4a) and the overall similarity of the observed trend pattern to the forced pattern (Fig. 4b) together suggest that the main trend features seen in observations are real climate changes and not data homogeneity artifacts.

Despite these uncertainties, our model-based assessment overall strongly suggests that the observed increase in summertime WBGT over land regions is detectable compared to natural variability, and is partly attributable to anthropogenic forcing. Summertime WBGT has continued to increase in recent years—consistent with climate models—despite the apparent ‘hiatus’ in global mean temperature and a slight reduction in observed relative humidity over land regions. Our results support the plausibility of CMIP5 model projections showing a pronounced continued increase of summertime WBGT during the 21st century, implying increases in heat stress levels, and likely consequences for human health, particularly in relatively warm regions and seasons.

Acknowledgments We thank several GFDL colleagues and three anonymous reviewers for helpful comments; Kate Willett for the HadISDH data set; and PCMDI and CMIP5 modeling groups for model data.

Author contributions TRK designed the study and wrote the paper. JJP performed most calculations for the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

American College of Sports Medicine (1984) Position stand on the prevention of thermal injuries during distance running. *Med J Aust* 141(12–13):876–879

- Cowan T, Purich A, Perkins S, Pezza A, Boschat G, Sadler K (2014) More frequent, longer, and hotter heat waves for Australia in the twenty-first century. *J Clim* 27:5851–5871
- Delworth TL, Mahlman JD, Knutson TR (1999) Changes in heat index associated with CO₂-induced global warming. *Clim Chang* 43(2):369–386
- Dole R et al (2011) Was there a basis for anticipating the 2010 Russian heat wave? *Geophys Res Lett* 38, L06702. doi:10.1029/2010GL046582
- Dunne JP, Stouffer RJ, John J (2013) Reductions in labour capacity from heat stress under climate warming. *Nat Clim Chang*. doi:10.1038/nclimate1827
- Epstein Y, Moran D S (2006) Thermal comfort and the heat stress indices. *Industrial Health* 44:388–398
- Fischer EM, Knutti R (2013) Robust projections of combined humidity and temperature extremes. *Nat Clim Chang* 3:126–130
- Fischer EM, Knutti R (2015) Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nat Clim Chang* 5:560–564
- Fischer EM, Schär C (2010) Consistent geographical patterns of changes in high-impact European heatwaves. *Nat Geosci* 3:398–403
- Flato G, Marotzke J, Abiodun, B, et al (2013) Evaluation of climate models. In: climate change 2013: The physical science basis. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Fyfe JC, Gillett NP, Zwiers FW (2013) Overestimated global warming over the past 20 years. *Nat Clim Chang* 3: 767–769. doi:10.1038/nclimate1972
- Gaffen D, Ross R (1998) Increased summertime heat stress in the US. *Nature* 396:529–530
- Grundstein A, Dowd J (2011) Trends in extreme apparent temperatures over the United States, 1949–2010. *J Appl Meteorol Climatol* 50:1650–1653
- Hartmann DL, Klein Tank AMG, Rusticucci M, et al (2013) Observations: atmosphere and surface. In: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Kalnay E et al (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77:437–471
- Knutson TR, Zeng F, Wittenberg AT (2013) Multimodel assessment of regional surface temperature trends: CMIP3 and CMIP5 twentieth century simulations. *J Clim*. doi:10.1175/JCLI-D-12-00567.1
- Knutson TR, Zeng F, Wittenberg AT (2014) Multimodel assessment of extreme annual-mean warm anomalies during 2013 over regions of Australia and the western tropical pacific [in “explaining extremes of 2013 from a climate perspective”]. *Bull Am Meteorol Soc* 95(9):S26–S30
- Lau N-C, Nath MJ (2012) A model study of heat waves over North America: meteorological aspects and projections for the twenty-first century. *J Clim* 25:4761–4784
- Lewis SC, Karoly DJ (2013) Anthropogenic contributions to Australia’s record summer temperatures of 2013. *Geophys Res Lett* 40:3705–3709. doi:10.1002/grl.50673
- Meehl GA, Tebaldi C (2004) More intense, more frequent and long lasting heat waves in the 21st century. *Science* 305:994–997
- Otto FE, Massey LN, van Oldenborgh GJ, Jones RG, Allen MR (2012) Reconciling two approaches to attribution of the 2010 Russian heat wave. *Geophys Res Lett* 39, L04702. doi:10.1029/2011GL050422
- Pal JS, Eltahir EAB (2015) Future temperature in southwest Asia projected to exceed a threshold for human habitability. *Nat Clim Chang*. doi:10.1038/NCLIMATE2833
- Perkins SE, Lewis SC, King AD, Alexander LV (2014) Increased simulated risk of the hot Australian summer of 2012/2013 due to anthropogenic activity as measure by heat wave frequency and intensity [in “explaining extremes of 2013 from a climate perspective”]. *Bull Am Meteorol Soc* 95(9):S34–S37
- Rahmstorf S, Coumou D (2011) Increase of extreme events in a warming world. *Proc Natl Acad Sci* 108(44): 17905–17909
- Schär C (2016) The worst heat waves to come. *Nat Clim Chang* 6:128–129
- Senevratne SI, Donat MG, Mueller B, Alexander LV (2014) No pause in the increase of hot temperature extremes. *Nat Clim Chang* 4:161–163
- Sherwood SC, Huber M (2010) An adaptability limit to climate change due to heat stress. *Proc Natl Acad Sci U S A* 107:9552–9555. doi:10.1073/pnas.0913352107
- Simmons AJ, Willett KM, Jones PD, Thorne PW, Dee DP (2010) Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets. *J Geophys Res* 115, D01110. doi:10.1029/2009JD012442
- Steadman N (1984) A universal scale of apparent temperature. *J Clim Appl Meteorol* 23:1674–1687

- Stott PA, Stone DA, Allen MR (2004) Human contribution to the European heatwave of 2003. *Nature* 432:610–614. doi:[10.1038/nature03089](https://doi.org/10.1038/nature03089)
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experimental design. *Bull Am Meteorol Soc* 93:485–498
- Willett KM, Sherwood S (2012) Exceedance of heat index thresholds for 15 regions under a warming climate using the wet-bulb globe temperature. *Int J Climatol* 32:161–177. doi:[10.1002/joc.2257](https://doi.org/10.1002/joc.2257)
- Willett KM, Gillett NP, Jones PD, Thorne PW (2007) Attribution of observed surface humidity changes to human influence. *Nature* 449:710–713. doi:[10.1038/nature06207](https://doi.org/10.1038/nature06207)
- Willett KM, Dolman AJ, Hurst DF, Rennie J, Thorne PW (2014a) Global climate [in “state of the climate in 2013”]. *Bull Am Meteorol Soc* 95(7):S5–S49
- Willett KM, Dunn RJH, Thorne PW, Bell S, de Podesta M, Parker DE, Jones PD, Williams CN Jr (2014b) HadISDH land surface multi-variable humidity and temperature record for climate monitoring. *Clim Past* 10: 1983–2006. doi:[10.5194/cp-10-1983-2014](https://doi.org/10.5194/cp-10-1983-2014)
- Yaglou CP, Minard D (1957) Control of heat casualties at military training centers. *Am Med Ass Arch Ind Health* 16:302–316
- Ying L, Shen Z, Piao S (2015) The recent hiatus in global warming of the land surface: scale-dependent breakpoint occurrences in space and time. *Geophys Res Lett* 42:6471–6478. doi:[10.1002/2015GL064884](https://doi.org/10.1002/2015GL064884)