• Original Paper •

HadISDH.extremes Part II: Exploring Humid Heat Extremes Using Wet Bulb Temperature Indices※

Kate M. WILLETT*

Met Office Hadley Centre, *FitzRoy Road*, *Exeter*, EX1 3*PB*, *U.K.*

ABSTRACT

Heat events may be humid or dry. While several indices incorporate humidity, such combined indices obscure identification and exploration of heat events by their different humidity characteristics. The new HadISDH.extremes global gridded monitoring product uniquely provides a range of wet and dry bulb temperature extremes indices. Analysis of this new data product demonstrates its value as a tool for quantifying exposure to humid verses dry heat events. It also enables exploration into "stealth heat events", where humidity is high, perhaps enough to affect productivity and health, while temperature remains moderate. Such events may not typically be identified as "heat events" by temperature-focused heat indices. Over 1973–2022, the peak magnitude of humid extremes (maximum daily wet bulb temperature over a month; T_wX) for the global annual mean increased significantly at $0.13 \pm 0.04°C$ $(10 \text{ yr})^{-1}$, which is slightly slower than the global annual mean T_w increase of $0.22 \pm 0.04°C$ $(10 \text{ yr})^{-1}$. The frequency of moderate humid extreme events per year (90th percentile daily maxima wet bulb temperature exceedance; T_wX90p) also increased significantly at 4.61 ± 1.07 d yr^{-1} $(10 \text{ yr})^{-1}$. These rates were slower than for temperature extremes, TX and TX90p, which respectively increased significantly at $0.27 \pm 0.04°C$ $(10 \text{ yr})^{-1}$ and 5.53 ± 0.72 d yr^{-1} $(10 \text{ yr})^{-1}$. Similarly, for the UK/Europe focus region, JJA-mean T_wX increased significantly, again at a slower rate than for TX and mean T_w. HadISDH.extremes shows some evidence of "stealth heat events" occurring where humidity is high but temperature remains more moderate.

Key words: wet bulb temperature, heat extremes, climate monitoring

Article Highlights:

- HadISDH.extremes is a new monitoring product for exploring large-scale, long-term features of humid and dry heat events.
- HadISDH.extremes shows significantly increasing humid heat extreme magnitude and frequency, but at a slower rate than for dry heat.
- HadISDH.extremes shows that increasing extreme heat magnitude is changing more slowly than mean temperature and humidity.

---

1. Introduction

Humans constantly regulate their body temperature to survive and optimise performance. We are very good at it, both with man-made interventions of clothing and heating/cooling systems, and naturally, through our ability to change our pore size, shiver and sweat. However, our natural cooling ability depends on the temperature, humidity, wind speed and radiation of our surrounding environment. Critically, when we need to reduce the temperature of our bodies we must be able to lose heat faster than we gain it. Heat gain can occur through several mechanisms: direct or indirect radiation from a heat source such as the sun; conduction from surrounding air or water; or internally from raising our metabolic rate by undertaking physical exertion.

There are various heat indices that combine these things to quantify the heat stress loading on a human. The wet bulb globe temperature (WBGT), humidex, environmental stress index, and apparent temperature and universal thermal comfort index (UTCI) are the most commonly used. These are described variously in Willett and Sherwood (2012), Zare et al. (2018), Schwingshackl et al. (2021) and Brimicombe et al. (2022). These indices typically require mea-

surements of temperature, humidity, radiation and wind speed. In some cases, approximations of the latter two are used, often assuming light winds and shaded conditions. In the case of the UTCI (Jendritzky et al., 2012), the calculation incorporates a physiological model of the human body, approximating the heat loading on an average person walking at 4 km h^{-1}. These added layers of complexity are very useful. However, they make it difficult to create long-term monitoring datasets of these quantities because the required observations are not available on a global scale for a sufficiently long (e.g., 30 year) time period. Combined heat indices also make it harder to distinguish the driving atmospheric and terrestrial processes that may have driven extremes in those quantities, because these differ depending on whether the heat event is dry or humid (Raymond et al., 2017, 2021; Zhang et al., 2021). Furthermore, the impact of heat can differ depending on whether the heat is humid or dry (Yu et al., 2021); yet for WBGT, the same value can be reached by either very hot and dry conditions or moderately hot and humid conditions (Budd, 2008). The behavioural decisions and mechanisms put in place to avoid ill-effects of excess heat may differ (Ahmed et al., 2014). In dry heat, staying in the shade, reducing activity levels, keeping hydrated, increasing air flow, and use of other passive cooling techniques such as evaporative cooling, night ventilation, ground cooling, and radiant cooling may be sufficient to keep cool. In humid heat, increased hydration, air flow, and passive cooling techniques may not be as effective. Activity levels must be far more limited and mechanical air cooling may be required.

The wet bulb temperature (T_w) has become a measure of interest relating to heat stress severity (Sherwood and Huber, 2010; Schär, 2016) because it is in some ways analogous to human body temperature regulation by means of evaporation of sweat from skin. Explicitly, the T_w is the temperature of the air cooled by evaporation, given the ambient temperature and amount of moisture in the air. If the air is below saturation [< 100 %rh relative humidity (RH)], then evaporation into it should occur. The process of evaporation removes energy from the water source. Hence, the less saturated the air is, the more evaporation that can occur, the lower the T_w will be. By covering a thermometer in a material wick that is continuously wetted through contact with a reservoir of water, evaporation can readily occur up to the point where the air becomes saturated – at which point the T_w is then equal to the dry bulb temperature T. The difference between T_w and T represents the degree to which evaporative cooling can be effective.

Taking this back to the human body analogy, our skin is the gateway for dissipating heat from the body. Although our skin temperature can vary quite a lot, on average it is between 33°C and 37°C. Generally, T_w is below this level, so water (sweat) can be readily evaporated away from the skin, leaving the skin cooler as a result, just like the wet bulb thermometer. However, should the air have a T_w of a similar level to the skin temperature, this means that very little

or no evaporative cooling of the skin can take place. Once the body can no longer dissipate heat effectively, the internal body temperature has to rise, which can prove fatal. A theoretical T_w of 35°C has been established as the critical level (Sherwood and Huber, 2010; Raymond et al., 2020). In practice, the T_w threshold above which sweat can no longer offload heat faster than heat gain can be much lower even in young, healthy adults undertaking moderate activity (Vecellio et al., 2022).

Clearly, T_w is a very useful measure of how inhabitable a place is in terms of the ability of the human body to survive. The ability to thrive and be productive can also be related to T_w because as T_w gets closer to skin temperature (~35°C) it becomes harder and more dangerous for the body to be physically active either for work or leisure purposes. Levels of activity, rest periods and water intake must be scaled appropriately to keep the body healthy. Animals too can be significantly impacted under high T_w conditions.

A new dataset, HadISDH.extremes.1.0.0.2022f (Willett, 2023a), provides global gridded monthly fields of wet and dry bulb temperature indices. These align with indices under the remit of the Expert Team on Sector-Specific Climate Indices (ET-SCI; ET-SCI, 2022), Climpact project (Climpact, 2022), and HadEX3 gridded extremes indices climate monitoring product (Dunn et al., 2020). Previous studies have presented T_w threshold exceedance values (Wang et al., 2019; Freychet et al., 2020, Raymond et al., 2020; Yu et al., 2021) but none have yet provided this as a globally consistent, updating monitoring product, alongside simultaneous indices for T. Hence, HadISDH.extremes provides a unique resource to study different types of heat events that may be high humidity–moderate temperature, low humidity–high temperature, or compound high humidity–high temperature. In terms of heat-related mortality, temperature is usually the dominant factor, correlating as well or better with mortality than indices that include humidity (Armstrong et al., 2019). However, high humidity can still lead to reduced productivity and non-fatal health impacts that are more difficult to detect and quantify.

Following the initial dataset paper (Willett, 2023a), here, several of the key indices from HadISDH.extremes are presented at a range of spatial and temporal scales. Two focus regions are used as case studies: China/East Asia and UK/Europe. These are used to explore the value of this product as a tool for monitoring heat extremes across the globe. Comparisons between the wet and dry bulb temperature indices are made to establish the different characteristics of these two contributing factors to heat stress and whether 'stealth heat events' are detectable. These are when high humidity heat may be occurring while temperature remains more moderate. Such events would not be detected using temperature-focused indices.

2. Methods

HadISDH.extremes (Willett, 2023a) is a gridded (5° by

5°) global monthly product from January 1973 to December 2022 (at time of writing). It is built within the framework of the existing Met Office Hadley Centre led International Surface Dataset for Humidity, HadISDH.land (Willett et al., 2013, 2014), which is a long-term, quality controlled, homogenised, gridded monthly mean land surface humidity monitoring product. A wide range of monthly extremes indices are available based on the maximum and minimum daily maxima and minima T_w over the month (T_wX and T_wN) and the equivalent for T (TX and TN). The indices used here are summarised in Table 1.

As HadISDH.extremes is designed as a climate monitoring product, data quality and long-term stability are a key focus. However, this is always a balance between data coverage and data quality. Persistent saturation is a specific issue for T_w. If the wet bulb thermometer (if this is the primary instrument used for measurement) partially or totally dries out, T_w will begin to track T, leading to spurious high (possibly extreme) T_w values, especially in hot climates. The quality control (QC) includes humidity-specific tests for this, and neighbour checks to avoid removal of real extremes, although of course no QC is perfect.

The data originate as sub-daily observations that could be as coarse as six-hourly. Even with hourly data the true maxima and minima are likely under- and overestimated, respectively. To minimise the error from this, strict data completeness requirements are in place:

• ≥ 4 observations per day with at least one in each 8-h tercile (0000–0700, 0800–1500, 1600–2300);
• ≤ 6 days missing per month;
• ≤ 20 days missing per year;
• ≤ 3 months missing per year;
• ≥ 15 years for each calendar month within the climatological period (1991–2020), with at least one in each decade (1991–2000, 2001–2010, 2011–2020).

This results in 4460 contributing stations. Data gaps remain over the high latitudes, Africa and large parts of Central and South America. Outside of North America, Europe, eastern China and Japan, most gridboxes contain only one to two stations [see Fig. 1 in (Willett 2023)]. Only gridboxes with at least 70% data completeness over the period 1973–2021 are used when calculating global and regional means to minimise temporal sampling bias.

Homogenisation to remove non-climate artefacts from

Table 1. Descriptions of HadISDH.extremes indices used within this study.

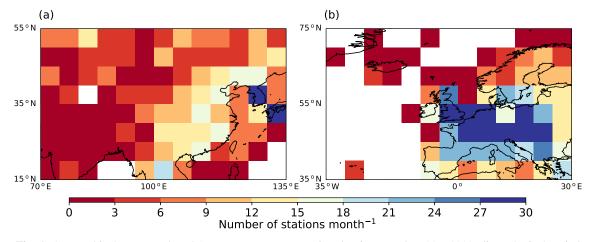| Index name	Index long name	Index description (NB: in all cases, climatological refers to the period 1991–2020)
T_wX	Maximum wet bulb temperature	Gridbox **mean** of station month maxima of daily maximum T_w
T_wX90p	90th percentile maximum wet bulb temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T_w exceeds the climatological 90th percentile of daily maxima
T_wX25	25°C maximum wet bulb temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T_w is equal to or exceeds 25°C
T_wX27	27°C maximum wet bulb temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T_w is equal to or exceeds 27°C
T_wX29	29°C maximum wet bulb temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T_w is equal to or exceeds 29°C
T_wX31	31°C maximum wet bulb temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T_w is equal to or exceeds 31°C
T_wX33	33°C maximum wet bulb temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T_w is equal to or exceeds 33°C
T_wX35	35°C maximum wet bulb temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T_w is equal to or exceeds 35°C
TX	Maximum temperature	Gridbox **mean** of station month maxima of daily maximum T
TX90p	90th percentile maximum temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T exceeds the climatological 90th percentile of daily maxima
TX25	25°C maximum temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T is equal to or exceeds 25°C
TX30	30°C maximum temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T is equal to or exceeds 30°C
TX35	35°C maximum temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T is equal to or exceeds 35°C
TX40	40°C maximum temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T is equal to or exceeds 40°C
TX45	45°C maximum temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T is equal to or exceeds 45°C
TX50	50°C maximum temperature exceedance	Gridbox **mean** of station percentages of days per month where the daily maximum T is equal to or exceeds 50°C

Fig. 1. Geographical extent and HadISDH.extremes mean station density over the 1991–2020 climatological period per 5° by 5° gridbox for the two case study regions of (a) China/East Asia (15°–55°N, 70°–135°E) and (b) UK/Europe (35°–70°N, 35°W–30°E).

the station time series is fundamental to long-term stability. However, this is very difficult to do, and validate, for large datasets with no digitised metadata available, especially at the daily or sub-daily level. It can lead to large uncertainty. HadISDH.extremes instead uses a homogeneity assessment approach. It uses the homogenisation adjustment information from the monthly mean HadISDH.land T and T_w stations to allocate a score (HQ Flag) for each gridbox month depending on the number of stations present, whether or not they contain an inhomogeneity and the size of that inhomogeneity. To remove the data with the largest inhomogeneities, while retaining reasonably consistent coverage over space and time, only those gridboxes with an HQ Flag = 0–6 are used here, as recommended in Willett (2023a). As with the QC, no method is perfect. It is also possible that inhomogeneities in the station data, which are not removed by the HQ Flag screening, might lead to periods where stations give systematically higher or lower readings than they should.

Furthermore, in addition to observation error, measurement error, and inhomogeneity, the conversion from the originally observed humidity variable to reported dewpoint temperature, and then to T_w for HadISDH, will contain some degree of error that is very difficult to quantify. Most humidity calculations are close approximations to some degree and further complicated by choice of precision of the input values. The gridbox monthly extremes indices are more vulnerable to these errors than the HadISDH.land gridbox monthly means, even if using the provided anomalies from the 1991–2020 climatological period. This is because they are based on a single hourly observation from the month rather than a mean where random errors at least may be reduced by averaging.

Anomalies are used where possible here to minimise error. This also minimises the impact of heterogeneity in altitude, aspect and land cover across the 5° gridboxes. Actual values, particularly extremes, could differ considerably between nearby stations. These differences affect the anomalies to a lesser extent because there is a larger correlation decay distance in the anomalies compared to the actual val-

ues. As Fig. 1 demonstrates, a gridbox mean could be from one single station or 30+, and this number can change over time. Clearly, caution is needed when interpreting the results of HadISDH.extremes at very small regional scales. Where HadISDH.extremes is used for specific feature analysis, cross-checking with other evidence such as national datasets or records of extreme events should be undertaken to verify the features. HadISDH.extremes is useful for identifying extreme humidity events in a consistent way across the globe, but given the vulnerability of extremes to error, users should always explore the nature of the data prior to making conclusions to ensure the validity and fit of conclusions to the data.

Two case study regions are chosen that have good data coverage and station density but represent different climatological zones. These are shown in Fig. 1. China/East Asia covers a wide range of climates, encompassing gridboxes where high humidity and temperature is common. UK/Europe is a generally climatologically drier region in terms of humidity. The China humidity observations have been found to contain a region-wide inhomogeneity over the early 2000s when manual wet bulb thermometers were changed to automated RH sensors (Freychet et al., 2020; Li et al., 2020). After homogenisation, Freychet et al. (2020) found that trends in T_w increased. This inhomogeneity is not detectable in the HQ Flag scores provided with HadISDH.extremes (Willett, 2023a). It is possible that the HadISDH.land homogenisation process is less able to detect it because the changes occur at a similar time across the whole region. This means that HadISDH.extremes T_w extremes indices presented over China may well be underestimates.

3. Analysis and validation of HadISDH.extremes

3.1. *Changes in peak extreme magnitude and moderate extreme frequencies*
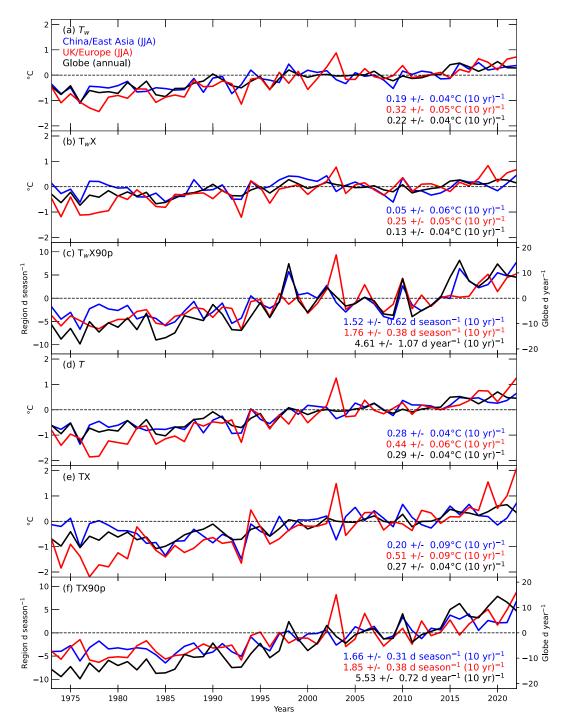
Figure 2 shows an overview of HadISDH.extremes

Fig. 2. Global mean annual and regional mean JJA anomaly time series and decadal trends from HadISDH. extremes (T_wX, T_wX90p, TX and TX90p) screened to remove gridboxes where the HQ Flag > 6, and the homogenised HadISDH.land (T_w and T). Regional extents are as presented in Fig. 1 but only gridboxes with 70% temporal completeness. Decadal trends in anomalies (relative to 1991–2020) are fitted using ordinary least-squares regression with 90th percentile confidence intervals corrected for AR(1) following Santer et al. (2008).

global (annual) and regional (seasonal – JJA) time series and trends in the magnitude of peak extremes (T_wX and TX) and the frequency of moderate extremes (T_wX90p and TX90p), alongside equivalent monthly mean quantities (T_w and T) from HadISDH.land.4.5.1.2022f. For T_wX90p and TX90p, the annual/seasonal totals are plotted. Global mean

T_wX and T_wX90p have increased significantly at 0.13 ± 0.04°C (10 yr)$^{-1}$ and 4.61 ± 1.07 d yr^{-1} (10 yr)$^{-1}$, respectively, while global mean T_w increased significantly faster than T_wX at 0.22 ± 0.04°C (10 yr)$^{-1}$. Global TX and TX90p increased significantly at 0.27 ± 0.04°C (10 yr)$^{-1}$ and 5.53 ± 0.72 d yr^{-1} (10 yr)$^{-1}$, respectively, alongside a global mean

significant T increase of 0.29 ± 0.04°C $(10 \text{ yr})^{-1}$. Collectively, Fig. 2 shows that humid heat has increased significantly in both magnitude and frequency but that this rate is slower than the rate of increase in means, and slower than the rate of increase in high temperature heat.

For the UK/Europe, the peak magnitudes of extremes are increasing significantly at a rate of 0.51 ± 0.09°C $(10 \text{ yr})^{-1}$ for TX compared to only 0.25 ± 0.05°C $(10 \text{ yr})^{-1}$ for T_wX. Over China/East Asia, the peak magnitude of extremes for TX is increasing significantly at 0.20 ± 0.09°C $(10 \text{ yr})^{-1}$, while there is in fact no significant increase in T_wX $[0.05 \pm 0.06$°C $(10 \text{ yr})^{-1}$, p-value $> 0.01]$. Trends in the peak magnitude of extremes are increasing faster over the UK/Europe compared to China/East Asia.

In terms of frequency of moderate heat events, although for both regions the TX90p exceedance increases slightly faster than the T_wX90p exceedance, the 90th percentile confidence intervals overlap considerably, meaning that these differences are not statistically significant. For the UK/Europe, T_wX90p increases significantly at 1.76 ± 0.38 d season^{-1} $(10 \text{ yr})^{-1}$ compared to 1.85 ± 0.38 d season^{-1} $(10 \text{ yr})^{-1}$ for TX90p. Over China/East Asia, the rate is slightly smaller, with T_wX90p increasing significantly at 1.52 ± 0.62 d season^{-1} $(10 \text{ yr})^{-1}$ compared to 1.66 ± 0.31 d season^{-1} $(10 \text{ yr})^{-1}$ for TX90p. Similarly, this difference between regions is not statistically significant.

For the UK/Europe region, humidity is a growing factor both in terms of peak magnitude of extremes and frequency of moderate events. Over China/East Asia, while the frequency of moderate humid heat events is increasing significantly, the peak magnitude of extremes appears not to be. Other parts of the distribution may change at different rates. Indeed, HadISDH.land regional mean JJA T_w increases faster than T_wX for both regions, at 0.32 ± 0.05°C $(10 \text{ yr})^{-1}$ for the UK/Europe and 0.19 ± 0.04°C $(10 \text{ yr})^{-1}$ for China/East Asia. For comparison, regional mean JJA T is increasing at a faster rate than T_w in both regions, at 0.44 ± 0.06°C $(10 \text{ yr})^{-1}$ for the UK/Europe and 0.28 ± 0.06°C $(10 \text{ yr})^{-1}$ for China/East Asia. This shows faster rates of increase in both mean T and T_w as well as extremes (TX and T_wX) over the UK/Europe compared to China/East Asia. Clearly, increasing temperature rather than humidity is the main driver of increasing heat stress over the UK/Europe and China/East Asia regions. This agrees broadly with the findings of Freychet et al. (2020) for China.

It is important to note the caveat that the trends over China/East Asia may be underestimated. Although trends in T_wX90p, TX90p, and TX from HadISDH.extremes, and in T_w and T from HadISDH.land, are comparable to those for the globe and UK/Europe, the HadISDH.extremes trend in T_wX is small and not significant. Freychet et al. (2020) demonstrated that inhomogeneity in the Chinese data leads to underestimated trends in T_w, and consequently, T_wX and any related quantities. This is discussed in relation to HadISDH.extremes in Willett (2023a). The fact that HadISDH.extremes shows no significant trend in T_wX, yet a comparable (with other regions) significant positive trend in

T_wX90p, suggests that the peak extremes (T_wX) are more sensitive to this inhomogeneity than the mean (T_w) and moderate extremes (T_wX90p). Overall, this region should be treated cautiously and T_w extremes from HadISDH.extremes seen as potential underestimates. The general agreement between regions and the globe shows that HadISDH.extremes enables a robust conclusion that high humidity heat stress is increasing in magnitude and frequency, albeit to a lesser degree than high temperature heat stress.

3.2. Locations of the most frequent threshold exceedance

Changes in peak magnitude of extremes and frequency of moderate heat events are useful for monitoring climate change. However, they may not necessarily correlate with societal impacts that are dependent on specific thresholds. As for T, there is a level of acclimatisation to T_w such that the level of impact associated with a threshold may differ from region to region (Shen and Zhu, 2015). The six T_wX thresholds from 25°C to 35°C are chosen to cover a wide range of societally relevant thresholds. Figure 3 shows gridbox percentages of days where T_wX is equal to or exceeds these thresholds over the period of record (1973–2022). These are based on the gridbox mean counts and so mask the fact that some stations within the gridbox exceeded the count while others experienced fewer days.

Various sources of error and uncertainty are discussed in section 2, especially in relation to using actual values of the indices rather than the anomalies, and when using indices tied to specific values such as T_wX35. Hence, treating these exceedance thresholds more broadly where possible is recommended. For example, aggregating the six thresholds to "moderate" (T_wX25 and T_wX27), "high" (T_wX29 and T_wX31) and "severe" (T_wX33 and T_wX35) extremes increases the number of data points and reduces the risk of features being driven by a few outliers.

Nearly everywhere outside of the high-latitude and high-elevation regions has experienced days of "moderate" extremes of T_wX25 and T_wX27. Over some parts of the tropics, particularly Southeast Asia, T_wX25 has been almost a daily occurrence. The gridbox with the highest percentage of days, at 96%, lies in the tropical Pacific centred at (7.5°N, 152.5°E). Chuuk International Airport, Micronesia, is the only contributing station. It has a T_wX consistently above 25°C, which appears to increase steadily over time. A data gap in the early 2010s is followed by a period of higher T_wX. For T_wX27, the gridbox with the highest percentage of days lies in the extratropics at (22.5°N, 87.5°E), where three Indian stations contribute. These stations have strong seasonal cycles in T_wX, exceeding 27°C during every summer season.

The "high" extremes thresholds of T_wX29 and T_wX31 have a narrower latitudinal range and much lower percentages, as expected. The gridboxes with the highest counts for each threshold are extratropical, over and around the Middle East. For T_wX29, the gridbox with the highest count is centred at (27.5°N, 57.5°E), which contains three stations from the
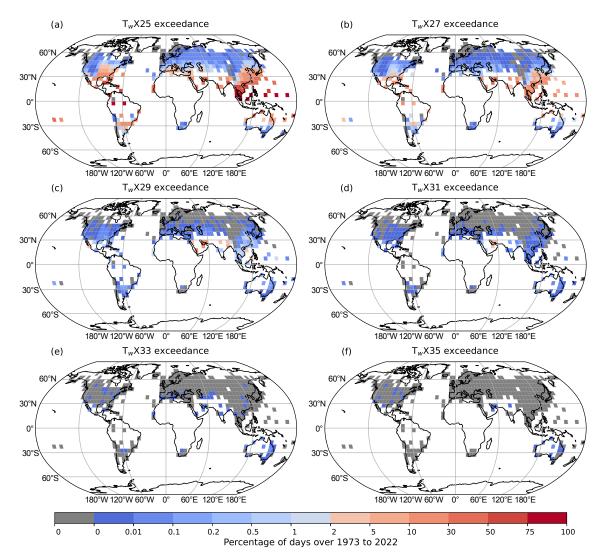
Fig. 3. Percentage of days exceeding various T_wX thresholds from January 1973 to December 2022 from HadISDH. extremes. All data are screened to remove gridboxes where the HQ Flag > 6 and where temporal completeness falls below 70%.

United Arab Emirates and one from Iran. For T_wX31, the gridbox with the highest count is nearby, centred at (27.5°N, 52.5°E), where Bahrain International Airport, Bahrain, and Doha International Airport, Qatar, contribute. From both visual inspection of the HadISDH.extremes stations and the regional consistency, it can be concluded that this extratropical Middle Eastern region experiences the most frequent 'high' heat extremes.

The 'severe' extremes thresholds of T_wX33 and T_wX35 are very limited in spatial extent and day counts, although day counts are apparent over several different continents. The gridboxes containing the highest day counts for both thresholds are over northern Australia. For T_wX33 and T_wX35, the gridbox with the highest count is centred at (17.5°S, 122.5°E), where three stations contribute: Derby Airport, Broome International, and Curtin Airport. Raymond et al. (2020) found robust evidence of T_wX exceeding 35°C over the Persian Gulf and Pakistan between 1979 and 2017, of which the former is also identified by HadISDH.extremes.

The additional locations noted here may be because Raymond et al. (2020) used the earlier version (2.0.1.2017f) of the HadISD dataset, which had fewer stations and was four years shorter. However, it has not been possible to verify the Australian exceedances with other sources. The exceedances over Australia are a good example of why users must seek verification for analysis on small-scale isolated features within HadISDH.extremes. Comparison with national Australian Bureau of Meteorology data holdings for Derby Airport and Curtin Airport [not shown (pers. comm., Blair Trewin)] found mixed results. There were several occurrences of T_wX above 35°C in both the HadISDH. extremes and national data holdings but these did not always match up in time and appeared to be far (~7°C) above "normal" peak values. In each case, the exceedances appeared suspect. Most of these instances occurred before wet bulb thermometers were replaced by RH probes as the primary instruments at the sites concerned, and are likely the result of wet bulb thermometers drying out. Conclusively,

these Australian exceedances are highly uncertain. Regardless, at larger regional scales the conclusions of widespread increasing extremes of T_w are robust.

For comparison, six set thresholds for TX have also been selected, from 25°C to 50°C (Fig. 4). For "moderate" extreme thresholds of TX25 and TX30, the gridboxes with the highest day counts occurred in the tropics over the western tropical Pacific Islands, South East Asia, and eastern Brazil. For the most part, these are close to the locations of peak frequency for "moderate" T_wX thresholds. For the "high" extreme thresholds, the peaks occurred over the Middle East, close to the region of peak "high" extreme threshold exceedance for T_wX. For the "severe" extremes, both threshold peaks are over Kuwait and Saudi Arabia in the Middle East.

3.3. *Changes in the frequency of threshold exceedance for the case study regions*

Focusing again on the China/East Asia and UK/Europe regions, Fig. 5 shows monthly time series of T_wX threshold exceedances in days. A regional mean of gridbox day counts is difficult to interpret directly in terms of societal impacts but useful for studying exposure relative to other regions and change in exposure over time. For both regions, every year on record had T_wX exceeding thresholds of 25°C, 27°C and 29°C. Trends are increasing significantly for these thresholds, although only over China/East Asia for T_wX29. For China/East Asia T_wX25 and T_wX27, "moderate" extreme threshold trends are quite large at 3.36 ± 0.81 d yr^{-1} (10 yr)$^{-1}$ and 2.84 ± 0.68 d yr^{-1} (10 yr)$^{-1}$, respectively. T_wX29 "high" extreme threshold trends are small but positive at 0.46 ± 0.12 d yr^{-1} (10 yr)$^{-1}$. For UK/Europe, exposure is far lower and T_wX25 and T_wX27 "moderate" extreme threshold trends are much smaller at 0.59 ± 0.20 d yr^{-1} (10 yr)$^{-1}$ and 0.14 ± 0.05 d yr^{-1} (10 yr)$^{-1}$, respectively. Both regions show regular exceedances of the T_wX31 "high" extreme threshold, with exceedances occurring annually for China/East Asia since 1994. T_wX31 trends are significant, but very small, for the China/East Asia region, at 0.01 ± 0.01 d yr^{-1} (10 yr)$^{-1}$. For the UK only (50°–60°N, 10°W–5°E; not shown), there have been days exceeding T_wX29 but not T_wX31. UK-only trends are not significant for any of the
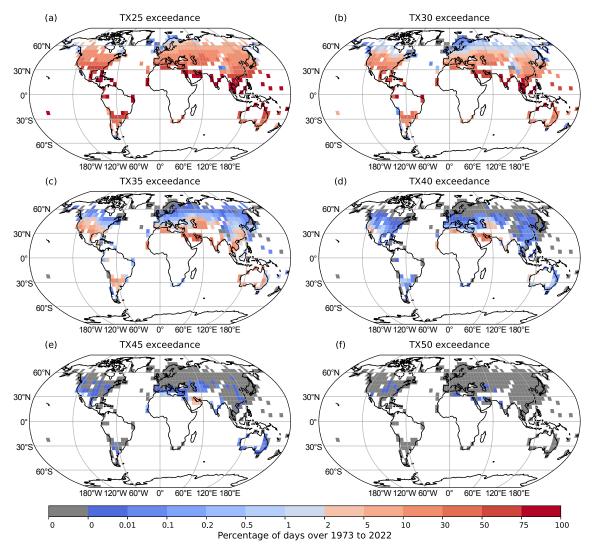

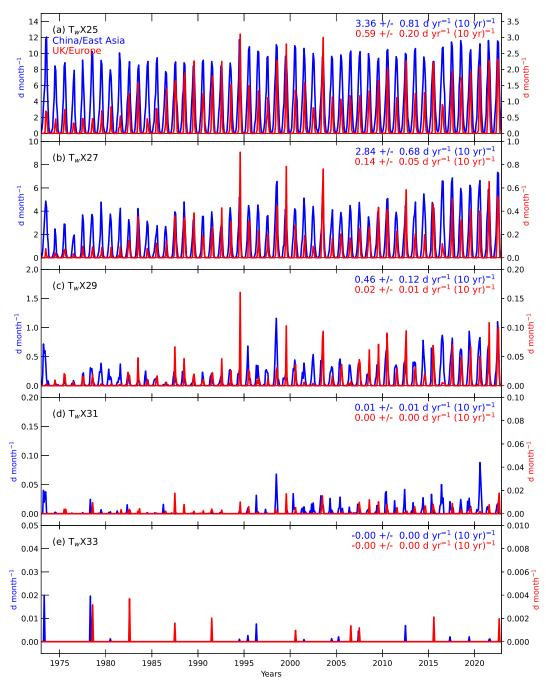Fig. 4. As in Fig. 3 but for TX thresholds.

Fig. 5. Time series of monthly T_wX threshold exceedance for the China/East Asia region (blue) and UK/ Europe region (red). Regional extents are as presented in Fig. 3 but only including gridboxes with at least 70% temporal completeness. The data have been screened to remove any gridboxes where the HQ Flag > 6. Decadal trends in annual total counts are shown for each region, fitted using ordinary least-squares regression with 90th percentile confidence intervals corrected for AR(1) following Santer et al. (2008).

thresholds. For the "severe" extreme thresholds, both wider regions have sporadic exceedances of T_wX33 over time with no discernible trend.

There is a single exceedance of T_wX35 in the China/ East Asia region in 2007 (not shown). This is from a gridbox over northern India (27.5°N, 77.5°E; also identifiable in Fig. 3f) in June. There are three contributing stations for this gridbox: Hissar, Jaipur and Indira Ghandi International Airport. As noted above, these isolated exceedances are uncer-

tain. Visual inspection of the station time series shows a T_wX35 exceedance for Hissar only. This value is around 5°C higher than all others in the time series, suggesting that it is uncertain. This exceedance occurred during the deadly South Asian heat wave in 2007 (ADPC, 2007; Reuters, 2007), but such an extreme value should still be treated with caution. There have been several heat waves over the region since then (Mazdiyasni et al., 2017) but these have not led to such high T_w.

A combined interpretation of the threshold exceedance indices, indices for peak magnitude of extremes, and moderate extreme event frequency, is useful here. Over the China/East Asia region, the lack of any significant trend in the peak magnitude of extremes (T_wX) does not mean that humid heat events are not becoming more of an issue, only that the most extreme extremes do not appear to be increasing. The significant positive trends in T_wX25, T_wX27, T_wX29 and T_wX31 (and also T_wX90p, shown above) show that there are now more days with "moderate" and "high" humidity heat. For the UK/Europe region, although there are significant trends in the peak magnitude of extremes,

there have not yet been corresponding increases in the frequency of "high" humid heat days where T_wX is equal to or greater than 31°C. Clearly, humid heat is far less of a problem for the UK/Europe region compared to the China/East Asia region. Nevertheless, "moderate" extreme threshold days of T_wX25 and T_wX27 could still have negative impacts on productivity over the UK/Europe region where the population is less acclimatised to, and perhaps less well adapted to, humid heat days.

Equivalent time series for the TX thresholds are shown in Fig. 6. For both regions, there are exceedances for TX25 and TX30 every year on record, and indeed all but one
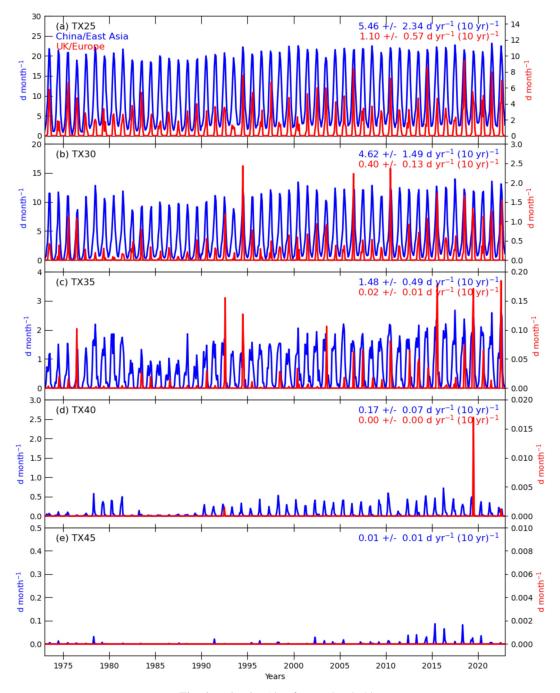

Fig. 6. As in Fig. 5 but for TX thresholds.

month for China/East Asia. China/East Asia also has exceedances every year for TX35, whereas for the UK/Europe exceedances have occurred annually only since 2005. Trends for TX25, TX30 and TX35 are significantly positive for both regions. These trends are 5.46 ± 2.34 d yr^{-1} (10 yr)$^{-1}$, 4.62 ± 1.49 d yr^{-1} (10 yr)$^{-1}$, and 1.48 ± 0.49 d yr^{-1} (10 yr)$^{-1}$, respectively, for the China/East Asia region. They are much smaller, at 1.10 ± 0.57 d yr^{-1} (10 yr)$^{-1}$, 0.40 ± 0.13 d yr^{-1} (10 yr)$^{-1}$, and 0.02 ± 0.01 d yr^{-1} (10 yr)$^{-1}$, respectively, for the UK/Europe region. For China/East Asia, trends are also significantly positive for TX40 and TX45, at 0.17 ± 0.07 d yr^{-1} (10 yr)$^{-1}$, and 0.01 ± 0.01 d yr^{-1} (10 yr)$^{-1}$, respectively. Every year on record has had exceedances for TX40. For TX45, there have been annual exceedances since 2001. Only the China/East Asia region has an exceedance for TX50. This is during May 2016 over western India and southeastern Pakistan (22.5°N, 72.5°E). There are three contributing stations: Ahmedabad, Rajkot and Surat. Only Ahmedabad shows the TX50 exceedance. There was a heat wave reported over the region at that time, where temperatures exceeded 50°C (World Weather Attribution, 2016). However, the reported station was in Rajasthan, further to the northeast than Ahmedabad. Overall, the trends in TX for both regions are larger than for T_wX.

3.4. *Assessing the co-occurrence of seasonal extreme humidity and temperature heat events*

There may be occasions where an extreme heat event would not be detectable from T, yet significant impacts may result from high T_w. To investigate the value of HadISDH. extremes for identifying these "stealth heat events" the co-occurrence, or lack thereof, of high numbers of days (as anomalies from the climatological mean) exceeding the T_wX and TX extreme thresholds over the JJA season for the two regions is explored (Fig. 7). The T_wX and TX thresholds are paired as follows: T_wX25 with TX25; T_wX27 with TX30; T_wX29 with TX35; T_wX31 with TX40; T_wX33 with TX45; and T_wX35 with TX50. There is no strong physical reason for pairing the T_wX and TX thresholds as done here but this covers the spread from generally "moderate" to "severe" extremes. Figure 7 shows that there is co-occurrence of very high JJA total exceedance anomalies within the threshold pairs, but not always. Only the T_wX27 with TX30 pair over the China/East Asia region shares the year with the highest anomaly of JJA exceedance, which is 2022.

Figure 7 identifies several potential "stealth heat event" JJA seasons where T_wX "moderate" to "high" threshold exceedances were relatively high while TX exceedances remained relatively low. Here, such seasons are identified where T_wX exceedance anomalies are positive and TX exceedance anomalies are negative or much lower, for most of the threshold pairs. Such an analysis is subjective and illustrative rather than conclusive. For the UK/Europe region, 1987, 1988, 1989, 1998, 1999, 2003, 2012 and 2017 are examples of these. Over China/East Asia, 1973, 1990, 1998, 2016 and 2020 are examples but, overall, there is lower variability and more frequent co-occurrence of high humidity

and high temperature events. Table 2 lists the top 10 years of JJA exceedance anomalies for each threshold and region. It identifies years of possible "humid and hot heat events" (in bold) where the year appears in both a T_wX and TX threshold top 10, years of potential "stealth heat events" (in blue) where the year appears in a T_wX threshold top 10 but not in any TX threshold top 10s, and possible "dry and hot heat events" (in red) where a year appears in a TX threshold top 10 but not in any T_wX threshold top 10s. The caveat here is that the time scales are broad. Treating the three threshold pairs collectively (T_wX25 and TX25, T_wX27 and TX30, T_wX29 and TX35), there are five potential "stealth heat events" for the UK/Europe region and for China/East Asia. For "dry and hot heat events", there are nine for the UK/Europe region and seven for the China/East Asia region. There are eight "humid and hot events" for the UK/Europe region and nine for the China/East Asia region. This demonstrates the value of a T_w-based extremes index product to be used alongside traditional T-focussed indices to capture the extreme humid heat events that would otherwise be missed. Furthermore, Freychet et al. (2020) noted different trends in T and T_w means, extremes and diurnal temperature ranges, concluding that different processes govern the changes in T compared to T_w. The majority of top 10 years for both variables are post 2000, in line with the significant long-term increasing trends of numbers of days equal to or exceeding extreme thresholds.

Of these top 10 JJA threshold exceedance anomalies, several are noteworthy regional extreme events. For example, over the China/East Asia region, detected heat events included:

• June to August 2022 — record-breaking heat wave for intensity, duration, spatial extent and impact (Guardian, 2022);

• July 2021 — the hottest month on record (at the time) for Asia and also the global average (NOAA, 2021);

• August 2020 — record heat over southern Asia and warmest Northern Hemisphere August on record (NOAA Climate.gov, 2020);

• July 2017 — a Chinese heat wave that was attributable to human influence (Sparrow et al., 2018);

• July 2016 — a heat wave extending over Southwest Asia and the Middle East (NASA Earth Observatory, 2016); and

• July 1994 — an extensive East Asian summer drought (Park and Schubert, 1997).

Over the UK/Europe region, detected heat events included:

• July 2022 — a record-breaking heat wave over southwestern Europe, particularly for the UK (Copernicus, 2022);

• July 2018 — a long heat wave and drought over northwestern Europe (Rösner et al., 2019);

• July 2006 — a record-breaking heat wave over western Europe (Rebetez et al., 2009; Russo et al., 2015);

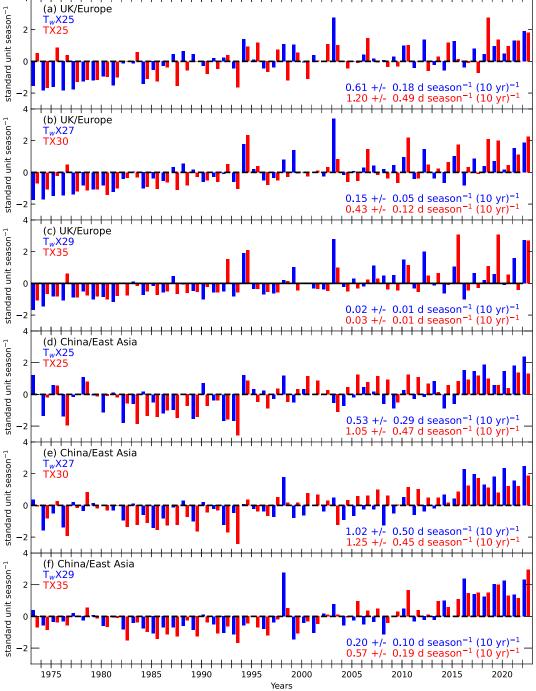• August 2003 — a record-breaking heat wave over Europe (Met Office, 2003; Stott et al., 2004; Rebetez et al.,

Fig. 7. JJA standardised anomaly time series of threshold exceedance for T_wX (blue) and TX (red) for the (a–c) UK/Europe and (d–f) China/East Asia region means. Regional extents are as presented in Fig. 3 but only gridboxes with at least 70% temporal completeness. The data have been screened to remove any gridboxes where the HQ Flag > 6. Decadal trends in seasonal total counts are shown for each region and threshold, fitted using ordinary least-squares regression with 90th percentile confidence intervals corrected for AR(1) following Santer et al. (2008).

2009; Russo et al., 2015); and

• August 1994 — a severe heat wave over Central Europe (Russo et al., 2015; Climate Change Post, 2018).

This demonstrates the ability of HadISDH.extremes to capture regional-scale extreme high temperature events. The humidity aspect is hard to verify as this is rarely reported on.

None of the potential "stealth heat event" years, as identified by years in blue in Table 2, appear to co-occur with notable heat waves. This is as expected and concurs with the theory that such events are often not lethal but could still have impacts on productivity and wellbeing, which are far harder to monitor.

Table 2. Comparison of top 10 years of JJA exceedances for the T_wX and TX threshold pairs for both regions. Only threshold pairs for where there are at least 10 JJA seasons with > 0 days of exceedances for both T_wX and TX are shown. Years shown in black (bold) are "humid and hot heat events" because they appear in both (paired) T_wX and TX threshold top 10s. Years shown in blue are potential "stealth heat events", where they are a top 10 for the T_wX threshold but do not appear in any of the top 10 TX thresholds. Years shown in red are "dry heat events", where they are a top 10 for the TX threshold but do not appear in any of the T_wX thresholds.

Region	Top 10 JJAs (high to low)					
	T_wX25	TX25	T_wX27	TX30	T_wX29	TX35
UK/Europe	2003	2018	**2003**	**1994**	**2022**	**2019**
	2022	**2022**	**2022**	**2022**	**2003**	**2015**
	1994	**2019**	**1994**	**2010**	1994	**2022**
	2021	1975	**2021**	2018	2012	**1994**
	2012	2006	2012	**2019**	2021	1992
	2015	**2021**	1999	**2015**	**2010**	**2010**
	1998	1973	**2015**	2006	**2015**	2018
	1999	2014	**2010**	**2021**	1999	**2003**
	2019	1995	**2019**	**2003**	2007	1976
	2010	**2010**	2017	1976	**2019**	2013
China/East Asia	2022	**2000**	**2022**	**2022**	1998	**2022**
	1973	2010	**2016**	**2017**	**2022**	**2019**
	2018	2007	2020	2000	**2016**	2010
	1998	2011	**2017**	**2010**	**2020**	**2017**
	2016	2005	1998	2007	**2019**	**2018**
	2017	**2021**	2019	**2016**	**2017**	**2016**
	2021	**2017**	**2021**	2011	2018	**2020**
	1994	2015	2018	**2021**	**2021**	2015
	2000	2001	2014	2001	2014	**2021**
	2020	2008	**2010**	2015	2003	2013

3.5. *Spatial analysis of trends*

HadISDH.extremes reveals that T_w extremes can differ in location to T equivalent extremes, and that seasons of high threshold exceedances do not always co-occur. This is also the case for location of the strongest trends to some degree. Figure 8 shows global maps of decadal trends in T_wX90p, $TX90p$, T_wX, TX, T_wX27 and $TX30$ over the period 1973–2022. Trends are almost exclusively positive and significant for the moderate extremes frequency (T_wX90p and $TX90p$) and peak magnitude of extremes (T_wX and TX). This is less so for specific threshold exceedances of T_wX27 and $TX30$, which is to be expected because thresholds are only regionally and not globally applicable. For T_wX27, significant trends are mostly positive, limited in extent to the tropics. For $TX30$ trends are almost exclusively positive, with widespread significance across the tropics, extratropics and much of the midlatitudes. The widespread positive significant trends in T_wX and TX are in line with previous findings using ERA5 from Yu et al. (2021, Fig. 3).

Broadly, there are similarities between T and T_w extremes. The larger trends in moderate extremes frequency for both lie over the tropics, while the larger trends in the peak magnitude of extremes for both lie over the midlatitudes, especially the European continent. As noted above, T_wX27 is not yet relevant for many regions outside of the tropics, whereas for $TX30$ larger trends are more latitudinally

extensive. At the gridbox level, there are differences in locations of the largest trends between T and T_w. Trends over Australia, Europe and the Persian Gulf are stronger for T extremes, whereas trends over Southeast Asia are stronger for T_w extremes. Other research has detected an impact of local irrigation on humid heat extremes using climate models (Krakauer et al., 2020). This small-scale regional detail is clearly important for studying societal impacts.

4. Discussion and conclusions

HadISDH.extremes provides the gridbox mean of monthly maximums of the daily maximum wet bulb temperature (T_wX). This is a measure of the peak magnitude of extremes. A significant increase in T_wX is detectable in the global annual mean and for the UK/Europe region JJA mean. These are smaller than both the equivalent trend in mean T_w and also the trends in maximum dry bulb temperature (TX). There is a similar story for the China/East Asia region, although the positive trends in T_wX are small and not significant. It is possible that the observations from China contain an inhomogeneity that reduces the moistening trend, especially in the peak magnitude of T_wX.

Indices presenting the frequency of "moderate" to "severe" events are also provided by HadISDH.extremes. These include the globally relevant gridbox mean monthly total of days exceeding the 90th percentile of daily maxima
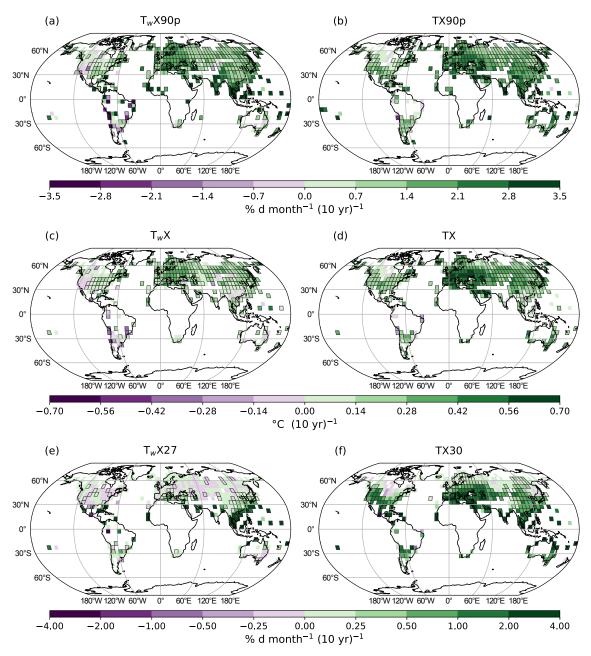
Fig. 8. Decadal trends in monthly anomalies of several T_wX and TX extremes indices from 1973 to 2021. The data have been screened to remove gridboxes where the HQ Flag > 6 and only include gridboxes with 70% data completeness. Decadal trends are fitted using ordinary least-squares regression with 90th percentile confidence intervals corrected for AR(1) following Santer et al. (2008).

(T_wX90p) and regionally relevant gridbox mean monthly total of days where T_wX is equal to or exceeding specific thresholds (T_wX25 to T_wX35). Equivalents are provided for T (TX90p, TX25 to TX50). Trends in the JJA frequency of exceedance of "moderate" thresholds (T_wX90p, T_wX25 and T_wX27) are significantly positive for both the China/East Asia and UK/Europe regions. For the "high" extremes thresholds (T_wX29 and T_wX31), only the China/East Asia region shows significant positive trends, although there are months where days exceed these thresholds for the UK/Europe region.

HadISDH.extremes shows that gridboxes over some

regions have apparently already experienced days where the T_wX reaches 35°C, including the Persian Gulf and Australia. Uncertainty can be large in these isolated exceedances of severe extreme thresholds and exceedances over Australia cannot be verified against national data holdings.

In all cases and regions, the trends in TX-related thresholds are larger than for T_wX, confirming that T is the key driver of heat extremes over these regions. However, it is clear that humidity is playing an important role too. There is some evidence that T_wX extremes can occur when TX is more moderate; in effect, these are "stealth heat events" that may not be detected by current temperature-focussed warning

systems, yet can have notable impacts on productivity and general health as opposed to mortality.

At the gridbox level, significant positive trends in both the peak magnitude of extremes (T_wX) and frequency of moderate extreme events (T_wX90p) are widespread. This is less so for specific thresholds because even the lowest threshold of T_wX25 has a limited spatial extent of relevance – it shows significant positive trends across most of the tropics (not shown). Collectively, this shows that the humid heat component of heat waves is increasing across the globe. For most regions outside of the tropics this only occasionally reaches levels that affect productivity or health. However, the widespread increasing trends show increasing risk for these regions. For regions in the tropics, T_w is already reaching levels that can potentially cause significant impacts.

HadISDH.extremes is based on hourly station data from all over the globe. These data have been quality controlled at the hourly level and various temporal completeness tests are made at the daily, monthly, annual, climatological and full record level to remove poorer quality, intermittent and short-record stations. The parent dataset, HadISDH.land, which is based on the same station data, compares well with estimates from ERA5 (Simmons et al., 2021), inferring that HadISDH.extremes should be of reasonable quality. However, HadISDH.extremes is more susceptible to errors and biases in the station data because its monthly values are based on either a single occasion in the month (maximum daily maxima) or counts of days exceeding a specific threshold. There is no averaging across the month and therefore no minimising of any random error. Monthly anomalies from the 1991–2020 climatology of these values are produced and averaged over the stations within each gridbox, reducing the effect of differing altitudes, aspects and land cover. For assessment of long-term trends, it is safer to use these gridbox mean anomalies than the gridbox mean of actual values. The results shown here are in line with other studies on T_w and humid heat extremes (Freychet et al., 2020; Raymond et al., 2020; Yu et al., 2021). However, users may well find differences at smaller spatial and temporal scales, especially if using actual values rather than the anomalies because these will be strongly dependent on the locations of the underlying station data. Furthermore, the HadISDH.extremes (T_wX and TX) will always likely be an underestimate of the true value because the origin observations are discrete in time, sampling at most hourly and often 3 hourly or less frequently. Caution is required over gridbox level interpretation of the threshold exceedance indices especially, with cross-validation against independent information recommended. Treating these as grouped "moderate" (T_wX25 and T_wX27), "high" (T_wX29 and T_wX31) and "severe" (T_wX33 and T_wX35) extremes thresholds for making comparative statements or looking at long-term changes is safer than identifying specific times and locations of day counts.

HadISDH.extremes can be used to make robust conclusions when the following criteria are taken into consideration. Firstly, any statistical analysis should be based on a data selection where the number of non-missing and non-zero (count) data points is high enough to provide sufficient degrees of freedom. Trends in time series of T_wX33, for example, where there may be many months or years of zero counts, could be strongly driven by outliers. Secondly, users should look for regional consistency in signals rather than drawing conclusions from single gridboxes to reduce the vulnerability to outliers. Thirdly, the time series themselves should be analysed to ensure that any fitted trend is a sensible representation of the data. Finally, where possible, users should cross-check with independent evidence such as national records or media reports of extremes. Further information on the dataset and advice on its usage can be found in Willett (2023a).

The known issue of a region-wide shift from manual wet bulb thermometers to automated RH sensors over China is likely still an issue for HadISDH.extremes. This is both because the homogenisation process used to provide homogeneity scores may be under-detecting this region-wide change and also because extreme values are more susceptible to inhomogeneity error than mean values. Hence, the HQ Flag scores over China may not fully identify poor quality gridboxes. Trends in T_w extremes are considerably smaller for the HadISDH.extremes China/East Asia region compared to those from a homogenised Chinese dataset (Freychet et al., 2020). However, significant positive trends are detectable in HadISDH.extremes in the days equal to or exceeding thresholds (specific and 90th percentile).

Following on from the concept of the HadISDH family (HadISDH.land, HadISDH.marine and HadISDH.blend), HadISDH.extremes is a coarse-resolution product designed for spatial and temporal stability with which to detect large-scale long-term features. It is designed for exploring current regional exposure to extremes, tracking the evolution of extremes exposure over time, and comparing with historical reconstructions from climate models. The latter aids model validation, which in turn provides confidence in future projections of such extremes and their related climate impacts. Heat extremes are often local and short-lived features. A region like the UK for example tends to experience heat waves of a few days with very high heat unlikely to simultaneously cover the entire country. However, there have been several regional-scale and longer-lived heat events that are detectable with HadISDH.extremes, such as the UK/European region heat waves in 2022, 2018, 2006, 2003 and 1994, and the China/East Asian region heat waves in 2022, 2021, 2020, 2017, 2016 and 1994.

Although several studies have now explored heat extremes around the globe based on T_w, HadISDH.extremes is the first dedicated product for monitoring them. It is envisaged that updates will continue alongside the rest of the HadISDH family. This provides a stable base with which to assess long-term and year-on-year changes in humid and dry heat extremes, alongside changes in the mean from HadISDH.land.

The provision of simultaneously observed dry and wet

bulb heat extremes, from a near-identical station base and identical methodological processing provides the opportunity to study co-occurrence of dry and humid heat extremes. It is possible for an event to be moderate in terms of T, yet impactful because of its extreme humidity. Such events may not be linked to mortality to the same degree as high temperature heat events, but their economic and general health and wellbeing impact can still be considerable (Parsons et al., 2022).

HadISDH.extremes provides the ability to assess the relative exposure to humid heat extremes and changes over time across different regions in addition to global and hemispheric analyses. The framework of HadISDH.extremes can be expanded to produce different indices as required. Other valuable future developments include a dedicated uncertainty model for the extremes and a higher resolution product over the data-rich regions. Such regions are limited geographically, and many stations suffer from intermittency, inhomogeneity or limited record length. Ultimately, this limits our ability to quantify and understand current exposure and predict future exposure in the detail required for devising optimum city- or district-level adaptation. Knowledge and understanding of extremes is considerably limited across much of the globe, especially Africa, owing to data sparsity and/or lack of sharing of the sub-daily data required to characterise the extremes. Furthermore, extremes are more susceptible to errors and biases than means, and more difficult to quality control and homogenise. To reduce the uncertainty in an extremes monitoring product such as HadiSDH.extremes requires a high quality, observation-dense network with detailed digitised metadata about the station and instrument setup and any changes, all of which are openly shared. At present, this is not the case for the majority of the land mass. However, while uncertainty can be large when assessing actual extreme values rather than anomalies, and very localised features, especially where observations are sparse, there is confidence in the large-scale, long-term changes presented by HadISDH.extremes.

HadISDH.extremes may serve as a proof-of-concept product for a similar approach to the HadEX family of indices. Such a product would ideally utilise dense networks of high-quality homogeneous T_w data provided as indices by countries themselves, presenting improved coverage and quality over HadISDH.extremes. Until such a product exists, HadISDH.extremes will hopefully prove a useful tool in developing our understanding of humid heat extremes and their evolution over time.

5. Sources of datasets

The HadISDH.extremes dataset is available in netCDF format under the Open Government Licence (https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/). The data can be accessed from the Centre for Environmental Data Analysis (CEDA) Archive (Willett, 2023b; https://catalogue.ceda.ac.uk/uuid/2d1613955e1b4cd1b156e5f3edbd7e66). It is also available alongside other related material and information from the Met Office Hadley Centre climate monitoring observations portal, HadOBS: www.metoffice.gov.uk/hadobs/hadisdh/.

Additionally, the land-based Met Office Hadley Centre's Integrated Surface Dataset of Humidity, HadISDH.land.4.5.1.2022f (Willett et al., 2023), was used within this paper. This is the latest version available at time of writing. It is downloadable as netCDF files from www.metoffice.gov.uk/hadobs/hadisdh under the Open Government Licence (https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/). It is also available from the CEDA Archive: https://catalogue.ceda.ac.uk/uuid/8956cf9e31334914ab4991796f0f645a.

REFERENCES

ADPC (Asian Disaster Preparedness Center), 2007: Disaster News. Available from http://www.adpc.net/V2007/IKM/EVENTS%20AND%20NEWS/DISASTER/2007/June/184%20killed%20in%20heatwave%20across.asp.

Ahmed, AS. F., KM. M K. Khan, AA. Maung Than Oo and RM. G. Rasul, 2014: Selection of suitable passive cooling strategy for a subtropical climate. *Int. J. Mech. Mater Eng*, **9**, 14, https://doi.org/10.1186/s40712-014-0014-7.

Armstrong, B., and Coauthors, 2019: The role of humidity in associations of high temperature with mortality: A multicountry, multicity study. *Environmental Health Perspectives*, **127**(9), 097007, https://doi.org/10.1289/EHP5430.

Brimicombe, C., C. Di Napoli, T. Quintino, F. Pappenberger, R. Cornforth, and H. L. Cloke, 2022: Thermofeel: A python thermal comfort indices library. *SoftwareX*, **18**, 101005, https://doi.org/10.1016/j.softx.2022.101005.

Budd, G. M., 2008: Wet-bulb globe temperature (WBGT)—its history and its limitations. *Journal of Science and Medicine in Sport*, **11**(1), 20–32, https://doi.org/10.1016/j.jsams.2007.07.003.

Climate Change Post, 2018: In the near future twice as many heat waves over Central Europe. Available from https://www.climatechangepost.com/news/2018/3/2/near-future-twice-many-heat-waves-over-central-eur/.

Climpact, 2022: Climpact open source R software for calculating

sector-specific extremes indices. Available from https://climpact-sci.org/.

Dunn, R. J. H., and Coauthors, 2020: Development of an updated global land in situ-based data set of temperature and precipitation extremes: HadEX3. *J. Geophys. Res.: Atmos.*, **125**, e2019JD032263, https://doi.org/10.1029/2019JD032263.

ET-SCI, 2022: Expert Team for Sector-Specific Indices webpage. Available from https://public.wmo.int/en/events/meetings/expert-team-sector-specific-climate-indices-et-sci.

Freychet, N., S. F. B. Tett, Z. Yan, and Z. Li, 2020: Underestimated change of wet-bulb temperatures over East and South China. *Geophys. Res. Lett.*, **47**, e2019GL086140, https://doi.org/10.1029/2019GL086140.

Guardian, 2022: Available from https://www.theguardian.com/world/2022/sep/07/china-reports-most-severe-heatwave-and-lowest-rainfall-on-record .

Jendritzky, G., R. de Dear, and G. Havenith, 2012: UTCI—Why another thermal index? *International Journal of Biometeorology*, **56**, 421–428, https://doi.org/10.1007/s00484-011-0513-7.

Krakauer, N. Y., B. I. Cook, and M. J. Puma, 2020: Effect of irrigation on humid heat extremes. *Environmental Research Letters*, **15**, 094010, https://doi.org/10.1088/1748-9326/ab9ecf.

Li, Z., Z. W. Yan, Y. N. Zhu, N. Freychet, and S. Tett, 2020: Homogenized daily relative humidity series in China during 1960–2017. *Adv. Atmos. Sci.*, **37**(4), 318–327, https://doi.org/10.1007/s00376-020-9180-0.

Mazdiyasni, O., and Coauthors, 2017: Increasing probability of mortality during Indian heat waves. *Science Advances*, **3**(6), e1700066, https://doi.org/10.1126/sciadv.1700066.

Met Office, 2003: The heatwave of 2003. Available from https://www.metoffice.gov.uk/weather/learn-about/weather/case-studies/heatwave.

NASA Earth Observatory, 2016: Extreme heat for an extreme year. Available from https://earthobservatory.nasa.gov/images/88547/extreme-heat-for-an-extreme-year.

NOAA, 2021. It's official: July was Earth's hottest month on record. News and Features. Available from https://www.noaa.gov/news/its-official-july-2021-was-earths-hottest-month-on-record#:~:text=July%202021%20has%20earned%20.

NOAA Climate.gov, 2020: August 2020: The warmest summer on record for the Northern Hemisphere comes to an end. Available from https://www.climate.gov/news-features/understanding-climate/august-2020-warmest-summer-record-northern-hemisphere-comes-end.

Park, C. K., and S. D. Schubert, 1997: On the nature of the 1994 East Asian summer drought. *J. Climate*, **10**(5), 1056–1070, https://doi.org/10.1175/1520-0442(1997)010<1056:OTNOTE>2.0.CO;2.

Parsons, L. A., Y. J. Masuda, T. Kroeger, D. Shindell, N. H. Wolff, and J. T. Spector, 2022: Global labor loss due to humid heat exposure underestimated for outdoor workers. *Environmental Research Letters*, **17**(1), 014050, https://doi.org/10.1088/1748-9326/ac3dae.

Raymond, C., D. Singh, and R. M. Horton, 2017: Spatiotemporal patterns and synoptics of extreme wet-bulb temperature in the contiguous United States. *J. Geophys. Res.: Atmos.*, **122**, 13 108–13 124, https://doi.org/10.1002/2017JD027140.

Raymond, C., T. Matthews, and R. M. Horton, 2020: The emergence of heat and humidity too severe for human tolerance. *Science Advances*, **6**(19), eaaw1838, https://doi.org/10.1126/sciadv.aaw1838.

Raymond, C., T. Matthews, R. M. Horton, E. M. Fischer, S. Fueglistaler, C. Ivanovich, L. Suarez-Gutierrez, and Y. Zhang, 2021: On the controlling factors for globally extreme humid heat. *Geophys. Res. Lett.*, **48**, e2021GL096082, https://doi.org/10.1029/2021GL096082.

Rebetez, M., O. Dupont, and M. Giroud, 2009: An analysis of the July 2006 heatwave extent in Europe compared to the record year of 2003. *Theor. Appl. Climatol.*, **95**, 1–7, https://doi.org/10.1007/s00704-007-0370-9.

Reuters, 2007: Dozens more people killed by South Asia heat wave. Available from https://www.reuters.com/article/environment-southasia-heatwave-dc-idUSDEL29044220070612.

Rösner, B., I. Benedict, C. C. van Heerwaarden, A. H. Weerts, W. Hazeleger, P. Bissolli, and K. Trachte, 2019: Sidebar 7.3: The long heat wave and drought in Europe in 2018. *State of the Climate in 2018*, J. Blunden and D. S. Arndt, Eds., Bulletin of the American Meteorological Society, 100 (9), S222–S223. https://doi.org/10.1175/2019BAMSStateofthe Climate.1#page=241.

Russo, S., J. Sillmann, and E. M. Fischer, 2015: Top ten European heatwaves since 1950 and their occurrence in the coming decades. *Environmental Research Letters*, **10**, 124003, https://doi.org/10.1088/1748-9326/10/12/124003.

Santer, B. D., and Coauthors, 2008: Consistency of modelled and observed temperature trends in the tropical troposphere. *International Journal of Climatology*, **28**, 1703–1722, https://doi.org/10.1002/joc.1756.

Schär, C., 2016: The worst heat waves to come. *Nature Climate Change*, **6**, 128–129, https://doi.org/10.1038/nclimate2864.

Schwingshackl, C., J. Sillmann, A. M. Vicedo-Cabrera, M. Sandstad, and K. Aunan, 2021: Heat stress indicators in CMIP6: Estimating future trends and exceedances of impact-relevant thresholds. *Earth's Future*, **9**, e2020EF001885, https://doi.org/10.1029/2020EF001885.

Shen, D. D., and N. Zhu, 2015: Influence of the temperature and relative humidity on human heat acclimatization during training in extremely hot environments. *Building and Environment*, **94**, 1–11, https://doi.org/10.1016/j.buildenv.2015.07.023.

Sherwood, S. C., and M. Huber, 2010: An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 9552–9555, https://doi.org/10.1073/pnas.0913352107.

Simmons, A. J., and Coauthors, 2021: Low frequency variability and trends in surface air temperature and humidity from ERA5 and other datasets. European Centre for Medium-range Weather Forecasting (ECMWF) Technical Memoranda 881, ECMWF, Shinfield Park, Reading, UK, 97 pp, https://doi.org/10.21957/ly5vbtbfd.

Sparrow, S., and Coauthors, 2018: Attributing human influence on the July 2017 Chinese heatwave: The influence of sea-surface temperatures. *Environmental Research Letters*, **13**, 114004, https://doi.org/10.1088/1748-9326/aae356.

Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature*, **432**, 610–614, https://doi.org/10.1038/nature03089.

Vecellio, D. J., S. T. Wolf, R. M. Cottle, and W. L. Kenney, 2022: Evaluating the 35°C wet-bulb temperature adaptability threshold for young, healthy subjects (PSU HEAT Project). *Journal of Applied Physiology*, **132**(2), 340–345, https://doi.org/10.1152/japplphysiol.00738.2021.

Wang, P. Y., L. R. Leung, J. Lu, F. F. Song, and J. P. Tang, 2019: Extreme wet-bulb temperatures in China: The significant role of moisture. *J. Geophys. Res.: Atmos.*, **124**, 11 944–11 960, https://doi.org/10.1029/2019JD031477.

Willett, K. M., 2023a: HadISDH.extremes Part I: A gridded wet bulb temperature extremes index product for climate monitoring. *Adv. Atmos. Sci.*, https://doi.org/10.1007/s00376-023-2347-8.

Willett, K. M., 2023b: HadISDH.extremes: Gridded global monthly land surface wet bulb and dry bulb temperature extremes index data version 1.0.0.2022f. NERC EDS Centre for Environmental Data Analysis, https://doi.org/10.5285/2d1613955e1b4cd1b156e5f3edbd7e66.

Willett, K. M., and S. Sherwood, 2012: Exceedance of heat index thresholds for 15 regions under a warming climate using the wet-bulb globe temperature. *International Journal of Climatology*, **32**(2), 161–177, https://doi.org/10.1002/joc.2257.

Willett, K. M., C. N. Williams Jr., R. J. H. Dunn, P. W. Thorne, S. Bell, M. de Podesta, P. D. Jones, and D. E. Parker, 2013: HadISDH: An updateable land surface specific humidity product for climate monitoring. *Climate of the Past*, **9**, 657–677, https://doi.org/10.5194/cp-9-657-2013.

Willett, K. M., R. J. H. Dunn, P. W. Thorne, S. Bell, M. de Podesta, D. E. Parker, P. D. Jones, and C. N. Williams Jr., 2014: HadISDH land surface multi-variable humidity and temperature record for climate monitoring. *Climate of the Past*, **10**, 1983–2006, https://doi.org/10.5194/cp-10-1983-2014.

Willett, K. M., R. J. H. Dunn, P. W. Thorne, S. Bell, M. de Podesta, D. E. Parker, P. D. Jones, and C. N. Williams Jr., 2023: HadISDH.land: Gridded global monthly land surface humidity data version 4.5.1.2022f. NERC EDS Centre for Environmental Data Analysis, https://doi.org/10.5285/8956cf9e31334914ab4991796f0f645a.

World Weather Attribution, 2016: Record high temperatures in India, 2016. Available from https://www.worldweatherattribution.org/india-heat-wave-2016/.

Yu, S., S. F. B. Tett, N. Freychet, and Z. W. Yan, 2021: Changes in regional wet heatwave in Eurasia during summer (1979–2017). *Environmental Research Letters*, **16**, 064094, https://doi.org/10.1088/1748-9326/ac0745.

Zare, S., N. Hasheminejad, H. E. Shirvan, R. Hemmatjo, K. Sarebanzadeh, and S. Ahmadi, 2018: Comparing Universal Thermal Climate Index (UTCI) with selected thermal indices/environmental parameters during 12 months of the year. *Weather and Climate Extremes*, **19**, 49–57, https://doi.org/10.1016/j.wace.2018.01.004.

Zhang, Y., I. Held, and S. Fueglistaler, 2021: Projections of tropical heat stress constrained by atmospheric dynamics. *Nature Geoscience*, **14**, 133–137, https://doi.org//10.1038/s41561-021-00695-3.