Earth's Future

RESEARCH ARTICLE
10.1029/2020EF001885

Heat Stress Indicators in CMIP6: Estimating Future Trends and Exceedances of Impact-Relevant Thresholds

Clemens Schwingshackl[1], Jana Sillmann[1], Ana Maria Vicedo-Cabrera[2,3], Marit Sandstad[1], and Kristin Aunan[1]

[1]Center for International Climate Research (CICERO), Oslo, Norway, [2]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland, [3]Oeschger Center for Climate Change Research, University of Bern, Bern, Switzerland

Key Points:

- All heat stress indicators increase statistically significantly with global mean temperature but trends reveal a substantial spread
- Exceedances of impact-relevant thresholds are strongly increasing globally including in several densely populated regions
- For assessing heat-related health impacts only indicators and thresholds should be chosen that were validated for the respective application

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

K. Aunan,
kristin.aunan@cicero.oslo.no

Citation:

Schwingshackl, C., Sillmann, J., Vicedo-Cabrera, A. M., Sandstad, M., & Aunan, K. (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

Received 31 OCT 2020
Accepted 29 JAN 2021
Corrected 7 JUN 2021

This article was corrected on 7 JUN 2021. See the end of the full text for details.

Author Contributions:

Conceptualization: Clemens Schwingshackl, Jana Sillmann, Kristin Aunan
Data curation: Marit Sandstad
Formal analysis: Clemens Schwingshackl
Funding acquisition: Kristin Aunan

Abstract

Global warming is leading to increased heat stress in many regions around the world. An extensive number of heat stress indicators (HSIs) has been developed to measure the associated impacts on human health. Here we calculate eight HSIs for global climate models participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6). We compare their future trends as function of global mean temperature, with particular focus on highly populated regions. All analyzed HSIs increase significantly ($p < 0.01$) in all considered regions. Moreover, the different HSIs reveal a substantial spread ranging from trends close to the rate of global mean temperature up to an amplification of more than a factor of two. Trends change considerably when normalizing the HSIs by accounting for the different scales on which they are defined, but the large spread and strong trends remain. Consistently, exceedances of impact-relevant thresholds are strongly increasing globally, including in several densely populated regions, but also show substantial spread across the selected HSIs. The indicators with the highest exceedance rates vary for different threshold levels, suggesting that the large indicator spread is associated both to differences in trend magnitude and the definition of threshold levels. These results highlight the importance of choosing indicators and thresholds that are appropriate for the respective impact under consideration. Additionally, further validation of HSIs regarding their capability to quantify heat impacts on human health on regional-to-global scales would be of great value for assessing global impacts of future heat stress more reliably.

Plain Language Summary

Heat stress caused by high levels of air temperature and humidity is rising globally due to climate change. Various indicators for heat stress have been developed to quantify different facets of how heat impacts people. We use data from climate models to calculate the future evolution of eight heat stress indicators for highly populated regions of the world. The trends of the different indicators vary substantially, with some indicators showing large increases while others only increase modestly. For estimating the severity of heat stress, we calculate how often each indicator exceeds threshold values that indicate different heat stress severity. Many thresholds will be exceeded much more often with rising temperatures. The increases are particularly large for some indicators while others only show small increases. Moreover, the indicators with the strongest trends are often not the ones that show the highest increase in threshold exceedances. For quantifying the impacts of heat stress caused by climate change it is thus important to choose indicators that are appropriate for the respective application. While several indicators were tested on small scales (e.g., in cities or single countries), for global heat stress assessments it is necessary to have more validation studies on regional-to-global scales.

1. Introduction

Human exposure to heat is growing globally as temperatures have been rising in the recent decades and are projected to further increase in the future (e.g., Diffenbaugh et al., 2007; Huang et al., 2011; Knutson & Ploshay, 2016; Russo et al., 2017). Without appropriate acclimatization and adaptation (Bobb et al., 2014; Gasparrini et al., 2015), heat-related impacts on human health are expected to strongly intensify with global warming as extreme temperatures over land are increasing faster than average temperatures (Sillmann et al., 2013; Vogel et al., 2017) and impacts of extreme heat are often nonlinear, for example, for heat-related loss of worker productivity (Dunne et al., 2013; Kjellstrom et al., 2013) or, more severe, for premature death caused by strong and extreme heat (McMichael et al., 2006; Gasparrini et al., 2015; Mora et al., 2017).

Methodology: Clemens Schwingshackl, Jana Sillmann
Project Administration: Kristin Aunan
Visualization: Clemens Schwingshackl
Writing – original draft: Clemens Schwingshackl
Writing – review & editing: Clemens Schwingshackl, Jana Sillmann, Ana Maria Vicedo-Cabrera, Marit Sandstad, Kristin Aunan

Additionally, health assessments including population aging project an increase in the overall vulnerability to heat with concomitant increases in mortality (Chen et al., 2020). While heat impacts on small scales (e.g., for different cities, age groups, seasons) have been examined in multiple studies (e.g., Armstrong et al., 2019; Barnett et al., 2010; Gasparrini et al., 2015; Morabito et al., 2014), global assessments quantifying the impacts of heat increase due to climate change are still difficult due to the lack of information on health risks at regional-to-global scales (Gasparrini et al., 2017).

Quantifying the impacts of extreme heat on human physiology is a nontrivial task as it involves modeling the complex heat transfer between the human body and the surrounding air (e.g., Cheng et al., 2012). Many attempts have been made to express human heat stress as function of ambient conditions, resulting in the development of a considerable number of heat stress indicators (HSIs) with varying complexity (de Freitas & Grigorieva, 2017). Here we use the term "heat stress" for heat-related impacts on human health that are associated with high levels of air temperature, humidity, and solar and thermal radiation. Many HSIs consider these variables in their formulation, and some also include additional variables such as wind speed.

A wide range of fields apply HSIs to quantify heat stress. Meteorological services use indicators such as the NOAA heat index (Rothfusz, 1990; Steadman, 1979), Humidex (Masterson & Richardson, 1979; as cited in Blazejczyk et al., 2012) or Perceived Temperature (Staiger et al., 2011) to provide heat warnings. Wet-bulb globe temperature is often used to assess occupational illness and injuries caused by heat stress (e.g., Garzon-Villalba et al., 2016), to estimate heat-related decreases of worker productivity (e.g., Kjellstrom et al., 2009; Orlov et al., 2019), to derive heat safety recommendations for athletes (Grundstein et al., 2015; Racinais et al., 2015), and for military purposes (Budd, 2008). Several HSIs were originally developed to measure the thermal comfort of humans. One important group of such indicators is called "apparent temperature" (Anderson et al., 2013), with the NOAA heat index being the most known representative of this group. More sophisticated estimations of thermal comfort consider in detail the heat transfer between the human body and the environment, with examples being the Physiological Equivalent Temperature (PET; Höppe, 1999) or the Universal Thermal Climate Index (UTCI; Fiala et al., 2012).

Despite the widespread usage of HSIs, epidemiological studies considering multiple indicators (i.e., temperature and different HSIs) often find comparable predictive skills for adverse health effects with no single indicator being superior to the others (e.g., Barnett et al., 2010; Burkart et al., 2011; Heo & Bell, 2018; Kim et al., 2011; Ragettli et al., 2017; Vaneckova et al., 2011). Additionally, the indicator showing the highest impacts on mortality or morbidity rates varies across geographic locations, seasons, or age groups (Barnett et al., 2010; Chung et al., 2009; Heo et al., 2019; Kang et al., 2020; Rodopoulou et al., 2015). To date, evidence from epidemiological studies suggest that HSI perform similarly well as temperature and, thus, do not seem to be particularly more suitable for predicting adverse health effects related to heat stress (e.g., Armstrong et al., 2019; Kent et al., 2014; Vaneckova et al., 2011). Moreover, different HSIs may be best suited for assessing particular endpoints, such as physiological heat stress, heat-related mortality, or specific morbidity outcomes (Gronlund et al., 2014; Ross et al., 2018; Zhang et al., 2012). Consequently, many health studies analyzing the effect of heat on human morbidity and mortality often do not use predefined HSIs but employ statistical methods based on temperature and humidity to assess the effects of heat stress on morbidity and mortality (e.g., Armstrong et al., 2019; Gasparrini et al., 2015; Li et al., 2016). As there is no clear evidence on the superiority of any of the HSIs, most studies on health impact projections rely on projections of mean or maximum temperature for estimating future heat-related impacts under climate change scenarios (Gasparrini et al., 2017; Wang et al., 2019).

In contrast to epidemiological studies, climate impact assessments often use HSIs because they constitute a rather simple way of estimating human heat stress from meteorological variables (e.g., Morabito et al., 2014; Spangler et al., 2018; Wang & Zhu, 2019), particularly on regional-to-global scales. Diffenbaugh et al. (2007) and Fischer and Schär (2010) analyzed future heat stress in the Mediterranean and Europe using the NOAA heat index and showed that dangerous heat conditions (as defined by the US National Weather Service) will be reached in central and southern Europe during a substantial number of days each year by the end of the century under moderate to high warming. Using the same indicator, Russo et al. (2017) showed that the probability of humid heatwaves (i.e., co-occurring high temperatures and elevated humidity levels) will strongly increase with global warming. Pal and Eltahir (2016) and Im et al. (2017) employed wet-bulb temperature to show that in some Asian regions even physiological limits of adaptation might be exceeded

Table 1
Definition of the Heat Stress Indicators (HSIs) Used in This Study (ϑ is Temperature in °C and e Vapor Pressure in Pa)

Heat stress indicator	Definition/Formula
TX	Daily maximum near-surface air temperature
AT	$AT = 0.92 \cdot \vartheta + 0.22\ C \cdot e/100\ Pa - 1.3\ °C$
HI	Formula can be found in Text S5 (Supporting Information)
Hu	$Hu = \vartheta + 5/9\ °C \cdot (e/100\ Pa - 10)$
T_{WB}	Formula can be found in Davies-Jones (2008) or Dunne et al. (2013)
T_{WBG}	$T_{WBG} = 0.7 \cdot T_{WB} + 0.3 \cdot \vartheta$
T_{WBGs}	$T_{WBGs} = 0.567 \cdot \vartheta + 0.393\ °C \cdot e/100\ Pa + 3.94\ °C$
UTCI	Formula can be found in the supplementary of Bröde et al. (2012)

under very high warming. In historical observations, significant trends were identified for wet-bulb globe temperature in various regions between the 1970s and the 2000s (Knutson & Ploshay, 2016; Willett & Sherwood, 2012). The frequency of wet-bulb globe temperature extremes is expected to keep increasing with rising global mean temperatures (Li et al., 2020). Wet-bulb globe temperature was further used to quantify future occupational heat exposure (Kjellstrom et al., 2013) and the projected economic costs of reduced worker productivity caused by heat stress increases due to climate change (Orlov et al., 2020). Future projections for HSIs that include temperature and humidity are less uncertain than temperature projections due to the co-variability between temperature and humidity (i.e., changes in one variable also impact the other), which makes trends in these HSIs generally rather robust (Fischer & Knutti, 2013; Willett & Sherwood, 2012).

Here, we provide a comparison of the projected evolution of eight different HSIs as function of global mean temperature for climate models participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6; Eyring et al., 2016). We only consider temperature, specific humidity, and pressure as input variables and do not use HSIs that account for other variables like wind or solar radiation. Consequently, the selected HSIs represent conditions in the shade without (or with only little) wind. To assess heat stress severity, we further calculate exceedances of impact-relevant thresholds already established in heat warning systems, occupational health, assessing decreases in worker productivity, and thermal comfort models. We apply these thresholds to identify regions for which a substantial increase of human heat stress is projected.

2. Methods and Data

2.1. Heat Stress Indicators

More than 170 HSIs are available for quantifying human heat stress (de Freitas & Grigorieva, 2017), but only a subset is appropriate to be used in climate impact studies. We require that HSIs can be calculated from climate model output (within feasible computing time), represent windless conditions in the shade, are expressed as equivalent temperature, and have been used in climate impact studies (see Supporting Information Text S1 for a complete list of criteria). The following eight HSIs are included in the analysis (see Table 1 for their definition):

2.1.1. Daily Maximum Near-surface Air Temperature (TX)

TX is often used as a simple heat indicator in climate science (e.g., Fischer & Schär, 2010; Sillmann et al., 2013; Shiogama et al., 2019). Several epidemiological studies apply TX in addition to daily mean temperature to estimate heat-related mortality or morbidity (e.g., Barnett et al., 2010; Heo & Bell, 2018; Zhang et al., 2014). In contrast to other HSIs, there is no set of absolute thresholds defined for TX that could be applied on global scales.

2.1.2. Apparent Temperature (AT)

The category "apparent temperature" refers to a group of different HSIs (Anderson et al., 2013), which were all developed based on the original AT formulation of Steadman (1984). We employ an AT version which

only uses temperature and vapor pressure as input variables (Zhao et al., 2015). Different versions of AT were applied in epidemiological studies and found to perform well for estimating heat-related health impacts (e.g., Gronlund et al., 2014; Mitchell et al., 2016; Urban & Kyselý, 2014; Vaneckova et al., 2011).

2.1.3. NOAA Heat Index (HI)

HI is used by the US National Oceanic and Atmospheric Administration (NOAA) for issuing heat warnings and was applied in a number of studies that investigated adverse health impacts due to heat stress (e.g., Burkart et al., 2011; Kent et al., 2014; Lin et al., 2012). HI is calculated as multiple linear regression with temperature and relative humidity as input variables (Rothfusz, 1990; Steadman, 1979).

2.1.4. Humidex (Hu)

Hu is used by meteorological services in Canada (Masterson & Richardson, 1979; as cited in Blazejczyk et al., 2012). It combines temperature and vapor pressure to calculate heat stress. In several studies, Hu was found to perform similarly well as other indicators for predicting heat-related mortality (e.g., Barnett et al., 2010; Heo & Bell, 2018; Vaneckova et al., 2011).

2.1.5. Wet-Bulb Temperature (T_{WB})

T_{WB} considers the cooling capacity of a human body through sweating by indicating the temperature an air parcel would have in case of complete water evaporation. We employ the formula derived by Davies-Jones (2008), which calculates T_{WB} from temperature, relative humidity, and pressure using equivalent potential temperature (Bolton, 1980). While T_{WB} has rarely been used in epidemiological studies, it was applied in several climate impact studies showing that thresholds for human adaptability to heat stress might be exceeded in several regions of the world under very high warming (Im et al., 2017; Pal & Eltahir, 2016).

2.1.6. Wet-Bulb Globe Temperature (T_{WBG})

T_{WBG} in the shade is defined as weighted mean of T_{WB} and near-surface air temperature. It is employed by several meteorological services and is frequently used in climate impact studies (e.g., Blazejczyk et al., 2012; Lemke & Kjellstrom, 2012). T_{WBG} is used for several purposes, specifically for assessments of occupational health (Budd, 2008; Garzon-Villalba et al., 2016) and for quantifying impacts of heat stress on worker productivity (Kjellstrom et al., 2009; Orlov et al., 2019). Further, T_{WBG} was applied in several epidemiological studies and found to perform as well as other indicators in quantifying adverse impacts of heat stress (e.g., Heo & Bell, 2018; Heo et al., 2019; Lin et al., 2012; Urban et al., 2019; Vaneckova et al., 2011).

2.1.7. Simplified Wet-bulb Globe Temperature (T_{WBGs})

T_{WBGs} uses a linear combination of temperature and vapor pressure to approximate T_{WBG} (American College of Sports Medicine, 1984; as cited in Willett & Sherwood, 2012). Many climate impact studies and also several epidemiologial studies employ T_{WBGs} instead of T_{WBG} due to its ease of use (e.g., Fischer & Knutti, 2013; Lin et al., 2012; Urban et al., 2019; Vaneckova et al., 2011; Willett & Sherwood, 2012; Zhao et al., 2015).

2.1.8. Universal Thermal Climate Index (UTCI)

UTCI was developed as multinode model of human heat transfer (Fiala et al., 2012). Here we use the polynomial UTCI approximation based on temperature and vapor pressure developed by Bröde et al. (2012). UTCI was shown to perform similarly well as other indicators for assessing heat-related mortality in Europe (Napoli et al., 2018; Urban & Kyselý, 2014).

2.2. Impact-Relevant Thresholds

The assessment of HSI trends does not allow for direct derivations of society-relevant impacts because HSI trends inherently depend on the scales on which the HSIs are defined. To classify the severity of human heat stress, we thus employ impact-relevant thresholds. While epidemiological studies usually do not consider absolute thresholds, safety regulations for occupational athletic health as well as meteorological heat warnings do rely on specific thresholds (Blazejczyk et al., 2012; Grundstein et al., 2015; Parsons, 2006; Racinais et al., 2015, see also Table S2). Due to the different purposes for which thresholds were developed, no uniform impact assessment scheme exists for HSIs and the thresholds applied in the present study are gathered from various sources. The selected thresholds are used in heat warning systems (HI, Hu), to screen

Table 2
Overview of the HSI Thresholds Used in This Study, Distinguishing Between Four Different Severity Levels (See Table S1 for Detailed Definitions)

	AT	HI	Hu	T_{WB}	T_{WBG} & T_{WBGs}	UTCI
Level 1	28.0 °C	27.0 °C	30.0 °C	-	29.0 °C	26.0 °C
Level 2	32.0 °C	32.0 °C	40.0 °C	-	30.5 °C	32.0 °C
Level 3	35.0 °C	41.0 °C	45.0 °C	-	32.0 °C	38.0 °C
Level 4	40.0 °C	54.0 °C	54.0 °C	35.0 °C	37.0 °C	46.0 °C

Sources for thresholds are: AT from Zhao et al. (2015), HI and Hu from Coccolo et al. (2016), T_{WB} from Sherwood and Huber (2010), T_{WBG} and T_{WBGs} from Kjellstrom et al. (2009), and UTCI from Bröde et al. (2012). For TX no absolute thresholds are defined.

for heat stress in industry and sport events (T_{WBG}, T_{WBGs}), in assessment scales of thermal comfort models (AT, UTCI), or for defining adaptability limits for heat stress (T_{WB}), as summarized in Table S1. For some HSIs, even multiple thresholds are defined for different geographic regions or work intensities (Table S2). Based on this collection, we distinguish four different levels for each HSI (see Table 2). Thresholds for Hu and UTCI stem from the publications originally introducing these indicators (Bröde et al., 2012; Masterson & Richardson, 1979). HI thresholds are adopted from the NOAA heat warning scale (https://www.wrh.noaa.gov/psr/general/safety/heat and Blazejczyk et al., 2012), which is used for issuing heat warnings in the USA. For AT, which is calculated here according to the formulation of Zhao et al. (2015), we employ the respective thresholds defined therein. For T_{WB} the only defined absolute threshold is the lethal value of 35 °C, which we adopt here (Sherwood & Huber, 2010). For T_{WBG} and T_{WBGs}, for which multiple thresholds are defined, we use a threshold from Kjellstrom et al. (2009), which describes heat impacts on workers for sustained medium level work. Additionally, the Supporting Information provides a comparison of threshold exceedances when applying different T_{WBG} and T_{WBGs} thresholds (see Section 4).

We employ impact-relevant thresholds as they allow for a global estimation of impacts related to heat stress, which would not be possible by only considering HSI trends. However, it is important to note that the application of globally uniform thresholds does not consider that the adaptation and vulnerability to heat stress varies in different regions of the world and for different population groups. Future adaption will furthermore be different across these groups and regions depending, for example, on the respective socioeconomic and demographic development. The employed impact-relevant thresholds can be interpreted as indication of heat stress severity but keeping in mind that their definition and application purpose might substantially differ across HSIs and also depend on geographic location or other contexts.

2.3. Data

The analysis is performed using 24 CMIP6 models (see Table S3) that pwarovide the necessary data for calculating HSIs (as of 30 September 2020), employing data from historical and SSP5-8.5 scenarios (O'Neill et al., 2016) and covering the time span 1981–2100. Daily HSIs are calculated using daily maximum near-surface air temperature, daily mean near-surface specific humidity, and daily mean sea level pressure. Ideally, values for all three variables would be taken at the same time of day, but standard CMIP6 output is not available for periods shorter than one day. Pressure and specific humidity (in contrast to relative humidity) do usually not vary much within one day. To avoid overestimation of heat stress, we thus use daily averages for these variables (Casanueva et al., 2019; Coffel et al., 2017). Note that a few adjustments are applied to pressure and humidity data (see Text S2 for details). The analysis is carried out over land areas (defined as grid cells with at least 50% land fraction and excluding Antarctica) and calculations are performed on the models' native grids. 29 February is removed for models that consider leap years.

Since many climate indicators scale linearly with global mean temperature (GMT; Seneviratne et al., 2016; Wartenburger et al., 2017), we express HSI trends and threshold exceedances as function of GMT. The data are linearly interpolated with GMT steps of 0.05 K before calculating multimodel statistics. Yearly GMT is calculated for the period 1981–2100 from monthly near-surface air temperature data for historical and SSP5-8.5 scenarios using 1850–1900 as reference period, and smoothed with a 20-years moving window. We also calculate GMT for SSP1-2.6, SSP2-4.5, and SSP3-7.0 scenarios to indicate the GMT change reached in 2081–2100 in each SSP. Figure S1 provides an overview of the GMT evolution for all climate models.

The intervals between threshold levels are different for each HSI (e.g., intervals for AT are 3 °C–5 °C, while intervals for UTCI are 6 °C–8 °C; see Table 2), reflecting that HSIs are defined each on their individual scale. To provide a scale-independent estimate, we normalize the HSI trends dividing them by the mean threshold interval (4.0 °C for AT, 9.0 °C for HI, 8.0 °C for Hu, 2.67 °C for T_{WBG} and T_{WBGs}, and 6.67 °C for UTCI). This approach is only applicable to the six indicators that have thresholds defined for all four levels, thus not to TX and T_{WB}. While the increase in threshold levels is approximately linear (Figure S2), the intervals

effectively vary with levels. The normalized values do not account for this variation, but rather indicate trends in terms of an average interval, which is representative for each HSI.

As reference data for bias adjustment (see Section 2.4), hourly 2 m temperature, 2 m dew point temperature, and surface pressure from the ECMWF reanalysis ERA5 (Hersbach et al., 2018, 2020) for the period 1981–2010 are used. Dew point temperature is converted into specific humidity (see Text S2). Maximum temperature is calculated as maximum, pressure and humidity as average of the hourly data. ERA5 data are interpolated to each climate model's grid by conservative remapping before calculating HSIs.

Since HSIs quantify heat effects on humans, we especially focus on densely populated regions. We use historical population data from ISIMIP2b (Frieler et al., 2017), which is based on the HYDE3.2 database (Klein Goldewijk et al., 2017) for the period 1981–2005. Population development for 2006–2100 follows the SSP5 trajectory (Jones & O'Neill, 2016; Samir & Lutz, 2017), consistent with the employed CMIP6 SSP5-8.5 scenario. Population data is temporally and spatially interpolated (see Text S3 for details). For global averages, we only consider grid boxes with more than 1000 inhabitants (see Figure S3).

Regional averages are calculated for nine domains (see Figure 1, top), selected from the regions defined by the IPCC Special Report on Extreme Events (SREX; Seneviratne et al., 2012). We use the five most populated regions in the world (SAS, EAS, and SEA in Asia, WAF and EAF in Africa) and the two most populated regions in Europe (MED and CEU) and America (CAM and ENA). For every region, area-weighted averages are calculated considering the land fraction of each grid cell.

2.4. Bias Adjustment

Although climate models incorporate our best current knowledge about the climate system, they can be prone to biases from various sources. For calculating exceedances above absolute thresholds, bias adjustment is thus required. We employ the quantile delta mapping (QDM) approach described by Cannon et al. (2015) to match the CMIP6 model distributions to the reference data set ERA5. QDM adjusts the model distribution in the application period to match the distribution of the reference data set in the reference period. Subsequently, it adds the climate change signal by considering changes between the model distribution in the reference and application periods for each quantile individually. QDM considers that future climate extremes might move beyond observed values in the reference period and reduces the risk of introducing artificial trends (Cannon et al., 2015; Maraun, 2016). To test the robustness of our results, we additionally perform bias adjustment using quantile mapping (Gudmundsson et al., 2012) and multivariate bias correction (Cannon, 2016, 2018) for comparison (see Text S4).

QDM is applied to the daily HSI distributions for each model on every land grid point separately and for each month of the year individually. The reference period is 1981–2010. We apply QDM on 50 quantiles, which is a reasonable compromise between allowing for enough flexibility and minimizing the risk of overfitting (Lakshmanan et al., 2015; Zscheischler et al., 2019). To get a smooth climate change signal, we use 30-years long application periods and select the central 10 years for our analysis, that is, data for the period 1991–2000 is obtained from the application period 1981–2010, data for 2000–2010 from the period 1991–2020, etc. Bias adjusted data cover the period 1991–2090 and are only used for calculating threshold exceedances.

3. Results

3.1. Trends in Heat Stress Indicators

The multimodel median evolution of the eight HSIs as function of GMT is shown in Figure 1 for nine selected SREX regions (an overview of the intermodel spread is shown in Figure S4). Indicated are yearly maximum values (denoted by subscript "X") relative to the 1850–1900 average. The HSIs reveal clear trends as function of GMT, although with substantially varying trend magnitude. All HSIs increase faster than GMT except $T_{\mathrm{WB},X}$ and $T_{\mathrm{WBG},X}$. TX_X, a frequently used indicator in climate studies, exhibits a moderate increase, well within the range of the other indicators and similar to the AT_X, $T_{\mathrm{WBGs},X}$, and UTCI_X trends. HI_X and Hu_X increase fastest and strongest in all regions. In the GMT range covered by all models in the period 1981–2100 (0.8 K–3.6 K, gray background), all HSIs increase mostly linearly. At very high GMT changes (beyond 5 K)

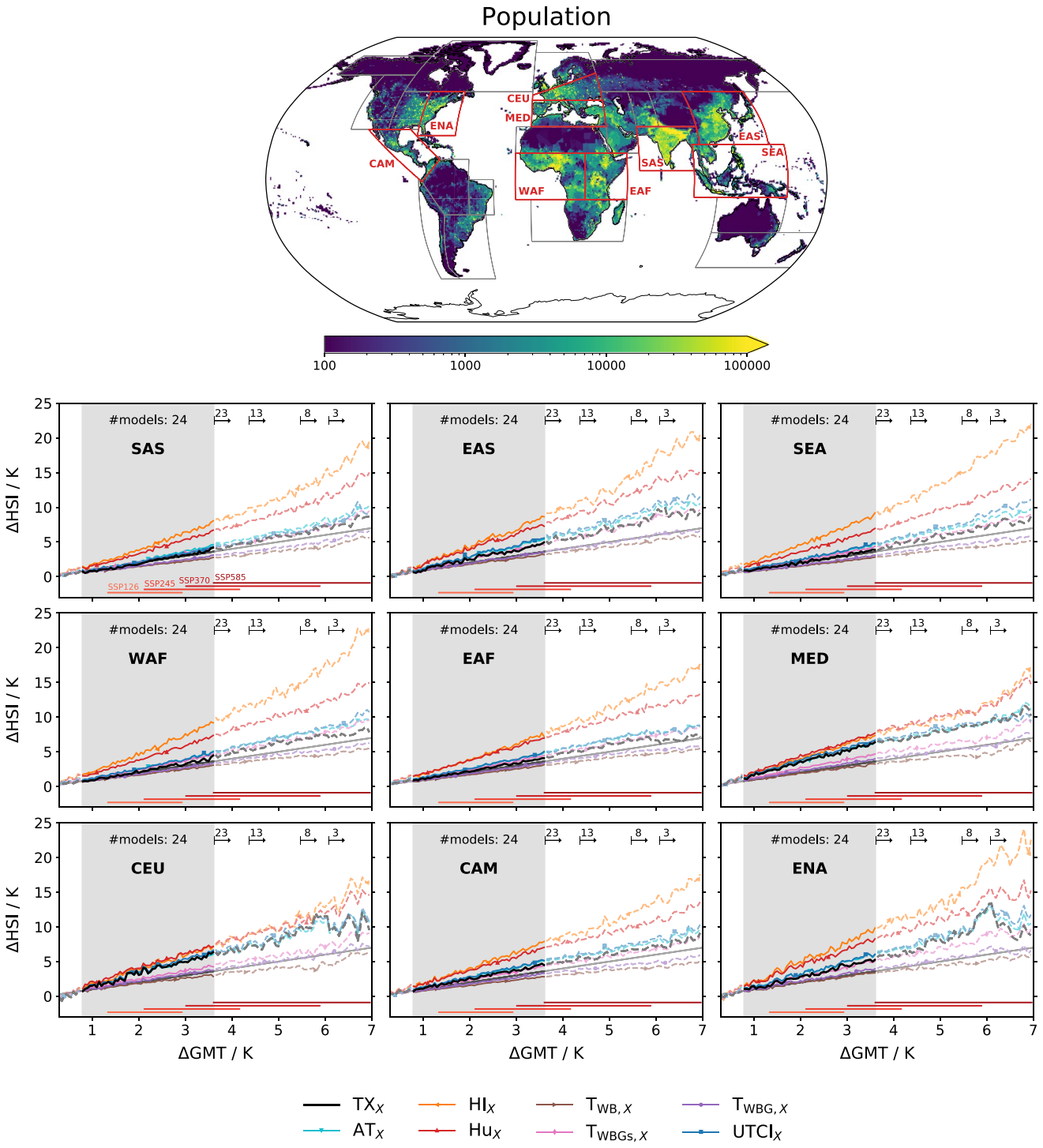

Figure 1. Global population and evolution of heat stress indicators (HSIs) as function of global mean temperature. Top: Map of global population in 2100 according to SSP5 and extent of SREX regions. Red colors highlight the nine regions that are discussed more closely in this study, selected according to their population (see Section 2.3 for details). Bottom: Multimodel median evolution of yearly maximum values (calculated as maximum of daily values and denoted by subscript "X") for eight HSIs as function of global mean temperature (GMT) in nine SREX regions. HSI and GMT changes are calculated relative to 1850–1900. Solid lines indicate the evolution of HSIs in the GMT range for which all models are available (gray background), dashed lines indicate a reduced model set. The gray solid line represents the identity line. The numbers at the top of each plot specify the total number of models and the number of models reaching the respective GMT levels (decreasing number toward higher GMT levels). The red bars at the bottom of each subplot indicate the GMT range reached by the models at the end of the 21st century (2081–2100 average) in the four Shared Socioeconomic Pathways SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5.

several HSIs show amplified trends and even sudden changes due to diverging trends in different models and the decreasing model number toward higher GMT levels (Figures S1 and S5). In several regions, models that warm up most in terms of GMT also seem to have an amplified trend in HSIs (Figure S5). HSIs increase by 2.1 K–3.6 K at ΔGMT = 2 K, 4.5 K–7.1 K at 4 K, and 6.8 K–11.6 K at 6 K (interquartile range across all regions and HSIs). When normalizing the HSI trends based on the threshold intervals (see Section 2.3), $T_{WBGs,X}$ increases fastest in all regions, followed by AT_X and $T_{WBG,X}$ (Figure S6). It is further remarkable that HI_X and Hu_X, which have the strongest trends in Figure 1, increase comparatively slowly when normalized.

The trends in HSIs as function of GMT in all 26 SREX regions are summarized in Figure 2 for absolute and normalized trends. Both reveal remarkable trend spread. When not normalized (Figure 2a), HI_X and Hu_X are the indicators with the strongest increase by far with a mean amplification of more than a factor of two compared to the GMT trend, while $T_{WB,X}$, and $T_{WBG,X}$ increase at a rate close to or even below the GMT trend. The other HSIs have 30%–60% amplified trends compared to GMT. The patterns of the indicator spread are mostly consistent across regions. The regions with the strongest trends (averaged over all HSIs) are NAS, MED, CEU, AMZ, and ENA. HSI trends are lowest in SAU and WSA. The patterns for normalized trends differ considerably (Figure 2b), with $T_{WBGs,X}$, $T_{WBG,X}$, and AT_X having the strongest trends, while HI_X, Hu_X, and $UTCI_X$ increase more slowly. This difference between absolute and normalized values highlights the strong scale-dependence of HSI trends. A comparison between SSP1-2.6 and SSP5-8.5 shows that the HSI trends are consistent across these SSPs (Figures S7 and S8), demonstrating that the increase of HSIs is in fact mostly determined by the GMT evolution. Trend magnitude and patterns are also similar when replacing daily maximum by daily minimum temperature (Figure S9).

3.2. Projected Exceedances of Impact-Relevant Thresholds

For quantifying increases in human heat stress due to climate change, the HSI exceedances beyond impact-relevant thresholds are more informative than trends. The yearly number of days exceeding the level 3 threshold for the nine selected regions are shown in Figure 3 (exceedances of the other thresholds can be found in Figures S10–S12). The average number of exceedance days (N_{exc}) is below 100 days/y in the current climate (ΔGMT ≈ 1 K) for all indicators except T_{WBGs}. In EAS, CEU, and ENA today's exceedance is virtually zero. With rising GMT levels N_{exc} is strongly increasing in several regions, particularly in SAS, SEA, WAF, EAF, and CAM (16 days/y to 86 days/y at ΔGMT = 2 K and 72 days/y to 175 days/y at 4 K; interquartile range across HSIs in these regions). N_{exc} increases more slowly in MED, EAS, and ENA, and exceedances in CEU stay very low even under high warming. The projections of N_{exc} strongly vary across HSIs. In general, T_{WBGs} and AT exhibit the most exceedance days in all regions, while T_{WBG} and UTCI exceedances are lowest. Differences across HSIs are particularly large in SAS, SEA, WAF, EAF, and CAM, which are all regions with potentially strong increases. In SEA and WAF, also the intermodel uncertainty (shaded areas around curves) is comparatively high. In contrast, all the other regions have a rather low uncertainty and show a mostly linear increase of N_{exc}. In SEA and WAF, N_{exc} follows a saturation curve for several HSIs, with level 3 exceeded almost every day of the year under very high warming (above 5 K) for all HSIs but T_{WBG}.

Figure 4 shows the fraction of populated land where the different thresholds are exceeded at least once per year, which can be interpreted as an indication for global heat stress. The spread across HSIs is considerable, but the intermodel spread is relatively low (on average contributing less than 10% to the total variance). Thresholds for level 1 and 2 are exceeded on a substantial fraction of populated land already today (38% and 21% respectively, on average), while level 3 and level 4 exceedances are presently close to zero (except for AT and T_{WBGs} at level 3). With global warming, all levels will be exceeded on a larger land fraction, with level 2 and level 3 increasing strongest (15%–17% more exceedance at ΔGMT = 4 K compared to 1 K, and 27% more at 6 K). Substantial level 4 exceedances only occur for AT and T_{WBGs} under more than 3 K global warming. When HSIs are calculated using daily minimum temperature instead of daily maximum temperature, global exceedances are much lower (Figure S13). However, particularly under high warming, level 1 and level 2 thresholds are exceeded on a considerable land fraction also for minimum temperatures.

The HSI with the highest exceedance percentage varies with threshold levels. For levels 1 and 2, most HSIs show similar exceedance percentages except for T_{WBG} which is lower. For level 3, exceedance percentages are highest for AT and T_{WBGs} and lowest for T_{WBG} with the remaining HSIs lying in between. For level 4, only AT and T_{WBGs} show notable increases, while the other HSIs stay close to zero even at very high warming

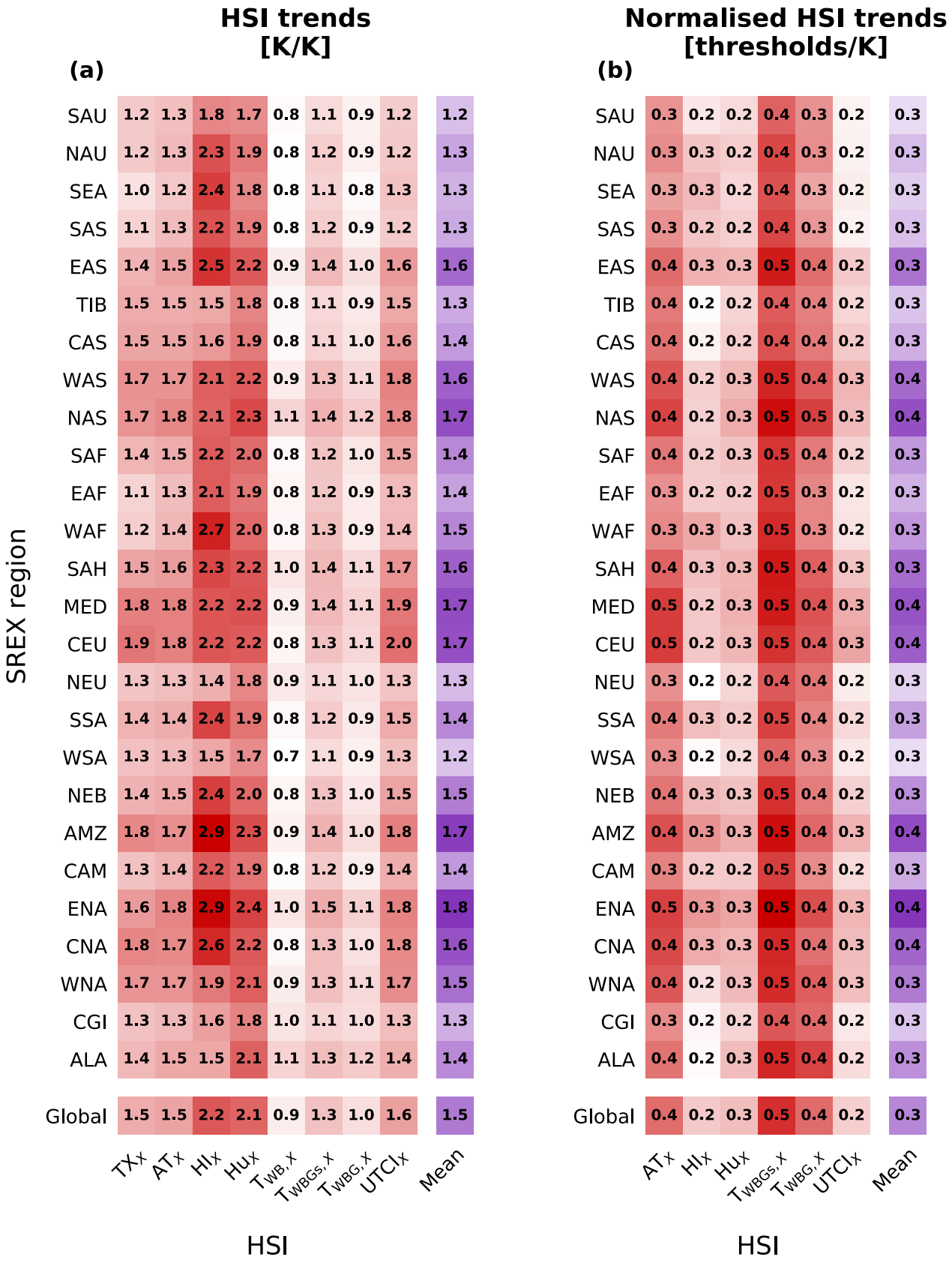
HSI trends
[K/K]
(a)
Normalised HSI trends
[thresholds/K]
(b)
SREX region
SAU 1.2 1.3 1.8 1.7 0.8 1.1 0.9 1.2 1.2
NAU 1.2 1.3 2.3 1.9 0.8 1.2 0.9 1.2 1.3
SEA 1.0 1.2 2.4 1.8 0.8 1.1 0.8 1.3 1.3
SAS 1.1 1.3 2.2 1.9 0.8 1.2 0.9 1.2 1.3
EAS 1.4 1.5 2.5 2.2 0.9 1.4 1.0 1.6 1.6
TIB 1.5 1.5 1.5 1.8 0.8 1.1 0.9 1.5 1.3
CAS 1.5 1.5 1.6 1.9 0.8 1.1 1.0 1.6 1.4
WAS 1.7 1.7 2.1 2.2 0.9 1.3 1.1 1.8 1.6
NAS 1.7 1.8 2.1 2.3 1.1 1.4 1.2 1.8 1.7
SAF 1.4 1.5 2.2 2.0 0.8 1.2 1.0 1.5 1.4
EAF 1.1 1.3 2.1 1.9 0.8 1.2 0.9 1.3 1.4
WAF 1.2 1.4 2.7 2.0 0.8 1.3 0.9 1.4 1.5
SAH 1.5 1.6 2.3 2.2 1.0 1.4 1.1 1.7 1.6
MED 1.8 1.8 2.2 2.2 0.9 1.4 1.1 1.9 1.7
CEU 1.9 1.8 2.2 2.2 0.8 1.3 1.1 2.0 1.7
NEU 1.3 1.3 1.4 1.8 0.9 1.1 1.0 1.3 1.3
SSA 1.4 1.4 2.4 1.9 0.8 1.2 0.9 1.5 1.4
WSA 1.3 1.3 1.5 1.7 0.7 1.1 0.9 1.3 1.2
NEB 1.4 1.5 2.4 2.0 0.8 1.3 1.0 1.5 1.5
AMZ 1.8 1.7 2.9 2.3 0.9 1.4 1.0 1.8 1.7
CAM 1.3 1.4 2.2 1.9 0.8 1.2 0.9 1.4 1.4
ENA 1.6 1.8 2.9 2.4 1.0 1.5 1.1 1.8 1.8
CNA 1.8 1.7 2.6 2.2 0.8 1.3 1.0 1.8 1.6
WNA 1.7 1.7 1.9 2.1 0.9 1.3 1.1 1.7 1.5
CGI 1.3 1.3 1.6 1.8 1.0 1.1 1.0 1.3 1.3
ALA 1.4 1.5 1.5 2.1 1.1 1.3 1.2 1.4 1.4
Global 1.5 1.5 2.2 2.1 0.9 1.3 1.0 1.6 1.5
TXx ATx HIx Hux TWB,x TWBGs,x TWBG,x UTCIx Mean
HSI
SAU 0.3 0.2 0.2 0.4 0.3 0.2 0.3
NAU 0.3 0.3 0.2 0.4 0.3 0.2 0.3
SEA 0.3 0.3 0.2 0.4 0.3 0.2 0.3
SAS 0.3 0.2 0.2 0.4 0.3 0.2 0.3
EAS 0.4 0.3 0.3 0.5 0.4 0.2 0.3
TIB 0.4 0.2 0.2 0.4 0.4 0.2 0.3
CAS 0.4 0.2 0.2 0.4 0.4 0.2 0.3
WAS 0.4 0.2 0.3 0.5 0.4 0.3 0.4
NAS 0.4 0.2 0.3 0.5 0.5 0.3 0.4
SAF 0.4 0.2 0.3 0.5 0.4 0.2 0.3
EAF 0.3 0.2 0.2 0.5 0.3 0.2 0.3
WAF 0.3 0.3 0.3 0.5 0.3 0.2 0.3
SAH 0.4 0.3 0.3 0.5 0.4 0.3 0.3
MED 0.5 0.2 0.3 0.5 0.4 0.3 0.4
CEU 0.5 0.2 0.3 0.5 0.4 0.3 0.4
NEU 0.3 0.2 0.2 0.4 0.4 0.2 0.3
SSA 0.4 0.3 0.2 0.5 0.4 0.2 0.3
WSA 0.3 0.2 0.2 0.4 0.3 0.2 0.3
NEB 0.4 0.3 0.3 0.5 0.4 0.2 0.3
AMZ 0.4 0.3 0.3 0.5 0.4 0.3 0.4
CAM 0.3 0.2 0.2 0.5 0.3 0.2 0.3
ENA 0.5 0.3 0.3 0.5 0.4 0.3 0.4
CNA 0.4 0.3 0.3 0.5 0.4 0.3 0.4
WNA 0.4 0.2 0.3 0.5 0.4 0.3 0.3
CGI 0.3 0.2 0.2 0.4 0.4 0.2 0.3
ALA 0.4 0.2 0.3 0.5 0.4 0.2 0.3
Global 0.4 0.2 0.3 0.5 0.4 0.2 0.3
ATx HIx Hux TWBGs,x TWBG,x UTCIx Mean
HSI

Figure 2.

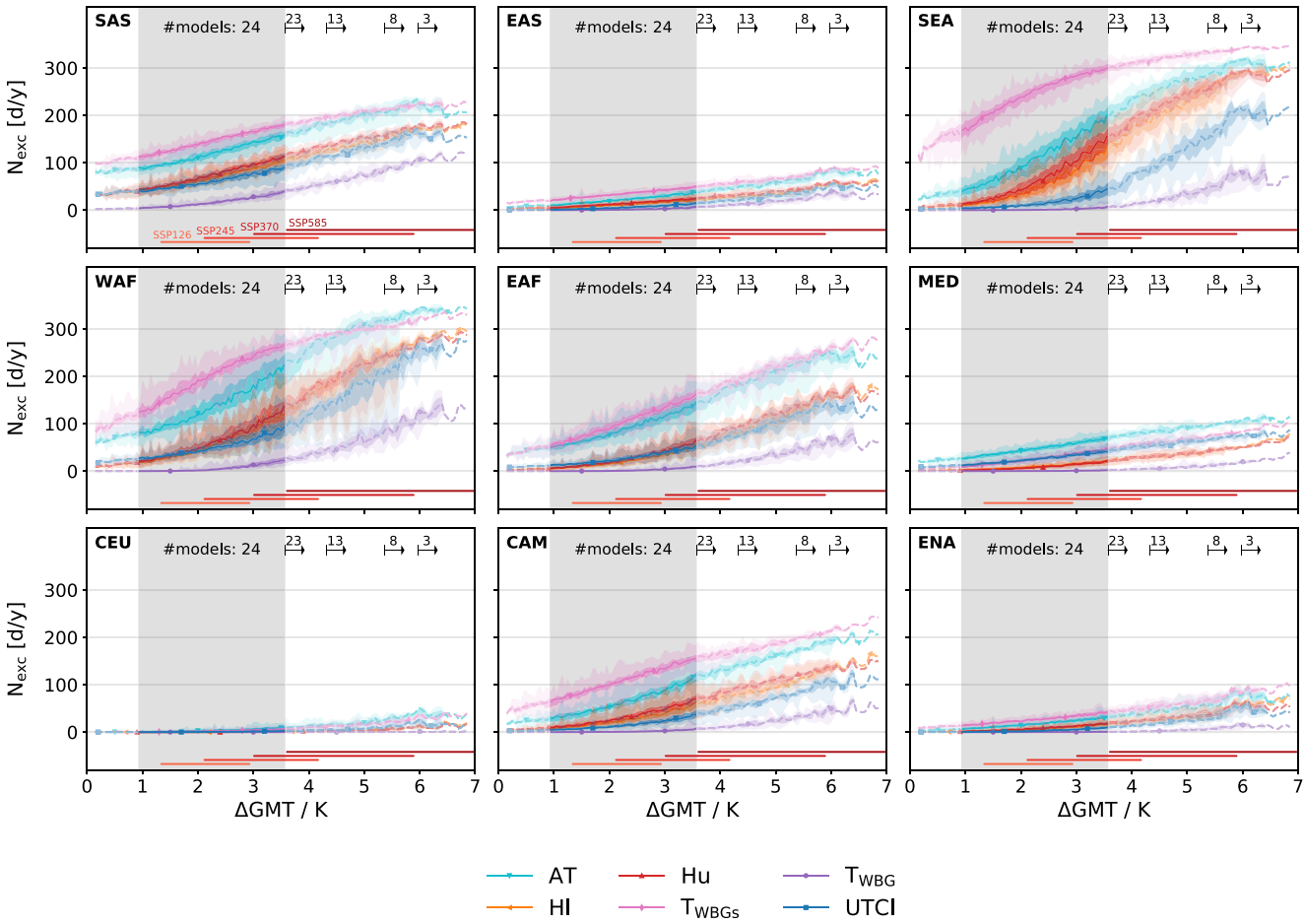

Figure 3. As in Figure 1 but for the number of days per year (N_{exc}) on which the level 3 threshold of each heat stress indicator is exceeded. Lines indicate the multimodel median, shading in dark colors indicates the multimodel interquartile range, and shading in light colors the total model range. T_{WB} and TX are omitted because no level 3 threshold is defined for T_{WB}, and for TX no absolute thresholds are defined at all. Bias adjusted data are used for calculating the threshold exceedance. Note that the models HadGEM3-GC31-LL, HadGEM3-GC31-MM, KACE-1-0-G, and UKESM1-0-LL use a 360-days calendar and values are thus multiplied by 365/360 to fit the other models.

levels. It is also noteworthy that exceedance percentages of T_{WBGs} and T_{WBG} are substantially different, given that the same thresholds are applied to them.

A geographically explicit representation of the yearly number of exceedance days at 3 K GMT increase for three representative HSIs (i.e., representing the HSI spread) is shown in Figure 5. The maps reveal a clear latitudinal gradient with high exceedance numbers at low latitudes and decreasing values toward the poles. Regions close to the equator are thus particularly in danger of exceeding thresholds during a substantial number of days each year under 3 K warming. In the Amazon region, central Africa, India, and southeast Asia level 1 (and partly level 2) thresholds of UTCI and AT would be exceeded almost every day of the year. Yet, differences between the selected HSIs are apparent: AT generally shows the highest exceedance numbers, T_{WBG} has the fewest, and UTCI lies in between (in agreement with Figures 3 and 4). This is particularly striking for the level 3 threshold, which is exceeded in many low latitude regions during more than half of

Figure 2. Multimodel median trends in yearly maximum values of the heat stress indicators (HSIs) in all 26 SREX regions and on the global land area. Trends in (a) absolute HSI values and in (b) HSI values normalized by average threshold intervals (only applicable to six HSIs, see Section 2.3 for details). Numbers indicate the trends in HSIs per change in global mean temperature and color strength represents the trend magnitude. The rightmost column (colored in violet) shows the mean trend across all HSIs. Linear regression is used to estimate trends for each model individually by pairing yearly GMT and yearly maximum HSI values for the GMT range 0.8 K–3.6 K (for which all models have data) and, subsequently, the multimodel median is calculated. Bold weighted fonts indicate that 90% of the models have a significant trend at the 1% level (estimated by a Wald test with t-distribution of the test statistic).

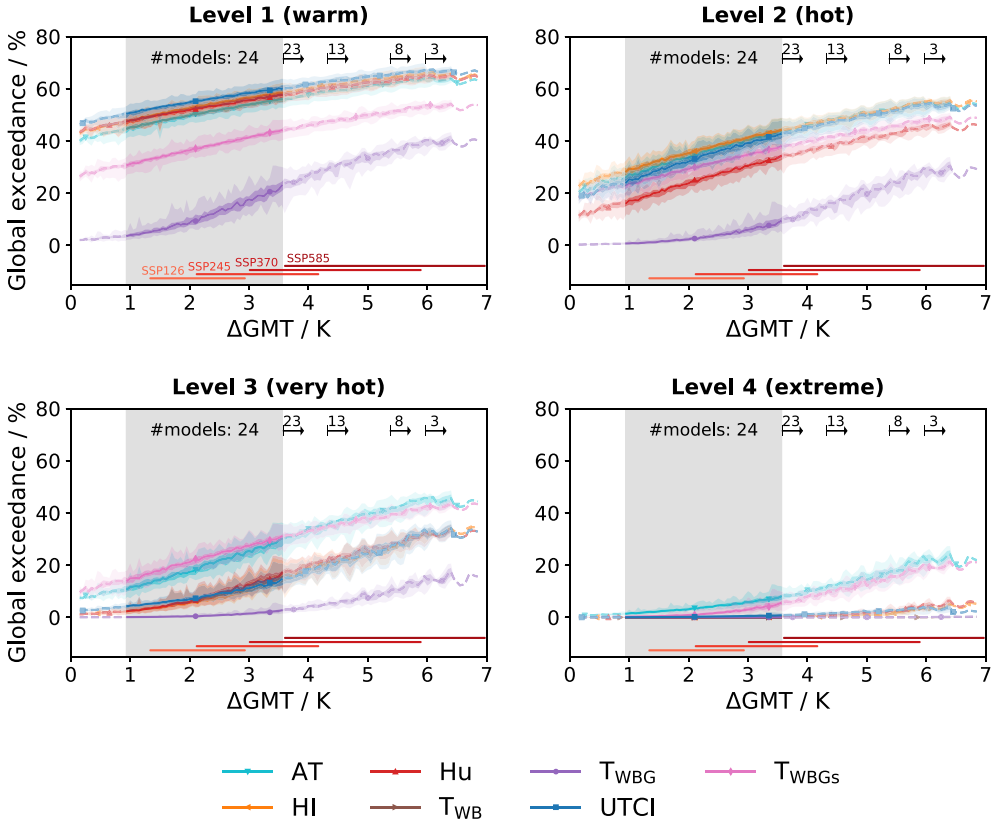

Figure 4. As in Figure 3 but for global area-weighted percentage of populated land that exceeds a threshold at least once per year. Only grid boxes with at least 1000 inhabitants are considered (see Section 2.3).

the year for AT, while UTCI, and especially T_{WBG} show much fewer exceedance days. The level 4 threshold is substantially exceeded only by AT, while exceedance days are low for UTCI and zero for T_{WBG}.

4. Discussion

Future trends in HSIs vary substantially across indicators and depend strongly on the scales on which the HSIs are defined. The varying definitions (Table 1) make a direct comparison of HSI trends challenging and the indicators with strongest (or smallest) trends can change substantially after normalization (Figure 2). Comparing future changes of different HSIs (e.g., Li et al., 2018; Matthews et al., 2017; Wang & Zhu, 2019) is thus only meaningful if differences and scale issues are considered and discussed. Importantly, the HSIs with the strongest trends in absolute terms (Figures 1 and 2a) do not necessarily correspond to the ones that show the highest threshold exceedances, particularly for the high levels 3 and 4 (Figure 4). These levels are most often exceeded by AT and T_{WBGs}, while HI and Hu despite having the strongest trends show comparatively low exceedance rates. After normalization, AT and T_{WBGs} are among the indicators with the strongest trends (Figure 2b) which corresponds better to the high threshold exceedances of these indicators. In contrast, T_{WBG} has low exceedance rates despite a strong normalized trend. The HSIs with the highest exceedances also vary across threshold levels, suggesting that the combination of both the trend magnitude and the threshold levels causes the high variability in projected HSI exceedances. This is of particular relevance, as different indicators are typically used depending on the impact under consideration. For instance, T_{WBG} and T_{WBGs} are widely applied in the context of occupational health (Budd, 2008; Garzon-Villalba et al., 2016) and worker productivity (Dunne et al., 2013; Kjellstrom et al., 2013; Lemke & Kjellstrom, 2012) while HI or Hu are used by meteorological services for issuing heat warnings. Recent developments, such as UTCI, are more sophisticated but still lack certain heat transfer aspects (e.g., clothing, Spector & Sheffield, 2014). While evidence about the advantage of using HSIs instead of simpler metrics such as minimum, mean, or maximum air temperature to predict heat-related mortality and morbidity

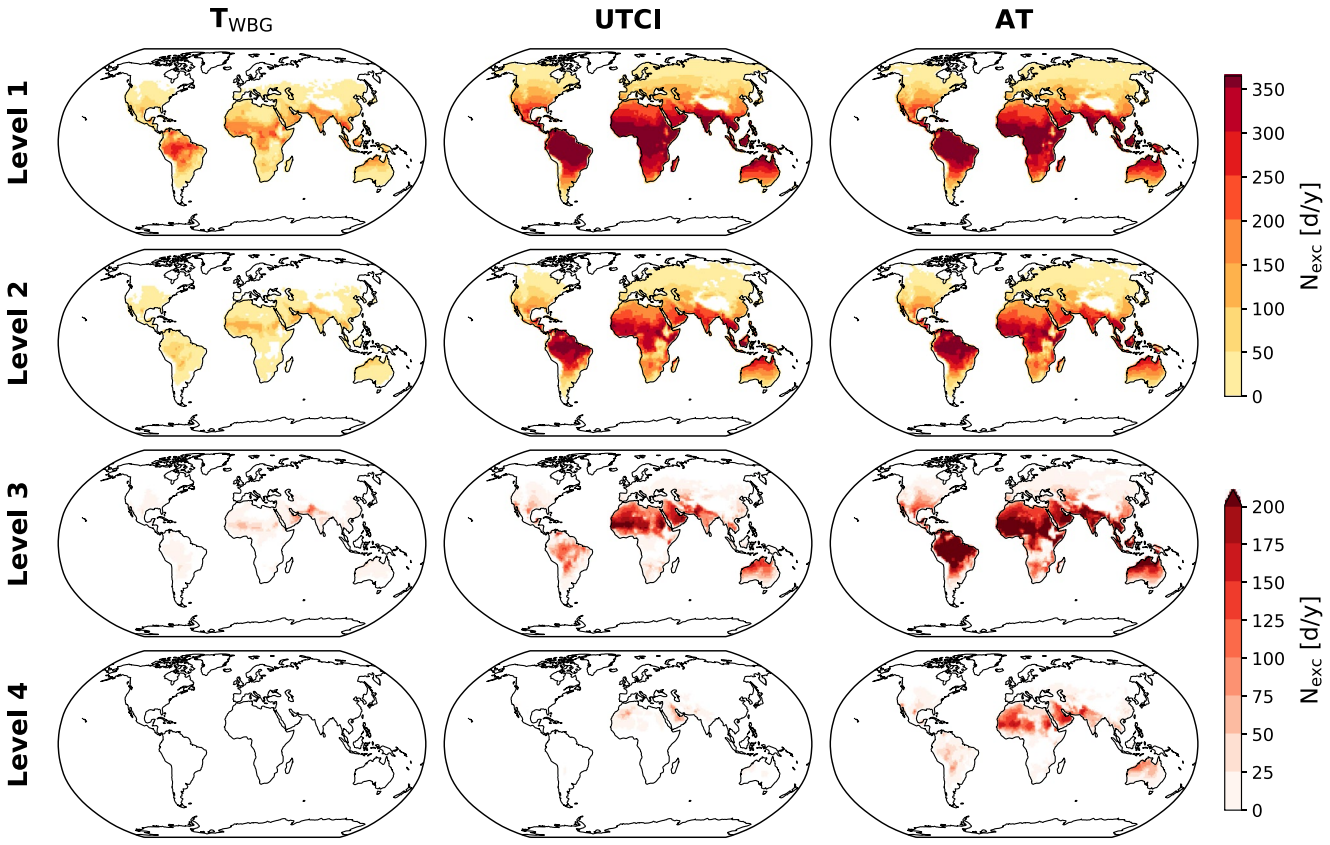

Figure 5. Multimodel median yearly exceedance (N_{exc}) of impact-relevant thresholds for the heat stress indicators T_{WBG}, UTCI, and AT at 3 K GMT increase (relative to 1850–1900). Note the different color maps and color bar limits for levels 1 and 2 and levels 3 and 4. Climate models are interpolated by conservative remapping to 1° resolution before calculating the multimodel median.

is still inconclusive (Barnett et al., 2010; Burkart et al., 2011; Heo & Bell, 2018; Kim et al., 2011; Ragettli et al., 2017; Vaneckova et al., 2011), the divergent future HSI trends and exceedances of impact-relevant thresholds call for a more thorough evaluation of HSIs at regional-to-global scales.

Many climate impact and epidemiological studies employ T_{WBGs} as an approximation for T_{WBG} (e.g., Blazejczyk et al., 2012; Matthews et al., 2017; Urban et al., 2019; Vaneckova et al., 2011; Zhao et al., 2015). According to our results, trends and threshold exceedances of T_{WBGs} and T_{WBG} differ substantially (see also Figure S14), providing a very strong indication that T_{WBGs} cannot be used as a simple replacement of T_{WBG}. Lemke and Kjellstrom (2012) tracked the source of T_{WBGs} and highlighted a potential inconsistency between the original citation and the currently used T_{WBGs} equation. Their recommendation to not use T_{WBGs} for replacing T_{WBG} is strongly supported by our results. Future research should thus apply T_{WBG} consistent with its definition (see Table 1; ideally also including mean radiant temperature to account for radiation) and not replace it by the simplified version T_{WBGs}.

Acclimatization to heat depends on the geographic region (Kalkstein & Davis, 1989; Xiao et al., 2014), and using globally uniform thresholds is thus a strong simplification. Additionally, human sensitivity to heat can depend on various other factors such as workload, age, or socioeconomic status (McMichael et al., 2006; Kjellstrom et al., 2009). Consequently, different threshold levels have been defined to integrate these factors at least partially (e.g., Grundstein et al., 2015; Kjellstrom et al., 2009). When applying different thresholds to one indicator, the exceedance rates can change considerably, as shown in Figure S15 for nine different T_{WBG} thresholds. Particularly for levels 1 and 2, the exceedance rates vary substantially when using different thresholds. For levels 3 and 4 the spread is not as large, but still exceedances of some thresholds are low even under high warming while other thresholds are exceeded frequently. Thresholds are generally much less exceeded when using daily minimum temperatures (Figure S13), indicating that the applied thresholds seem

to be more relevant for daytime conditions, while during night time lower thresholds and the persistence of warm nights need to be considered (e.g., Laaidi et al., 2012).

In this study, we do not consider that humans have the potential to adapt to increasing heat stress, for example, by improving house insulation, installing air conditioning, or changing behavior (e.g., adjust working and rest times, avoid direct sunlight, adjust clothing, visit cooling centers if available). Although HSIs scale well with GMT (Figure 1), choosing only the SSP5 scenario for the analysis neglects that different socioeconomic developments affect the ability of people to adapt to increasing heat exposure (Russo et al., 2019). The calculated heat stress for ambient conditions might thus overestimate the real heat stress experienced by humans, also because people use to spend much time indoors. Yet, especially in low-income countries, which belong to the regions most affected by increasing heat stress, housing quality is often poor and indoor conditions can be assumed to be similar to the outdoor state (Lemke & Kjellstrom, 2012; Spector & Sheffield, 2014). In certain regions, the HSIs we employ (representing conditions in the shade) might thus be a reasonable approximation even for indoor conditions. For estimating realistic heat stress for outdoor activities, variables such as solar radiation or wind speed would be important to consider, which we, however, did not include in our analysis. With regard to the threshold levels, adaptation strategies could also alter the levels at which critical human reactions to heat stress emerge (e.g., Bobb et al., 2014; Gasparrini et al., 2015).

The maximum ΔGMT reached in SSP5-8.5 varies greatly across models (from 3.6 K to 7.0 K, see Figure S1). Only few models reach very high ΔGMT values, making projections for these high warming levels rather uncertain. Yet, as long as these very-high warming scenarios cannot be physically excluded they should be considered (Forster et al., 2019) since they could potentially have dramatic impacts on human health due to the nonlinear impacts caused by increasing heat stress (Dunne et al., 2013; Gasparrini et al., 2015; Matthews et al., 2017).

5. Conclusions

Heat stress indicators (HSIs) are used in a wide range of fields, such as heat warning systems (Blazejczyk et al., 2012), occupational health and worker productivity assessments (Garzon-Villalba et al., 2016; Kjellstrom et al., 2013; Orlov et al., 2019), heat safety estimates of sports events (Racinais et al., 2015; Thorsson et al., 2020), and epidemiological studies (e.g., Gronlund et al., 2014; Heo & Bell, 2018; Morabito et al., 2014; Ragettli et al., 2017; Vaneckova et al., 2011). In the present study we calculated eight HSIs (TX, AT, HI, Hu, T_{WB}, T_{WBG}, T_{WBGs}, UTCI) as function of global mean temperature for future climate projections from CMIP6 models. All HSIs show statistically significant future trends in all regions. The trends as well as the exceedances of impact-relevant thresholds show substantial spread across indicators, highlighting the need for choosing appropriate HSIs according to the respective context of interest (e.g., for meteorological heat warnings, occupational health, or worker productivity). Due to this context dependency, thresholds for the same level might not be directly comparable across different HSIs, which likely contributes to the indicator spread and partly explains that the HSI with the highest exceedance varies for different threshold levels. The indicator dependency and the large uncertainty of future heat stress estimations also make the choice of HSIs a critical task. In line with findings showing that no HSI can be identified to be superior than others for quantifying epidemiological impacts (e.g., Barnett et al., 2010; Burkart et al., 2011; Heo & Bell, 2018; Kim et al., 2011; Ragettli et al., 2017; Vaneckova et al., 2011), results of the present study do not allow for recommendations on the usage of any particular HSI – except that our results confirm the findings of Lemke and Kjellstrom (2012) that the usage of T_{WBGs} should be avoided. Trends in TX_X, which is a widely used heat indicator in climate science (e.g., Shiogama et al., 2019; Sillmann et al., 2013), are in line with trends in the other HSIs. As recent findings even question the importance of humidity as contributing factor to higher mortality (Armstrong et al., 2019), TX_X, a comparatively simple indicator, might indeed provide reliable estimates of future heat stress trends. As no thresholds are defined for TX_X, the usage of other HSIs is nevertheless necessary for estimating human-relevant threshold exceedances. The large variability of threshold exceedances identified in this study highlights the need for further research to better connect HSIs to observed impacts on regional-to-global scales, which calls for close collaboration between climate science and health impact communities.

Data Availability Statement

This study contains modified Copernicus Climate Change Service Information 2019. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus Information or Data it contains. UTCI code was adapted from the comfort_tool script of the Center for the Built Environment (https://github.com/CenterForTheBuiltEnvironment). The computations were performed using the NS9188K project account and data was stored and shared on project account NS9252K on resources provided by UNINETT Sigma2 — the National Infrastructure for High Performance Computing and Data Storage in Norway. This publication contains Supporting Information, which includes additionally the following references: Alduchov and Eskridge (1996); ECMWF (2016); Fischereit and Schlünzen (2018); d'Ambrosio Alfano et al. (2011). The data that support the findings of this study are openly available from the NIRD Research Data Archive under doi:10.11582/2021.00011. CMIP6 data are available from https://esgf-node.llnl.gov and ERA5 data from cds.climate.copernicus.eu. Population data can be downloaded from https://esg.pik-potsdam.de/projects/isimip.

Acknowledgments
We acknowledge the World Climate Research Program, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 820655 (EXHAUSTION) and from the Belmont Forum Collaborative Research Action on Climate, Environment, and Health, supported by the Norwegian Research Council (contract No 310672, HEATCOST).

References

Alduchov, O. A., & Eskridge, R. E. (1996). Improved Magnus form approximation of saturation vapor pressure. *Journal of Applied Meteorology*, *35*(4), 601–609. https://doi.org/10.1175/1520-0450(1996)035⟨0601:IMFAOS⟩2.0.CO;2

American College of Sports Medicine. (1984). Prevention of thermal injuries during distance running – position stand. *Medical Journal of Australia*, *141*(12–13), 876–879. https://doi.org/10.5694/j.1326-5377.1984.tb132981.x

Anderson, G. B., Bell, M. L., & Peng, R. D. (2013). Methods to calculate the heat index as an exposure metric in environmental health research. *Environmental Health Perspectives*, *121*(10), 1111–1119. https://doi.org/10.1289/ehp.1206273

Armstrong, B., Sera, F., Vicedo-Cabrera, A. M., Abrutzky, R., Åström, D. O., Bell, M. L., et al. (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

Barnett, A., Tong, S., & Clements, A. (2010). What measure of temperature is the best predictor of mortality? *Environmental Research*, *110*(6), 604–611. https://doi.org/10.1016/j.envres.2010.05.006

Blazejczyk, K., Epstein, Y., Jendritzky, G., Staiger, H., & Tinz, B. (2012). Comparison of UTCI to selected thermal indices. *International Journal of Biometeorology*, *56*(3), 515–535. https://doi.org/10.1007/s00484-011-0453-2

Bobb, J. F., Peng, R. D., Bell, M. L., & Dominici, F. (2014). Heat-related mortality and adaptation to heat in the United States. *Environmental Health Perspectives*, *122*(8), 811–816. https://doi.org/10.1289/ehp.1307392

Bolton, D. (1980). The computation of equivalent potential temperature. *Monthly Weather Review*, *108*(7), 1046–1053. https://doi.org/10.1175/1520-0493(1980)108⟨1046:TCOEPT⟩2.0.CO;2

Bröde, P., Fiala, D., Błażejczyk, K., Holmér, I., Jendritzky, G., Kampmann, B., et al., (2012). Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). *International Journal of Biometeorology*, *56*, (3), 481–494. https://doi.org/10.1007/s00484-011-0454-1

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

Burkart, K., Schneider, A., Breitner, S., Khan, M. H., Krämer, A., & Endlicher, W. (2011). The effect of atmospheric thermal conditions and urban thermal pollution on all-cause and cardiovascular mortality in Bangladesh. *Environmental Pollution*, *159*(8–9), 2035–2043. https://doi.org/10.1016/j.envpol.2011.02.005

Cannon, A. J. (2016). Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure. *Journal of Climate*, *29*(19), 7045–7064. https://doi.org/10.1175/JCLI-D-15-0679.1

Cannon, A. J. (2018). Multivariate quantile mapping bias correction: An N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics*, *50*(1), 31–49. https://doi.org/10.1007/s00382-017-3580-6

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *Journal of Climate*, *28*(17), 6938–6959. https://doi.org/10.1175/JCLI-D-14-00754.1

Casanueva, A., Kotlarski, S., Herrera, S., Fischer, A. M., Kjellstrom, T., & Schwierz, C. (2019). Climate projections of a multivariate heat stress index: The role of downscaling and bias correction. *Geoscientific Model Development*, *12*(8), 3419–3438. https://doi.org/10.5194/gmd-12-3419-2019

Chen, K., Vicedo-Cabrera, A. M., & Dubrow, R. (2020). Projections of ambient temperature- and air pollution-related mortality burden under combined climate change and population aging scenarios: A review. *Current Environmental Health Reports*, *7*(3), 243–255. https://doi.org/10.1007/s40572-020-00281-6

Cheng, Y., Niu, J., & Gao, N. (2012). Thermal comfort models: A review and numerical investigation. *Building and Environment*, *47*, 13–22. https://doi.org/10.1016/j.buildenv.2011.05.011

Chung, J.-Y., Honda, Y., Hong, Y.-C., Pan, X.-C., Guo, Y.-L., & Kim, H. (2009). Ambient temperature and mortality: An international study in four capital cities of East Asia. *Science of The Total Environment*, *408*(2), 390–396. https://doi.org/10.1016/j.scitotenv.2009.09.009

Coccolo, S., Kämpf, J., Scartezzini, J.-L., & Pearlmutter, D. (2016). Outdoor human comfort and thermal stress: A comprehensive review on models and standards. *Urban Climate*, *18*, 33–57. https://doi.org/10.1016/j.uclim.2016.08.004

Coffel, E. D., Horton, R. M., & de Sherbinin, A. (2017). Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century. *Environmental Research Letters*, *13*(1), 014001. https://doi.org/10.1088/1748-9326/aaa00e

d'Ambrosio Alfano, F. R., Palella, B. I., & Riccio, G. (2011). Thermal environment assessment reliability using temperature–humidity indices. *Industrial Health*, *49*(1), 95–106. https://doi.org/10.2486/indhealth.ms1097

Davies-Jones, R. (2008). An efficient and accurate method for computing the wet-bulb temperature along pseudoadiabats. *Monthly Weather Review*, *136*(7), 2764–2785. https://doi.org/10.1175/2007MWR2224.1

de Freitas, C. R., & Grigorieva, E. A. (2017). A comparison and appraisal of a comprehensive range of human thermal climate indices. *International Journal of Biometeorology*, *61*(3), 487–512. https://doi.org/10.1007/s00484-016-1228-6

Diffenbaugh, N. S., Pal, J. S., Giorgi, F., & Gao, X. (2007). Heat stress intensification in the Mediterranean climate change hotspot. *Geophysical Research Letters*, *34*(11). https://doi.org/10.1029/2007GL030000

Dunne, J. P., Stouffer, R. J., & John, J. G. (2013). Reductions in labor capacity from heat stress under climate warming. *Nature Climate Change*, *3*(6), 563. https://doi.org/10.1038/nclimate1827

ECMWF. (2016). Part IV: Physical processes. In *Clouds and large-scale precipitation*, IFS Documentation (Vol. Cy41r2, pp. 93–114). Reading: ECMWF. Retrieved from https://www.ecmwf.int/node/16648

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., et al. (2016). Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fiala, D., Havenith, G., Bröde, P., Kampmann, B., & Jendritzky, G. (2012). UTCI-Fiala multi-node model of human heat transfer and temperature regulation. *International Journal of Biometeorology*, *56*(3), 429–441. https://doi.org/10.1007/s00484-011-0424-7

Fischer, E. M., & Knutti, R. (2013). Robust projections of combined humidity and temperature extremes. *Nature Climate Change*, *3*(2), 126. https://doi.org/10.1038/nclimate1682

Fischer, E. M., & Schär, C. (2010). Consistent geographical patterns of changes in high-impact European heatwaves. *Nature Geoscience*, *3*(6), 398. https://doi.org/10.1038/ngeo866

Fischereit, J., & Schlünzen, K. H. (2018). Evaluation of thermal indices for their applicability in obstacle-resolving meteorology models. *International Journal of Biometeorology*, *62*(10), 1887–1900. https://doi.org/10.1007/s00484-018-1591-6

Forster, P., Maycock, A., McKenna, C., & Smith, C. (2019). Latest climate models confirm need for urgent mitigation. *Nature Climate Change*. https://doi.org/10.1038/s41558-019-0660-0

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., et al. (2017). Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development*, *10*(12), 4321–4345. https://doi.org/10.5194/gmd-10-4321-2017

Garzon-Villalba, X. P., Mbah, A., Wu, Y., Hiles, M., Moore, H., Schwartz, S. W., et al. (2016). Exertional heat illness and acute injury related to ambient wet bulb globe temperature. *American Journal of Industrial Medicine*, *59*(12), 1169–1176. https://doi.org/10.1002/ajim.22650

Gasparrini, A., Guo, Y., Hashizume, M., Kinney, P. L., Petkova, E. P., Lavigne, E., et al. (2015). Temporal variation in heat–mortality associations: A multicountry study. *Environmental Health Perspectives*, *123*(11), 1200–1207. https://doi.org/10.1289/ehp.1409070

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., et al. (2015). Mortality risk attributable to high and low ambient temperature: A multicountry observational study. *The Lancet*, *386*(9991), 369–375. https://doi.org/10.1016/S0140-6736(14)62114-0

Gasparrini, A., Guo, Y., Sera, F., Vicedo-Cabrera, A. M., Huber, V., Tong, S., et al. (2017). Projections of temperature-related excess mortality under climate change scenarios. *The Lancet Planetary Health*, *1*(9), e360–e367. https://doi.org/10.1016/S2542-5196(17)30156-0

Gronlund, C. J., Zanobetti, A., Schwartz, J. D., Wellenius, G. A., & O'Neill, M. S. (2014). Heat, heat waves, and hospital admissions among the elderly in the United States, 1992–2006. *Environmental Health Perspectives*, *122*(11), 1187–1192. https://doi.org/10.1289/ehp.1206132

Grundstein, A., Williams, C., Phan, M., & Cooper, E. (2015). Regional heat safety thresholds for athletics in the contiguous United States. *Applied Geography*, *56*, 55–60. https://doi.org/10.1016/j.apgeog.2014.10.014

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Technical note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. *Hydrology and Earth System Sciences*, *16*(9), 3383–3390. https://doi.org/10.5194/hess-16-3383-2012

Heo, S., & Bell, M. L. (2018). Heat waves in South Korea: Differences of heat wave characteristics by thermal indices. *Journal of Exposure Science and Environmental Epidemiology*, *29*(6), 790–805. https://doi.org/10.1038/s41370-018-0076-3

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018). ERA5 hourly data on single levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.adbb2d47

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*, 1999–2049. https://doi.org/10.1002/qj.3803

Heo, S., Bell, M. L., & Lee, J.-T. (2019). Comparison of health risks by heat wave definition: Applicability of wet-bulb globe temperature for heat wave criteria. *Environmental Research*, *168*, 158–170. https://doi.org/10.1016/j.envres.2018.09.032

Höppe, P. (1999). The physiological equivalent temperature - A universal index for the biometeorological assessment of the thermal environment. *International Journal of Biometeorology*, *43*(2), 71–75. https://doi.org/10.1007/s004840050118

Huang, C., Barnett, A. G., Wang, X., Vaneckova, P., FitzGerald, G., & Tong, S. (2011). Projecting future heat-related mortality under climate change scenarios: A systematic review. *Environmental Health Perspectives*, *119*(12), 1681–1690. https://doi.org/10.1289/ehp.1103456

Im, E.-S., Pal, J. S., & Eltahir, E. A. B. (2017). Deadly heat waves projected in the densely populated agricultural regions of South Asia. *Science Advances*, *3*(8). https://doi.org/10.1126/sciadv.1603322

Jones, B., & O'Neill, B. C. (2016). Spatially explicit global population scenarios consistent with the shared socioeconomic pathways. *Environmental Research Letters*, *11*(8), 084003. https://doi.org/10.1088/1748-9326/11/8/084003

Kalkstein, L. S., & Davis, R. E. (1989). Weather and human mortality: An evaluation of demographic and interregional responses in the United States. *Annals of the Association of American Geographers*, *79*(1), 44–64. https://doi.org/10.1111/j.1467-8306.1989.tb00249.x

Kang, M., Kim, K. R., & Shin, J.-Y. (2020). Event-based heat-related risk assessment model for South Korea using maximum perceived temperature, wet-bulb globe temperature, and air temperature data. *International Journal of Environmental Research and Public Health*, *17*(8), 2631. https://doi.org/10.3390/ijerph17082631

Kent, S. T., McClure, L. A., Zaitchik, B. F., Smith, T. T., & Gohlke, J. M. (2014). Heat waves and health outcomes in Alabama (USA): The importance of heat wave definition. *Environmental Health Perspectives*, *122*(2), 151–158. https://doi.org/10.1289/ehp.1307262

Kim, Y.-M., Kim, S., Cheong, H.-K., & Kim, E.-H. (2011). Comparison of temperature indexes for the impact assessment of heat stress on heat-related mortality. *Environmental Health and Toxicology*, *26*. https://doi.org/10.5620/eht.2011.26.e2011009

Kjellstrom, T., Holmer, I., & Lemke, B. (2009). Workplace heat stress, health and productivity – An increasing challenge for low and middle-income countries during climate change. *Global Health Action*, *2*(1), 2047. https://doi.org/10.3402/gha.v2i0.2047

Kjellstrom, T., Kovats, R. S., Lloyd, S. J., Holt, T., & Tol, R. S. J. (2009). The direct impact of climate change on regional labor productivity. *Archives of Environmental & Occupational Health*, *64*(4), 217–227. https://doi.org/10.1080/19338240903352776

Kjellstrom, T., Lemke, B., & Otto, M. (2013). Mapping occupational heat exposure and effects in South-East Asia: Ongoing time trends 1980–2011 and future estimates to 2050. *Industrial Health*, *51*(1), 56–67. https://doi.org/10.2486/indhealth.2012-0174

Klein Goldewijk, K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene – HYDE 3.2. *Earth System Science Data*, *9*(2), 927–953. https://doi.org/10.5194/essd-9-927-2017

Knutson, T. R., & Ploshay, J. J. (2016). Detection of anthropogenic influence on a summertime heat stress index. *Climatic Change*, *138*(1), 25–39. https://doi.org/10.1007/s10584-016-1708-z

Laaidi, K., Zeghnoun, A., Dousset, B., Bretin, P., Vandentorren, S., Giraudet, E., et al. (2012). The impact of heat islands on mortality in Paris during the August 2003 heat wave. *Environmental Health Perspectives*, *120*(2), 254–259. https://doi.org/10.1289/ehp.1103532

Lakshmanan, V., Gilleland, E., McGovern, A., & Tingley, M. (2015). Machine learning and data mining approaches to climate science. In *Proceedings of The 4th International Workshop on Climate Informatics*. https://doi.org/10.1007/978-3-319-17220-0

Lemke, B., & Kjellstrom, T. (2012). Calculating workplace WBGT from meteorological data: A tool for climate change assessment. *Industrial Health*, *50*(4), 267–278. https://doi.org/10.2486/indhealth.MS1352

Li, C., Sun, Y., Zwiers, F., Wang, D., Zhang, X., Chen, G., et al. (2020). Rapid warming in summer wet bulb globe temperature in China with human-induced climate change. *Journal of Climate*, *33*(13), 5697–5711. https://doi.org/10.1175/JCLI-D-19-0492.1

Li, D., Yuan, J., & Kopp, R. E. (2020). Escalating global exposure to compound heat-humidity extremes with warming. *Environmental Research Letters*, *15*(6), 064003. https://doi.org/10.1088/1748-9326/ab7d04

Li, J., Chen, Y. D., Gan, T. Y., & Lau, N.-C. (2018). Elevated increases in human-perceived temperature under climate warming. *Nature Climate Change*, *8*(1), 43–47. https://doi.org/10.1038/s41558-017-0036-2

Li, T., Horton, R. M., Bader, D. A., Zhou, M., Liang, X., Ban, J., et al. (2016). Aging will amplify the heat-related mortality risk under a changing climate: Projection for the elderly in Beijing, China. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep28161

Lin, Y.-K., Chang, C.-K., Li, M.-H., Wu, Y.-C., & Wang, Y.-C. (2012). High-temperature indices associated with mortality and outpatient visits: Characterizing the association with elevated temperature. *Science of The Total Environment*, *427–428*, 41–49. https://doi.org/10.1016/j.scitotenv.2012.04.039

Maraun, D. (2016). Bias correcting climate change simulations - A critical review. *Current Climate Change Reports*, *2*(4), 211–220. https://doi.org/10.1007/s40641-016-0050-x

Masterson, J., & Richardson, F. (1979). *Humidex, a method of quantifying human discomfort due to excessive heat and humidity, Environment Canada* (Vol. 151, pp. 1–79). Downsview, Ontario: Atmospheric Environment Service.

Matthews, T. K. R., Wilby, R. L., & Murphy, C. (2017). Communicating the deadly consequences of global warming for human heat stress. *Proceedings of the National Academy of Sciences*, 3861–3866. https://doi.org/10.1073/pnas.1617526114

McMichael, A. J., Woodruff, R. E., & Hales, S. (2006). Climate change and human health: Present and future risks. *The Lancet*, *367*(9513), 859–869. https://doi.org/10.1016/S0140-6736(06)68079-3

Mitchell, D., Heaviside, C., Vardoulakis, S., Huntingford, C., Masato, G., Guillod, B. P., et al. (2016). Attributing human mortality during extreme heat waves to anthropogenic climate change. *Environmental Research Letters*, *11*(7), 074006. https://doi.org/10.1088/1748-9326/11/7/074006

Mora, C., Dousset, B., Caldwell, I. R., Powell, F. E., Geronimo, R. C., Bielecki, C. R., et al. (2017). Global risk of deadly heat. *Nature Climate Change*, *7*(7), 501–507. https://doi.org/10.1038/NCLIMATE3322

Morabito, M., Crisci, A., Messeri, A., Capecchi, V., Modesti, P. A., Gensini, G. F., et al. (2014). Environmental temperature and thermal indices: What is the most effective predictor of heat-related mortality in different geographical contexts? *Science World Journal*, *2014*, 1–15. https://doi.org/10.1155/2014/961750

Napoli, C. D., Pappenberger, F., & Cloke, H. L. (2018). Assessing heat-related health risk in Europe via the Universal Thermal Climate Index (UTCI). *International Journal of Biometeorology*, *62*(7), 1155–1165. https://doi.org/10.1007/s00484-018-1518-2

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., et al. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Orlov, A., Sillmann, J., Aaheim, A., Aunan, K., & de Bruin, K. (2019). Economic losses of heat-induced reductions in outdoor worker productivity: A case study of Europe. *Economics of Disasters and Climate Change*, *3*(3), 191–211. https://doi.org/10.1007/s41885-019-00044-0

Orlov, A., Sillmann, J., Aunan, K., Kjellstrom, T., & Aaheim, A. (2020). Economic costs of heat-induced reductions in worker productivity due to global warming. *Global Environmental Change*, *63*, 102087. https://doi.org/10.1016/j.gloenvcha.2020.102087

Pal, J. S., & Eltahir, E. A. (2016). Future temperature in southwest Asia projected to exceed a threshold for human adaptability. *Nature Climate Change*, *6*(2), 197–200. https://doi.org/10.1038/nclimate2833

Parsons, K. (2006). Heat stress standard ISO 7243 and its global application. *Industrial Health*, *44*(3), 368–379. https://doi.org/10.2486/indhealth.44.368

Racinais, S., Alonso, J. M., Coutts, A. J., Flouris, A. D., Girard, O., González-Alonso, J., et al. (2015). Consensus recommendations on training and competing in the heat. *Scandinavian Journal of Medicine & Science in Sports*, *25*, 6–19. https://doi.org/10.1111/sms.12467

Ragettli, M. S., Vicedo-Cabrera, A. M., Schindler, C., & Röösli, M. (2017). Exploring the association between heat and mortality in Switzerland between 1995 and 2013. *Environmental Research*, *158*, 703–709. https://doi.org/10.1016/j.envres.2017.07.021

Rodopoulou, S., Samoli, E., Analitis, A., Atkinson, R. W., de 'Donato, F. K., & Katsouyanni, K. (2015). Searching for the best modeling specification for assessing the effects of temperature and humidity on health: A time series analysis in three European cities. *International Journal of Biometeorology*, *59*(11), 1585–1596. https://doi.org/10.1007/s00484-015-0965-2

Ross, M. E., Vicedo-Cabrera, A. M., Kopp, R. E., Song, L., Goldfarb, D. S., Pulido, J., et al. (2018). Assessment of the combination of temperature and relative humidity on kidney stone presentations. *Environmental Research*, *162*, 97–105. https://doi.org/10.1016/j.envres.2017.12.020

Rothfusz, L. P. (1990). *The heat index equation*. Fort Worth, TX: National Weather Service Technical Attachment (SR 90–23).

Russo, S., Sillmann, J., Sippel, S., Barcikowska, M. J., Ghisetti, C., Smid, M., et al. (2019). Half a degree and rapid socioeconomic development matter for heatwave risk. *Nature Communications*, *10*(1), 136. https://doi.org/10.1038/s41467-018-08070-4

Russo, S., Sillmann, J., & Sterl, A. (2017). Humid heat waves at different warming levels. *Scientific Reports*, *7*(1), 7477. https://doi.org/10.1038/s41598-017-07536-7

Samir, K., & Lutz, W. (2017). The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. *Global Environmental Change*, *42*, 181–192. https://doi.org/10.1016/j.gloenvcha.2014.06.004

Seneviratne, S. I., Donat, M. G., Pitman, A. J., Knutti, R., & Wilby, R. L. (2016). Allowable CO_2 emissions based on regional and impact-related climate targets. *Nature*, *529*(7587), 477–483. https://doi.org/10.1038/nature16542

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., et al. (2012). Changes in climate extremes and their impacts on the natural physical environmenty. In C. B. Field, et al. (Eds.), *Managing the risks of extreme events and disasters to advance climate change adaptation. A special report of working groups I and II of the Intergovernmental Panel on Climate Change (IPCC)* (pp. 109–230). Cambridge, United Kingdom and NewYork, NY, USA: Cambridge University Press.

Sherwood, S. C., & Huber, M. (2010). An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences*, *107*(21), 9552–9555. https://doi.org/10.1073/pnas.0913352107

Shiogama, H., Hasegawa, T., Fujimori, S., Murakami, D., Takahashi, K., Tanaka, K., et al. (2019). Limiting global warming to 1.5 °C will lower increases in inequalities of four hazard indicators of climate change. *Environmental Research Letters*, *14*(12), 124022. https://doi.org/10.1088/1748-9326/ab5256

Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., & Bronaugh, D. (2013). Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *Journal of Geophysical Research: Atmosphere*, *118*(4), 1716–1733. https://doi.org/10.1002/jgrd.50203

Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., & Bronaugh, D. (2013). Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of Geophysical Research: Atmosphere*, *118*(6), 2473–2493. https://doi.org/10.1002/jgrd.50188

Spangler, K. R., Weinberger, K. R., & Wellenius, G. A. (2018). Suitability of gridded climate datasets for use in environmental epidemiology. *Journal of Exposure Science and Environmental Epidemiology*, *29*(6), 777–789. https://doi.org/10.1038/s41370-018-0105-2

Spector, J. T., & Sheffield, P. E. (2014). Re-evaluating occupational heat stress in a changing climate. *Annals of Occupational Hygiene*, *58*(8), 936–942. https://doi.org/10.1093/annhyg/meu073

Staiger, H., Laschewski, G., & Grätz, A. (2011). The perceived temperature – A versatile index for the assessment of the human thermal environment. Part a: Scientific basics. *International Journal of Biometeorology*, *56*(1), 165–176. https://doi.org/10.1007/s00484-011-0409-6

Steadman, R. G. (1979). The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. *Journal of Applied Meteorology*, *18*(7), 861–873. https://doi.org/10.1175/1520-0450(1979)018⟨0861:TAOSPI⟩2.0.CO;2

Steadman, R. G. (1984). A universal scale of apparent temperature. *Journal of Climate and Applied Meteorology*, *23*(12), 1674–1687. https://doi.org/10.1175/1520-0450(1984)023⟨1674:AUSOAT⟩2.0.CO;2

Thorsson, S., Rayner, D., Palm, G., Lindberg, F., Carlström, E., Börjesson, M., et al. (2020). Is physiological equivalent temperature (PET) a superior screening tool for heat stress risk than wet-bulb globe temperature (WBGT) index? Eight years of data from the Gothenburg half marathon. *British Journal of Sports Medicine*. https://doi.org/10.1136/bjsports-2019-100632

Urban, A., Hondula, D. M., Hanzlíková, H., & Kyselý, J. (2019). The predictability of heat-related mortality in Prague, Czech Republic, during summer 2015—A comparison of selected thermal indices. *International Journal of Biometeorology*, *63*(4), 535–548. https://doi.org/10.1007/s00484-019-01684-3

Urban, A., & Kyselý, J. (2014). Comparison of UTCI with other thermal indices in the assessment of heat and cold effects on cardiovascular mortality in the Czech Republic. *International Journal of Environmental Research and Public Health*, *11*(1), 952–967. https://doi.org/10.3390/ijerph110100952

Vaneckova, P., Neville, G., Tippett, V., Aitken, P., FitzGerald, G., & Tong, S. (2011). Do biometeorological indices improve modeling outcomes of heat-related mortality? *Journal of Applied Meteorology and Climatology*, *50*(6), 1165–1176. https://doi.org/10.1175/2011JAMC2632.1

Vogel, M. M., Orth, R., Cheruy, F., Hagemann, S., Lorenz, R., Hurk, B., et al. (2017). Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks. *Geophysical Research Letters*, *44*(3), 1511–1519. https://doi.org/10.1002/2016GL071235

Wang, S., & Zhu, J. (2019). Amplified or exaggerated changes in perceived temperature extremes under global warming. *Climate Dynamics*, *54*(1–2), 117–127. https://doi.org/10.1007/s00382-019-04994-9

Wang, Y., Wang, A., Zhai, J., Tao, H., Jiang, T., Su, B., et al. (2019). Tens of thousands additional deaths annually in cities of China between 1.5 °C and 2.0 °C warming. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-11283-w

Wartenburger, R., Hirschi, M., Donat, M. G., Greve, P., Pitman, A. J., & Seneviratne, S. I. (2017). Changes in regional climate extremes as a function of global mean temperature: An interactive plotting framework. *Geoscientific Model Development*, *10*(9), 3609–3634. https://doi.org/10.5194/gmd-10-3609-2017

Willett, K. M., & Sherwood, S. (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

Xiao, J., Peng, J., Zhang, Y., Liu, T., Rutherford, S., Lin, H., et al. (2014). How much does latitude modify temperature–mortality relationship in 13 eastern US cities? *International Journal of Biometeorology*, *59*(3), 365–372. https://doi.org/10.1007/s00484-014-0848-y

Zhang, K., Li, Y., Schwartz, J. D., & O'Neill, M. S. (2014). What weather variables are important in predicting heat-related mortality? A new application of statistical learning methods. *Environmental Research*, *132*, 350–359. https://doi.org/10.1016/j.envres.2014.04.004

Zhang, K., Rood, R. B., Michailidis, G., Oswald, E. M., Schwartz, J. D., Zanobetti, A., et al. (2012). Comparing exposure metrics for classifying 'dangerous heat' in heat wave and health warning systems. *Environment International*, *46*, 23–29. https://doi.org/10.1016/j.envint.2012.05.001

Zhao, Y., Ducharne, A., Sultan, B., Braconnot, P., & Vautard, R. (2015). Estimating heat stress from climate-based indicators: Present-day biases and future spreads in the CMIP5 global climate model ensemble. *Environmental Research Letters*, *10*(8), 084013. https://doi.org/10.1088/1748-9326/10/8/084013

Zscheischler, J., Fischer, E. M., & Lange, S. (2019). The effect of univariate bias adjustment on multivariate hazard estimates. *Earth System Dynamics*, *10*(1), 31–43. https://doi.org/10.5194/esd-10-31-2019

Erratum

The authors learned that data of one of the 25 climate models that they used is faulty and the model should not be included. The paper has been corrected to remove this model and update the results. The plots and numbers have changed only very slightly and the conclusions of the paper are unaffected. This corrected version may be considered the version of record.