



OPEN

DATA DESCRIPTOR

Wet-Bulb Globe Temperature, Universal Thermal Climate Index, and Other Heat Metrics for US Counties, 2000–2020

Keith R. Spangler¹✉, Shixin Liang^{1,2} & Gregory A. Wellenius¹

Epidemiologic research on extreme heat consistently finds significant impacts on human morbidity and mortality. However, most of these analyses do not use spatially explicit measures of heat (typically assessing exposures at major cities using the nearest weather station), and they frequently consider only ambient temperature or heat index. The field is moving toward more expansive analyses that use spatially resolved gridded meteorological datasets and alternative assessments of heat, such as wet-bulb globe temperature (WBGT) and universal thermal climate index (UTCI), both of which require technical geoscientific skills that may be inaccessible to many public health researchers. To facilitate research in this domain, we created a database of population-weighted, spatially explicit daily heat metrics – including WBGT, UTCI, heat index, dewpoint temperature, net effective temperature, and humidex – for counties in the conterminous United States derived from the ERA5-Land gridded data set and using previously validated equations and algorithms. We also provide an R package to calculate these metrics, including gold-standard algorithms for estimating WBGT and UTCI, to facilitate replication.

Background & Summary

Exposure to moderate and extreme heat is associated with increased risk of illness and death¹. Although most epidemiologic studies contributing to this conclusion use dry-bulb temperature (i.e., ambient air temperature measured in the shade) as the exposure metric of interest^{2–6}, others have assessed alternative metrics that incorporate humidity – such as the heat index, a combined temperature and humidity metric used by the U.S. National Weather Service⁷, or the humidex, which is used by the Meteorological Service of Canada – to better estimate the physiologic impact of heat on the human thermoregulatory system^{8–10}. Increasingly, public health research on heat is considering metrics that combine additional meteorological conditions, including solar radiation and wind speed, to further contextualize the actual heat stress experienced by populations^{11–13}.

Of particular interest is the wet-bulb globe temperature (WBGT), a thermal index originally developed in the 1950s to establish epidemiologically relevant thermal thresholds to prevent heat-related illnesses at US military training camps¹⁴. The WBGT is a weighted average of the ambient, wet-bulb, and globe temperatures, which collectively incorporate thermal, solar, and convective heat transfers from ambient temperature, humidity, solar radiation, and wind speed¹⁵. In contrast to simpler, more commonly used metrics, such as ambient temperature or heat index, WBGT is measured in conditions of direct solar radiation and is partially mitigated by wind speed, making it an appealing measure for estimating thermal conditions experienced by outdoor workers and athletes. In recognition of this utility, it has been approved by the International Organization for Standardization (ISO), the American Conference of Governmental Industrial Hygienists, and other national and international organizations for use as a thermal stress screening tool¹⁶.

However, it should be noted that WBGT has limitations, including its potential underestimation of thermal stress in conditions where sweating is restricted, susceptibility to measurement errors, and variability based on clothing and activity¹⁴. Others have also noted that the scale of the measurement is prone to misinterpretation, given that extreme values of WBGT are much lower than what would be considered extreme by ambient

¹Boston University School of Public Health, Department of Environmental Health, Boston, MA, USA. ²Boston University, Department of Mathematics & Statistics, Boston, MA, USA. ✉e-mail: krspangl@bu.edu

temperature standards¹⁶. Newer metrics have been developed that may reduce some of these limitations; of interest here is the Universal Thermal Climate Index (UTCI), a heat metric derived from human energy balance models with the goal of being universally applicable, physiologically relevant, and appropriate for use in a range of bioclimatic applications¹⁷. Although the UTCI is correlated with WBGT and may similarly reflect thermal perception¹⁸, some have advocated for replacing WBGT with UTCI in operational use cases, particularly in athletics^{19,20}.

While WBGT and UTCI have clear utility in bioclimatic contexts, there are challenges for population health researchers hoping to use them in their analyses. For example, measuring WBGT requires specialized equipment that is not widely deployed operationally, and UTCI similarly requires data that are frequently unavailable in meteorological data sets. Although it would be advantageous for public health researchers to have easy access to population-scale estimates of WBGT, UTCI, and other heat metrics, no such database presently exists. Furthermore, no single measure of heat will be universally superior in all contexts^{10,19}, suggesting the need for a single dataset containing multiple metrics for intercomparison.

Concurrently, heat-health researchers are increasingly aware of the value in using gridded meteorological data sets, both for their potential to provide population-applicable estimates of weather experienced across large areas and to avail rural areas without weather stations to epidemiologic analyses²¹. Recent studies have demonstrated the utility of a range of gridded data sets for such applications^{22–25}. Although these data sets are a valuable source of information, they contain massive amounts of data that require time, computational resources, and expertise to process that are not available to many public health researchers.

Given these barriers to enhanced epidemiologic research, the broader goal of this data set is to provide: (1) a data set of pre-processed gridded data, population-weighted to the daily county level; and (2) code for calculating various heat metrics, including wet-bulb globe temperature and the Universal Thermal Climate Index using gold-standard approaches. The potential reuse value of these data extends not only to epidemiologic analyses, but to any research that requires daily, county-level estimates of temperature or other heat metrics.

Methods

Overview. Our dataset provides daily minimum, maximum, and mean values of ambient temperature, dew-point temperature, net effective temperature, heat index, humidex, wet-bulb globe temperature, and the Universal Thermal Climate Index, all population-weighted at the county level. We used a high-resolution reanalysis dataset to calculate these hourly heat metrics across the entire contiguous United States (CONUS) from January 2000 to December 2020. We developed an accompanying R package (*heatmetrics*) that can be used to replicate the calculations on other data sets (Fig. 1).

Reanalysis data. We derived a series of heat metrics using data from the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 Land product (ERA5-Land)^{26,27}. The ERA5-Land variables we obtained included: (1) two-meter air temperature, (2) two-meter dew point temperature, (3) surface pressure, (4) ten-meter zonal and meridional wind velocity vectors, (5) surface solar radiation downward, (6) surface thermal radiation downward, (7) surface net solar radiation, and (8) surface net thermal radiation. The ERA5-Land data are available hourly at a spatial resolution of 0.1 degrees (approximately 9 km) globally over land. We also obtained total sky direct solar radiation at surface from ERA5^{28,29}, the reanalysis from which ERA5-Land is derived. We interpolated this from the 0.25-degree ERA5 grid to the 0.1-degree ERA5-Land grid using nearest-neighbor interpolation, following the approach of Yan *et al.*³⁰.

Converting to local standard time. Data from ERA5-Land are provided hourly in Coordinated Universal Time (UTC), also known as Greenwich Mean Time (GMT). Since many users need minimum and maximum values relative to the *local* day, we created raster stacks of the calculated hourly data that aligned with the local time zone. To do this, we created rasters of ERA5-Land grids containing centroids of latitude and longitude and then used the *lutz* R package³¹ to identify the time zone of each grid cell, which was represented as a numeric offset from UTC time. We created separate rasters for standard time and daylight saving time (for example, Eastern Standard Time has a UTC offset of –5 hours while Eastern Daylight Time has an offset of –4 hours). While most locations had consistent time zones throughout the study period, we manually accounted for time zone changes that occurred in the state of Indiana in 2006 and 2007 by creating shapefiles of affected counties and adjusting the ERA5-Land grid offsets accordingly. The final result was eight rasters representing grid-level UTC offsets: four each for local standard time (LST) and local daylight time (LDT) for years 2000–2005, 2006, 2007, and 2008–2020. We used these time zone rasters as masks to subset the calculated heat metrics corresponding to the local day. The result was stacks of 24 hourly observations for each day that reflected data from 00 local time to 23 local time. We note that although hourly time steps are the highest temporal resolution available in reanalysis data, “true” daily minimum and maximum temperatures can occur between hours. Others have found that, in some instances, this can lead to small differences in estimated relative risks of exposure to ambient heat, but that such differences have considerably smaller magnitudes than the differences in relative risks found between different heat metrics³².

Calculating heat metrics. We calculated hourly heat metrics across all of CONUS between January 2000 and December 2020 on the stacks of ERA5-Land data converted to local time. We then calculated, at the pixel level, the minimum, maximum, and mean values for each variable on every local day. The final results were rasters of CONUS for daily observations of each of the heat metrics described in this section. The database may be updated in the future as ERA5-Land data are updated.

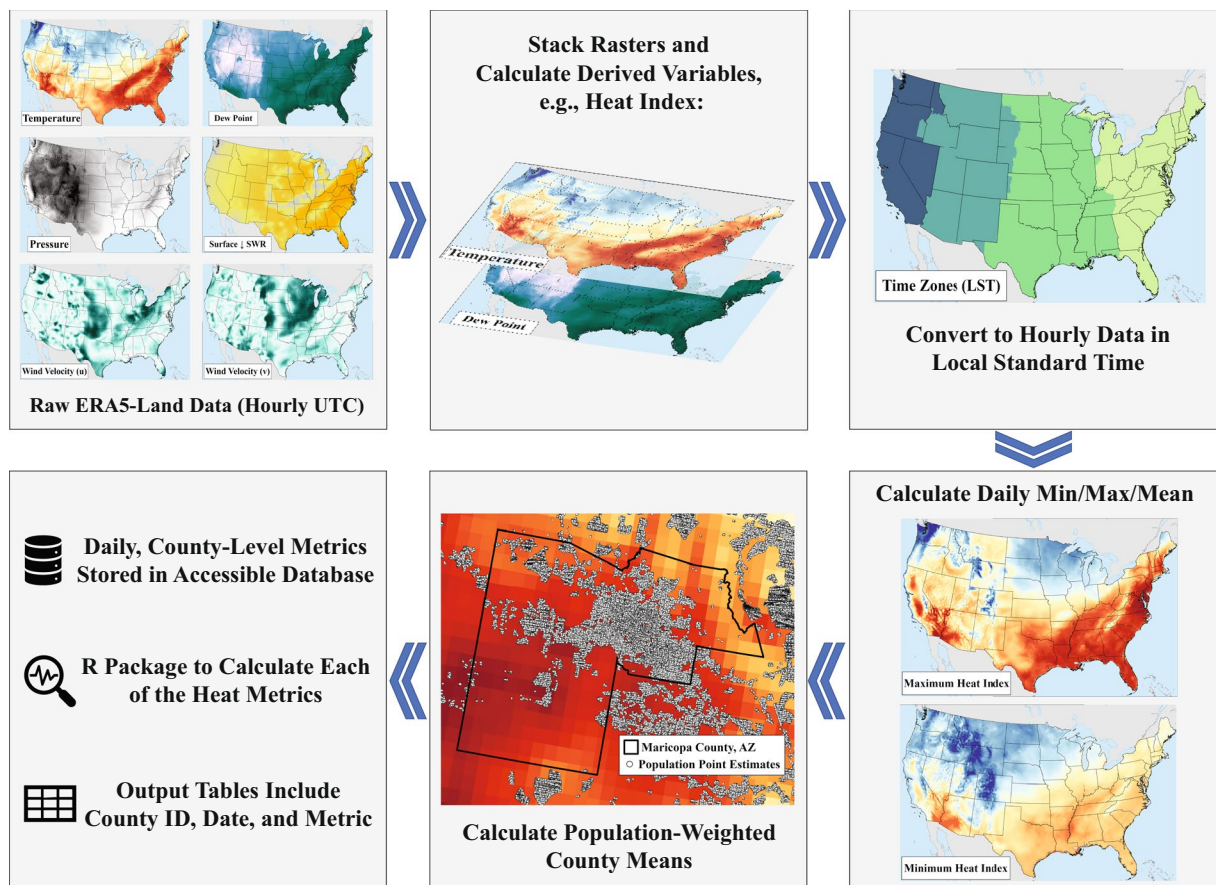


Fig. 1 Schematic overview of the creation of the *heatmetrics* database.

Wet-bulb globe temperature. Lemke and Kjellstrom³³ summarized and compared approaches to estimating wet-bulb globe temperature, and they found that the algorithm by Liljegren *et al.*¹⁵ performed most accurately across a range of conditions. We therefore employed this algorithm in our calculations here. In contrast to most estimates of WBGT – which use relatively simple equations – the Liljegren approach applies fundamental laws of physics, thermodynamics, and mass balance to separately model natural wet bulb (T_w) and globe temperatures (T_G), which together comprise 90% of the WBGT (the remaining 10% is the ambient temperature [T_A]; Eq. 1).

$$WBGT = 0.70 \cdot T_w + 0.20 \cdot T_G + 0.1 \cdot T_A \quad (1)$$

The Liljegren algorithm was originally written in the C programming language but was translated to Javascript and made available in a web interface by the Occupational Safety and Health Administration (OSHA)³⁴. We translated this Javascript into an R package and processed it on raster stacks of the ERA5-Land data.

The WBGT algorithm requires as inputs the following variables (Table 1): year, month, day, hour, minute, UTC offset, averaging time of meteorological observations, latitude, longitude, temperature, relative humidity, incident solar radiation, wind speed, surface pressure, height of wind speed measurement, vertical temperature difference between observed temperature and wind-speed-height temperature, and urban or rural land cover. Although most of these variables are provided directly by ERA5-Land, a few needed to be derived first. Relative humidity was calculated using air temperature and dew-point temperature, and the wind speed was calculated using the zonal and meridional wind velocity vectors (Table 1). To identify ERA5-Land grid cells as urban or rural, we used the 2011 National Land Cover Database³⁵ and assigned grid cells as “urban” (1) if at least 33% of the 30-meter by 30-meter land cover pixels within each ERA5-Land grid cell were classified as “developed, low intensity,” “developed, medium intensity,” or “developed, high intensity.” All other grid cells were classified as “rural” (0). Note that the contribution of urban/rural to the WBGT calculation is marginal and only affects the conversion of wind speed from higher altitudes to two-meter equivalents.

In addition to the pre-processing described above, we made a few small modifications to the Liljegren algorithm, as described here and as comments throughout the R source code. First, in the calculation of the natural wet-bulb temperature, the original algorithm used a static enhancement factor of 1.004 when calculating the saturation vapor pressure; however, this factor assumes a barometric pressure of at least 800 hPa. To accommodate lower pressures for high-elevation locations, we applied an enhancement factor that is a function of barometric pressure (second quantity of Eq. 2), following Equation 8 in Buck (1981, p. 1532)³⁶.

Variable	Calculation	Units/Notes
Year	Provided by ERA5-Land	Numeric (YYYY)
Month	Provided by ERA5-Land	Numeric (MM), 1–12
Day	Provided by ERA5-Land	Numeric (DD), 1–31
Hour	Provided by ERA5-Land	Numeric (HH), 0–23 in local time
Minute	Set to 0 for all observations	Top-of-hour observations
UTC Offset	Set to 0 for all observations because we calculated on the original UTC data	Number of hours difference from UTC, if applicable
Averaging Time	Set to 0 for all observations	Observations are provided as either instantaneous or already averaged over the hour
Latitude	ERA5-Land grid centroids	Decimal-degrees North (°N)
Longitude	ERA5-Land grid centroids	Decimal-degrees East (°E)
Ambient Temperature	Provided by ERA5-Land	Degrees Celsius (°C)
Relative Humidity	$e = 610.94 \cdot \exp\left(\frac{17.625 \cdot T_D}{243.04 + T_D}\right)$	T_D = dew point temperature (°C); see Lawrence (2005) ⁴⁹
	$e_s = 610.94 \cdot \exp\left(\frac{17.625 \cdot T_A}{243.04 + T_A}\right)$	T_A = ambient air temperature (°C); see Lawrence (2005) ⁴⁹
	$RH = 100 \cdot \left[\frac{e}{e_s}\right]$	e = vapor pressure (Pascals [Pa]) and e_s = saturation vapor pressure (Pascals [Pa])
Surface Solar Radiation Downward	Provided by ERA5-Land	Watts per square-meter (W/m ²)
Wind Speed	$w_s = \sqrt{u^2 + v^2}$	u = zonal wind velocity (m/s) and v = meridional wind velocity (m/s)
Surface Pressure	Provided by ERA5-Land	Pascals (Pa)
Wind-Speed Height	Set to 10 meters for all observations	N/A
Temperature Lapse	Set to -0.052°C for all observations	Assumes vertical lapse rate of $-6.5^\circ\text{C} / \text{km}$
Urban	Assigned 1 (urban) to pixels classified as urban or built up in the National Land Cover Database 2011 dataset and 0 (rural) otherwise	N/A

Table 1. Inputs to wet-bulb globe temperature (WBGT) algorithm.

$$e_{sat} = \left(6.1121 \cdot \exp\left[\frac{17.502 \cdot (T_{air} - 273.15)}{T_{air} - 32.18}\right] \right) \cdot (1.0007 \cdot 3.46 \cdot 10^{-6} \cdot P_{air}) \quad (2)$$

Where T_{air} is the ambient temperature in Kelvin and P_{air} is the barometric pressure in hPa.

Second, we fixed a small error in the stability classes, which are used to estimate the 2-meter wind speeds when the input wind speeds are measured at a different height (as is the case in ERA5-Land, which reports 10-meter wind speeds). In the original algorithm, nighttime conditions with wind speeds between 2 and 2.5 m/s were given stability classes of “E” and “F” for positive and negative lapse rates, respectively; we changed this to the correct values of “D” and “E,” consistent with guidance from EPA documentation (see Table 6–7 in reference)³⁷.

Third, we updated the algorithm for calculating the heat of vaporization to follow the approach of Meyra *et al.*³⁸, which was found to be more accurate than the traditional Watson equation. Although we believe this to be a more-accurate method, we find that it typically changes the estimate of the wet-bulb temperature by less than 0.1 degrees Celsius.

Finally, we changed the minimum wind speed from 0.13 m/s to 0.5 m/s for a more-conservative estimate of WBGT that prevents unreasonably high estimates that result from very low wind speeds. Lemke and Kjellstrom (2012)³³ recommend a more-conservative minimum wind speed of 1 m/s for assessing WBGT effects on outdoor workers, noting that typical bodily movement results in an apparent wind speed on the skin of at least 1 m/s. Our value of 0.5 m/s is a balance between these values and helps capture the WBGT for stationary individuals, as our index is not exclusively for outdoor workers. Note that this wind speed adjustment was done as a pre-processing step and that the minimum wind speed is only directly set to 0.5 m/s in the WBGT function of the *heatmetrics* R package when supplying wind speeds at a height other than two meters.

We conducted a sensitivity analysis to determine the impact that these changes to the Liljegren algorithm had on the final county-level mean values. We found that daily maximum WBGT estimates for August 2020 across all available CONUS counties using WBGT algorithms with and without the aforementioned alterations were extremely similar, with $r^2 = 99.99\%$, a mean absolute difference of 0.08°C , and a maximum absolute difference of 0.25°C .

Universal Thermal Climate Index (UTCI). We calculated UTCI following the approach of Di Napoli *et al.*³⁹, as implemented in the ECMWF python library, *thermofeel*⁴⁰. This method uses the sixth-order polynomial regression approximation given by Bröde *et al.* (2011), which is a highly accurate approximation of UTCI with a root-mean square error of 1.1 degrees Celsius⁴¹. The equation takes as inputs the ambient temperature, 10-meter wind speed, vapor pressure, and mean radiant temperature (T_{mrt}). We calculated mean radiant temperature

Variable	Calculation	Variables/Units	Ref.
Ambient Temperature	Provided directly by ERA5-Land	°C	N/A
Dew-Point Temperature	Provided directly by ERA5-Land	°C	N/A
Net Effective Temperature	$37 - \left(\frac{37 - T_A}{0.68 - 0.0014 \cdot RH + \frac{1}{1.76 + 1.4 \cdot w^{0.75}}} \right) - \left[0.29 \cdot T_A \cdot \left(1 - \frac{RH}{100} \right) \right]$	RH – rel. hum. (%) T _A – air temp. (°C) w – wind speed (m/s)	50
Humidex	$e = 6.1094 \cdot \exp\left(\frac{17.625 \cdot T_D}{243.04 + T_D}\right)$	T _D – dew-point T (°C)	49
	$T_A + \frac{5}{9} \cdot [e - 10]$	T _A – air temp. (°C) e – vapor pressure (hPa)	51
Heat Index	<i>Weathermetrics</i> R package following algorithm of National Weather Service, using ambient temperature and dew point temperature as inputs	°C	7
Wet-Bulb Globe Temperature (WBGT)	Liljegren <i>et al.</i> approach as described in the Methods section here and in the reference.	°C	15
Universal Thermal Climate Index (UTCI)	Mean radiant temperature (T _{mrt}) estimated using Di Napoli <i>et al.</i> (2020) approach	°C	42
	Bröde <i>et al.</i> (2011) approach as implemented by Di Napoli <i>et al.</i> (2021) and Brimicombe <i>et al.</i> (2022). See Methods for details.	°C	39–41

Table 2. Summary of the methods used to calculate the heat metrics in the data set.

following the approach of Di Napoli *et al.*⁴², which approximates T_{mrt} using hourly measurements of total downward surface solar radiation (direct and diffuse), surface *net* solar radiation, downward surface thermal radiation, surface *net* thermal radiation, downward direct surface solar radiation, and cosine of the solar zenith angle (cza). As was done in Di Napoli *et al.*⁴², we calculated the average *daytime* cza in order to minimize errors that arise during sunrise and sunset hours, described comprehensively by Hogan and Hirahara⁴³. We used this same integrated cza approach in the calculation of WBGT and note that, although it is the most-accurate approach, others have found that it actually has a very small impact (<0.01 °C on average) on the overall estimation of UTCI⁴⁰. Although the Bröde *et al.* (2011) algorithm is suitable for wind speeds up to 30.3 m/s, we followed Di Napoli *et al.* (2021) in capping wind speeds at 17 m/s and marking as missing (“NA”) observations above this threshold, based on findings of extremely low UTCI values at these tropical-storm-force wind speeds⁴⁴. Finally, consistent with recommendations in Bröde *et al.*⁴¹, we constrained the vapor pressure input to be consistent with a relative humidity of ≥5%, the lower bound for which their algorithm is validated: in cases where relative humidity was less than 5%, we set the vapor pressure equal to the saturation vapor pressure multiplied by 0.05. This had only a minimal impact on the final county UTCI values: 99.93% of county-day maximum UTCI values were completely unaffected. Among the county-days that did have the adjustment, the mean absolute difference in maximum UTCI was 0.06 °C and the maximum absolute difference was 0.58 °C.

Other heat metrics. Most of the other heat metrics in this data set use relatively straightforward equations (Table 2). The one exception is heat index: we calculated this variable using the *weathermetrics* R package⁷, which implements the approach to calculating heat index that is used by the National Weather Service. Note that, for consistency with how we calculated WBGT, we set the minimum wind speed to 0.5 m/s for all variable calculations that use wind speed.

Calculating population-weighted county means. Daily heat metric values are reported as population-weighted county mean values. We used high-resolution (approximately 250 m × 250 m) population data from the Joint Research Centre (JRC) of the European Commission⁴⁵ to calculate spatial weights for each ERA5-Land grid cell within each county based on the proportion of the county population residing in that grid cell. For example, if the sum of the high-resolution population points within a particular grid cell in County A were equal to 10% of the total county population, then that grid cell would get weighted 10% toward the overall county mean. Temperatures in the more densely populated parts of counties were thus weighted more heavily than the less-populous parts, resulting in metrics that are likely more relevant to population-based studies. To account for potential population shifts over the period analyzed, we used two sets of population estimates based on availability in the JRC data: we used 2000 population distributions for county means from 2000–2009 and 2015 populations for 2010–2020.

To account for missing data, we added flag variables to indicate whether the county estimates for a particular day were based on non-missing grid cells representing less than 50% of the population. This is applicable because ERA5-Land is available only over land areas in which the grid cell is comprised of no more than 50% ocean; this ensures that the meteorological data are representative of land areas, but it also means that some small island and coastal areas are excluded. An additional source of missing data is from particular hourly values being marked as NA (for example, hourly UTCI values when wind speeds exceed 17 m/s). A grid cell was marked as NA for a particular variable-day if fewer than 21 hourly observations were available. We calculated the percent of county populations that were represented by non-missing ERA5-Land data for every variable on every day and added flags as follows: “0” means the estimate is based on data representing ≥50% of the population, “1” covers 10–49% of the population, “2” covers <10%, and “3” means the data are completely missing (variable-days in this case are marked as NA). These flags affect only an extremely small portion of the data set: >99.7% of

Variable Name (Short)	Variable Name (Long)	Description / Format	Units
StCoFIPS	State-county Federal Information Processing Standard (FIPS) Identifier	Unique county identifier: concatenation of two-digit state identifier and three-digit county identifier	N/A
Date	Date	Local day in the format YYYYMMDD	N/A
Tmin_C	Daily Minimum Ambient Temperature	Lowest 2-meter ambient temperature observed from hourly data from 00 LST to 23 LST	°C
Tmax_C	Daily Maximum Ambient Temperature	Highest 2-meter ambient temperature observed from hourly data from 00 LST to 23 LST	°C
Tmean_C	Daily Mean Ambient Temperature	2-meter ambient temperature averaged over hourly observations from 00 LST to 23 LST	°C
TDmin_C	Daily Minimum Dew Point Temperature	Lowest dew point temperature observed from hourly data from 00 LST to 23 LST	°C
TDmax_C	Daily Maximum Dew Point Temperature	Highest dew point temperature observed from hourly data from 00 LST to 23 LST	°C
TDmean_C	Daily Mean Dew Point Temperature	Dew point temperature averaged over hourly observations from 00 LST to 23 LST	°C
NETmin_C	Daily Minimum Net Effective Temperature	Lowest net effective temperature observed from hourly data from 00 LST to 23 LST	°C
NETmax_C	Daily Maximum Net Effective Temperature	Highest net effective temperature observed from hourly data from 00 LST to 23 LST	°C
NETmean_C	Daily Mean Net Effective Temperature	Net effective temperature averaged over hourly observations from 00 LST to 23 LST	°C
HImin_C	Daily Minimum Heat Index	Lowest heat index observed from hourly data from 00 LST to 23 LST	°C
HImax_C	Daily Maximum Heat Index	Highest heat index observed from hourly data from 00 LST to 23 LST	°C
HImean_C	Daily Mean Heat Index	Heat index averaged over hourly observations from 00 LST to 23 LST	°C
HXmin_C	Daily Minimum Humidex	Lowest humidex observed from hourly data from 00 LST to 23 LST	°C
HXmax_C	Daily Maximum Humidex	Highest humidex observed from hourly data from 00 LST to 23 LST	°C
HXmean_C	Daily Mean Humidex	Humidex averaged over hourly observations from 00 LST to 23 LST	°C
WBGmin_C	Daily Minimum Wet-Bulb Globe Temperature	Lowest wet-bulb globe temperature (WBGT) observed from hourly data from 00 LST to 23 LST	°C
WBGmax_C	Daily Maximum Wet-Bulb Globe Temperature	Highest wet-bulb globe temperature (WBGT) observed from hourly data from 00 LST to 23 LST	°C
WBGmean_C	Daily Mean Wet-Bulb Globe Temperature	Wet-bulb globe temperature (WBGT) averaged over hourly observations from 00 LST to 23 LST	°C
UTCImin_C	Daily Minimum Universal Thermal Climate Index	Lowest Universal Thermal Climate Index (UTCI) observed from hourly data from 00 LST to 23 LST	°C
UTCImax_C	Daily Maximum Universal Thermal Climate Index	Highest Universal Thermal Climate Index (UTCI) observed from hourly data from 00 LST to 23 LST	°C
UTCImean_C	Daily Mean Universal Thermal Climate Index	Universal Thermal Climate Index (UTCI) averaged over hourly data from 00 LST to 23 LST	°C
Flag_T	Ambient temperature flag	Indicator of the percent of county population represented by the county-day ambient temperature estimate. 0: $\geq 50\%$, 1: 10–49%, 2: $< 10\%$, 3: 0% (NA)	N/A
Flag_TD	Dew point temperature flag	Indicator of the percent of county population represented by the county-day dew point temperature estimate. 0: $\geq 50\%$, 1: 10–49%, 2: $< 10\%$, 3: 0% (NA)	N/A
Flag_NET	Net effective temperature flag	Indicator of the percent of county population represented by the county-day net effective temperature estimate. 0: $\geq 50\%$, 1: 10–49%, 2: $< 10\%$, 3: 0% (NA)	N/A
Flag_HI	Heat index flag	Indicator of the percent of county population represented by the county-day heat index estimate. 0: $\geq 50\%$, 1: 10–49%, 2: $< 10\%$, 3: 0% (NA)	N/A
Flag_HX	Humidex flag	Indicator of the percent of county population represented by the county-day humidex estimate. 0: $\geq 50\%$, 1: 10–49%, 2: $< 10\%$, 3: 0% (NA)	N/A
Flag_WBGT	Wet-bulb globe temperature flag	Indicator of the percent of county population represented by the county-day WBGT estimate. 0: $\geq 50\%$, 1: 10–49%, 2: $< 10\%$, 3: 0% (NA)	N/A
Flag_UTCI	Universal Thermal Climate Index flag	Indicator of the percent of county population represented by the county-day UTCI estimate. 0: $\geq 50\%$, 1: 10–49%, 2: $< 10\%$, 3: 0% (NA)	N/A

Table 3. Description of variables available in the heatmetrics database.

county-days have no flag, and only two counties in all of CONUS are missing entirely from the data set (Monroe County, Florida [containing the Key West archipelago] and Nantucket County, Massachusetts [containing the island of Nantucket]).

Data Records

The *heatmetrics* data are accessible via figshare⁴⁶. The variables currently available for download are described in Table 3. At present, the data set includes population-weighted estimates at the county level, which can be queried using the state-county Federal Information Processing Standard (FIPS) identifier. We also provide a separate data set of *unweighted* county mean values, which were created by taking a simple average of all grid cells in a county, which are also available via figshare⁴⁷. Variable names are the same between the two data sets, so users should take care to download the applicable file for their needs and rename variables as appropriate if using both data sets simultaneously.

Technical Validation

The *heatmetrics* data set employs existing algorithms and an established reanalysis product that have all been peer-reviewed and frequently cited in the literature. Please see accompanying references and citations therein for the input data set used, ERA5-Land, for model development and validation²⁷. The WBG algorithm used here is based on the Liljegren approach, which was found to be accurate to within 1 °C in the developers' testing¹⁵, and independently verified as being the most accurate across different estimation methods³³. Similarly, we followed the approach of Di Napoli *et al.*³⁹, as implemented by Brimicombe *et al.*⁴⁰ for operational distribution through the European Centre for Medium-Range Weather Forecasts (ECMWF); this algorithm employs the UTCI approximation reported by Bröde *et al.* (2011), which was found to have a root-mean square error of approximately 1.1 °C.

Disclaimers. This data set contains modified Copernicus Climate Change Service information (2022), as described and cited in the manuscript. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. The data set and software are provided by the manuscript authors “as is” with no warranty of any kind.

Code availability

We developed the *heatmetrics* R package to facilitate replication of these methods to other meteorological data sets. The package is available to download via figshare⁴⁸.

Received: 3 August 2021; Accepted: 18 May 2022;

Published online: 17 June 2022

References

- Sarofim, M. C. *et al.* In *The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment* (ed. Crimmins, A. *et al.*) Ch. 2: Temperature-related death and illness <https://doi.org/10.7930/J0MG7MDX> (U.S. Global Change Research Program, 2016).
- Gasparrini, A. *et al.* Temporal Variation in Heat-Mortality Associations: A Multicountry Study. *Environ. Health Persp.* **123**, 1200–1207, <https://doi.org/10.1289/ehp.1409070> (2015).
- Medina-Ramon, M. & Schwartz, J. Temperature, temperature extremes, and mortality: a study of acclimatisation and effect modification in 50 US cities. *Occup. Environ. Med.* **64**, 827–833, <https://doi.org/10.1136/oem.2007.033175> (2007).
- Knowlton, K. *et al.* The 2006 California heat wave: impacts on hospitalizations and emergency department visits. *Environ. Health Persp.* **117**, 61–67, <https://doi.org/10.1289/ehp.11594> (2009).
- Bobb, J. F., Peng, R. D., Bell, M. L. & Dominici, F. Heat-related mortality and adaptation to heat in the United States. *Environ. Health Persp.* **122**, 811–816, <https://doi.org/10.1289/ehp.1307392> (2014).
- Weinberger, K. R., Harris, D., Spangler, K. R., Zanobetti, A. & Wellenius, G. A. Estimating the number of excess deaths attributable to heat in 297 United States counties. *Environ. Epidemiol.* **4** <https://doi.org/10.1097/EE9.0000000000000096> (2020).
- Anderson, G. B., Bell, M. L. & Peng, R. D. Methods to calculate the heat index as an exposure metric in environmental health research. *Environ. Health Persp.* **121**, 1111–1119, <https://doi.org/10.1289/ehp.1206273> (2013).
- Metzger, K. B., Ito, K. & Matte, T. D. Summer Heat and Mortality in New York City: How Hot Is Too Hot? *Environ. Health Persp.* **118**, 80–86, <https://doi.org/10.1289/ehp.0900906> (2010).
- Wellenius, G. A. *et al.* Heat-related morbidity and mortality in New England: Evidence for local policy. *Environ. Res.* **156**, 845–853, <https://doi.org/10.1016/j.envres.2017.02.005> (2017).
- Barnett, A. G., Tong, S. & Clements, A. C. A. What measure of temperature is the best predictor of mortality? *Environ. Res.* **110**, 604–611, <https://doi.org/10.1016/j.envres.2010.05.006> (2010).
- Heo, S. & Bell, M. L. Heat waves in South Korea: differences of heat wave characteristics by thermal indices. *J. Expo. Sci. Environ. Epidemiol.* **29**, 790–805, <https://doi.org/10.1038/s41370-018-0076-3> (2019).
- Urban, A., Hondula, D. M., Hanzlikova, H. & Kysely, J. The predictability of heat-related mortality in Prague, Czech Republic, during summer 2015—a comparison of selected thermal indices. *Int. J. Biometeorol.* **63**, 535–548, <https://doi.org/10.1007/s00484-019-01684-3> (2019).
- Heo, S., Bell, M. L. & Lee, J. 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> (2019).
- Budd, G. M. Wet-bulb globe temperature (WBGT)—its history and its limitations. *Journal of science and medicine in sport* **11**, 20–32, <https://doi.org/10.1016/j.jsams.2007.07.003> (2007).
- Liljegren, J. C., Carhart, R. A., Lawday, P., Tschopp, S. & Sharp, R. Modeling the Wet Bulb Globe Temperature Using Standard Meteorological Measurements. *J. Occup. Environ. Hyg.* **5**, 645–655, <https://doi.org/10.1080/15459620802310770> (2008).
- Dambrosio Alfano, F. R., Malchaire, J., Palella, B. I. & Riccio, G. WBGT index revisited after 60 years of use. *Ann. Occup. Hyg.* **58**, 955–970, <https://doi.org/10.1093/annhyg/meu050> (2014).
- Jendritzky, G., de Dear, R. & Havenith, G. UTCI—Why another thermal index? *Int J Biometeorol* **56**, 421–428, <https://doi.org/10.1007/s00484-011-0513-7> (2011).
- Zare, S. *et al.* 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> (2018).
- Grundstein, A. & Vanos, J. There is no ‘Swiss Army Knife’ of thermal indices: the importance of considering ‘why?’ and ‘for whom?’ when modelling heat stress in sport. *British journal of sports medicine* **55**, 822–824, <https://doi.org/10.1136/bjsports-2020-102920> (2021).

20. Brocherie, F. & Millet, G. P. Is the wet-bulb globe temperature (WBGT) index relevant for exercise in the heat? *Sports Medicine* **45**, 1619–1621, <https://doi.org/10.1007/s40279-015-0386-8> (2015).
21. Spangler, K. R., Weinberger, K. R. & Wellenius, G. A. Suitability of gridded climate datasets for use in environmental epidemiology. *J. Expo. Sci. Environ. Epidemiol.* **29**, 777–789, <https://doi.org/10.1038/s41370-018-0105-2> (2018).
22. Weinberger, K. R., Spangler, K. R., Zanobetti, A., Schwartz, J. D. & Wellenius, G. A. Comparison of temperature-mortality associations estimated with different exposure metrics. *Environ. Epidemiol.* **3**, e072, <https://doi.org/10.1097/EE9.0000000000000072> (2019).
23. Isaksen, T. B. *et al.* Increased mortality associated with extreme-heat exposure in King County, Washington, 1980–2010. *Int J Biometeorol* **60**, 85–98, <https://doi.org/10.1007/s00484-015-1007-9> (2015).
24. Royé, D., Íñiguez, C. & Tobias, A. Comparison of temperature–mortality associations using observed weather station and reanalysis data in 52 Spanish cities. *Environmental research* **183**, 109237, <https://doi.org/10.1016/j.envres.2020.109237> (2020).
25. Vaidyanathan, A. *et al.* Assessment of extreme heat and hospitalizations to inform early warning systems. *Proceedings of the National Academy of Sciences - PNAS* **116**, 5420–5427, <https://doi.org/10.1073/pnas.1806393116> (2019).
26. Copernicus Climate Change Service. ERA5-Land Hourly Data from 2001 to Present. <https://doi.org/10.24381/cds.e2161bac>. Accessed March 1, 2022.
27. Muñoz-Sabater, J. *et al.* ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth system science data* **13**, 4349–4383, <https://doi.org/10.5194/essd-13-4349-2021> (2021).
28. Hersbach, H. *et al.* ERA5 Hourly Data on Single Levels from 1979 to Present. <https://doi.org/10.24381/cds.adbb2d47>. Accessed March 8, 2022.
29. Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly journal of the Royal Meteorological Society* **146**, 1999–2049, <https://doi.org/10.1002/qj.3803> (2020).
30. Yan, Y., Xu, Y. & Yue, S. A high-spatial-resolution dataset of human thermal stress indices over South and East Asia. *Scientific data* **8**, 229, <https://doi.org/10.1038/s41597-021-01010-w> (2021).
31. Teucher, A. lutz: Look Up Time Zones of Point Coordinates. <https://CRAN.R-project.org/package=lutz> (2019).
32. Davis, R. E., Hondula, D. M. & Patel, A. P. Temperature Observation Time and Type Influence Estimates of Heat-Related Mortality in Seven U.S. Cities. *Environmental health perspectives* **124**, 795–804, <https://doi.org/10.1289/ehp.1509946> (2016).
33. Lemke, B. & Kjellstrom, T. Calculating Workplace WBGT from Meteorological Data: A Tool for Climate Change Assessment. *Ind. Health* **50**, 267–278, <https://doi.org/10.2486/indhealth.ms1352> (2012).
34. Occupational Safety and Health Administration. OSHA Outdoor WBGT Calculator. <https://perma.cc/T6GH-EL3K>. Accessed September, 2019.
35. Dewitz, J. and U.S. Geological Survey. National Land Cover Database (NLCD) all Land Cover Science Products. <https://doi.org/10.5066/P9KZCM54>. Accessed March 11, 2022.
36. Buck, A. L. New Equations for Computing Vapor Pressure and Enhancement Factor. *J. Appl. Meteorol.* **20**, 1527–1532. [https://doi.org/10.1175/1520-0450\(1981\)020<1527:NEFCVP>2.0.CO;2](https://doi.org/10.1175/1520-0450(1981)020<1527:NEFCVP>2.0.CO;2) (1981).
37. Environmental Protection Agency. *Meteorological Monitoring Guidance for Regulatory Modeling Applications*. Report No. EPA-454/R-99-005. <https://perma.cc/2NK4-FLJX> (Office of Air and Radiation, Office of Air Quality Planning and Standards, 2000).
38. Meyra, A. G., Kuz, V. A. & Zarragoicoechea, G. J. Universal behavior of the enthalpy of vaporization: an empirical equation. *Fluid Phase Equilib.* **218**, 205–207, <https://doi.org/10.1016/j.fluid.2003.12.011> (2004).
39. Di Napoli, C., Barnard, C., Prudhomme, C., Cloke, H. L. & Pappenberger, F. ERA5-HEAT: A global gridded historical dataset of human thermal comfort indices from climate reanalysis. *Geoscience data journal* **8**, 2–10, <https://doi.org/10.1002/gdj3.102> (2021).
40. Brimicombe, C. *et al.* Thermofeel: A python thermal comfort indices library. *SoftwareX* **18**, 101005, <https://doi.org/10.1016/j.softx.2022.101005> (2022).
41. Bröde, P. *et al.* Deriving the Operational Procedure for the Universal Thermal Climate Index UTCI. *Int J Biometeorol* **56**, 481–494, <https://doi.org/10.1007/s00484-011-0454-1> (2011).
42. Di Napoli, C., Hogan, R. J. & Pappenberger, F. Mean radiant temperature from global-scale numerical weather prediction models. *Int J Biometeorol* **64**, 1233–1245, <https://doi.org/10.1007/s00484-020-01900-5> (2020).
43. Hogan, R. J. & Hirahara, S. Effect of solar zenith angle specification in models on mean shortwave fluxes and stratospheric temperatures. *Geophysical research letters* **43**, 482–488, <https://doi.org/10.1002/2015GL066868> (2016).
44. Pappenberger, F. *et al.* Global forecasting of thermal health hazards: the skill of probabilistic predictions of the Universal Thermal Climate Index (UTCI). *Int J Biometeorol* **59**, 311–323, <https://doi.org/10.1007/s00484-014-0843-3> (2014).
45. M. Schiavina, S. Freire and K. MacManus. GHS-POP R2019A - GHS Population Grid Multitemporal (1975-1990-2000-2015). <https://data.europa.eu/euodp/en/data/dataset/0c6b9751-a71f-4062-830b-43c9f432370f> Accessed September 3, 2020.
46. Spangler, K. R., Liang, S. & Wellenius, G. A. Daily, County-Level Wet-Bulb Globe Temperature, Universal Thermal Climate Index, and Other Heat Metrics for the Contiguous United States, 2000–2020. *figshare* <https://doi.org/10.6084/m9.figshare.19419836> (2022).
47. Spangler, K. R., Liang, S. & Wellenius, G. A. *UNWEIGHTED* Daily, County-Level Wet-Bulb Globe Temperature, Universal Thermal Climate Index, and Other Heat Metrics for the Contiguous United States, 2000–2020. *figshare* <https://doi.org/10.6084/m9.figshare.19419881> (2022).
48. Spangler, K. R., Liang, S. & Wellenius, G. A. heatmetrics R Package. *figshare* <https://doi.org/10.6084/m9.figshare.19739965> (2022).
49. Lawrence, M. G. The relationship between relative humidity and the dewpoint temperature in moist air - A simple conversion and applications. *B. Am. Meteorol. Soc.* **86**, 225–233, <https://doi.org/10.1175/Bams-86-2-225> (2005).
50. Li, P. W. & Chan, S. T. Application of a weather stress index for alerting the public to stressful weather in Hong Kong. *Meteorol. Appl.* **7**, 369–375, <https://doi.org/10.1017/S1350482700001602> (2000).
51. Smoyer-Tomic, K. E. & Rainham, D. G. Beating the heat: development and evaluation of a Canadian hot weather health-response plan. *Environ. Health Perspect.* **109**, 1241–1248, <https://doi.org/10.1289/ehp.011091241> (2001).

Acknowledgements

This work was supported by grant R01-ES029950 from the US National Institutes of Health/National Institute of Environmental Health Sciences and grant 216033-Z-19-Z from the Wellcome Trust. The funders had no role in study design or collection, analysis, or interpretation of data, writing of the report, or decision to submit the article for publication

Author contributions

K.S. – conceptualization, methodology, software, formal analysis, resources, data curation, writing – original draft; S.L. – software, data curation, writing – review & editing; G.W. – conceptualization, resources, writing – review & editing, supervision, project administration, funding acquisition.

Competing interests

Dr. Wellenius has received consulting income from the Health Effects Institute (Boston, MA) and recently served as a paid visiting scientist at Google LLC (Mountain View, CA). The authors declare that they have no competing conflicts of interest with respect to the content of this manuscript.

Additional information

Correspondence and requests for materials should be addressed to K.R.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022