

SCIENTIFIC DATA

OPEN Data Descriptor: GHWR, a multi-method global heatwave and warm-spell record and toolbox

Ehsan Raei¹, Mohammad Reza Nikoo¹, Amir AghaKouchak², Omid Mazdidasni² & Mojtaba Sadegh³

Received: 15 February 2018

Accepted: 7 August 2018

Published: 30 October 2018

Heatwaves are extended periods of unusually high temperatures with significant societal and environmental impacts. Despite their significance, there is not a generalized definition for heatwaves. In this paper, we introduce a multi-method global heatwave and warm-spell data record and analysis toolbox (named GHWR). In addition to a comprehensive long-term global data record of heatwaves, GHWR allows processing and extracting heatwave records for any location efficiently. We use traditional constant temperature threshold methods, as well as spatially and temporally localized threshold approaches to identify heatwaves. GHWR includes binary (0/1) occurrence records of heatwaves/warm-spells, and annual summary files with detailed information on their frequency, duration, magnitude and amplitude. GHWR also introduces the standardized heat index (SHI) as a generalized statistical metric to identify heatwave/warm-spells. SHI has direct association with the probability distribution function of long-term daily temperatures for any given calendar day and spatial grid. Finally, GHWR offers a unique opportunity for users to select the type of heatwave/warm-spell information from a plethora of methods based on their needs and applications.

Design Type(s)	source-based data analysis objective • protocol optimization design
Measurement Type(s)	increased temperature
Technology Type(s)	data item extraction from journal article
Factor Type(s)	
Sample Characteristic(s)	Earth (Planet) • temperature of environmental material

¹Shiraz University, Shiraz, Iran. ²University of California, Irvine, CA, USA. ³Boise State University, Boise, ID, USA. Correspondence and requests for materials should be addressed to M.S. (email: mojtabasadeqh@boisestate.edu)

Background & Summary

Heatwaves are generally defined as prolonged periods of high temperature and excess heat¹ with significant impacts on humans and natural ecosystems. In the recent decades, heatwaves have become one of the deadliest natural hazards in some regions^{2,3}, and are recognized as “*silent killers*”⁴. Heatwaves leave marked societal impacts ranging from damaging infrastructure, stressing emergency response units and hospitals, and overburdening the electricity grids, to paralyzing the railroad and air transportation systems^{2,5–7}. The extent of such impacts depends largely on the societal preparedness to respond to the extreme events^{3,8}. Other negative impacts of heatwaves and warm-spells include reducing work capacity and labor productivity⁹, crop yield^{9,10}, and livestock productivity².

A comprehensive analysis of heatwaves should consider four key characteristics of such events, namely: 1. intensity, 2. duration, 3. frequency, and 4. spatial extent^{1,2,11–13}. In this study, we create a comprehensive multi-method global heatwave and warm-spell record and toolbox, GHWR, based on different methodologies in the literature. Here, a warm-spell corresponds to unusually warm temperatures even if they are not extremely high (e.g., a 5-day event in winter with an average temperature of +5 °C in a region where the expected average temperature based on long-term climatology is –5 °C). GHWR can be used to analyze changes in statistics of heatwaves and warm-spells around the world, and to quantify their impacts on different sectors and target groups.

Heatwave and warm-spell identification methodologies in this data set include five general categories: 1. constant temperature threshold approach, 2. location and calendar day specific temperature threshold (PDF-derived threshold approach), 3. location specific threshold based on upper tail of summer temperature distribution (summer-derived threshold approach), 4. excess heat factor (EHF) approach, and 5. the herein introduced standardized heat index (SHI) method. Heatwaves are complex in nature and cannot be simply compared across regions, natural systems, and target sectors¹⁴. The idea behind SHI is to provide standardized and spatially consistent heatwave and warm-spell information. SHI and EHF methods employ average daily temperature, while other approaches are based on daily max, min or average temperature.

In this data record, we cover a wide range of temperature thresholds to address different potential applications. The threshold approaches (items 1–3), consider 2, 3, 4, 5, 6, 7 and 10 consecutive days as required temporal persistence for heatwave and warm-spells. The GHWR data set is generated using the CPC global air temperature data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (<http://www.esrl.noaa.gov/psd/>). Gridded air temperature at the surface is provided for the entire globe with a spatial resolution of 0.5° and a daily temporal resolution for 1979 to present. GHWR incorporates all the required characteristics of heatwaves and warm-spells (intensity-duration-frequency-spatial extent) in one easily accessible depository, and can be used in different studies ranging from human mortality and morbidity, impacts on the energy and agriculture sectors, and also spatiotemporal statistical analysis of heatwaves and warm-spells. It is noteworthy that GHWR (toolbox) relaxes the dependency of GHWR (data set) on the source CPC temperature data, and can perform a comprehensive heatwave/warm-spell analysis on any gridded daily temperature source data for any location.

Existing heatwave datasets mainly focus on fixed-threshold definitions emanated from community efforts such as the Expert Team on Climate Change Detection and Indices (ETCCDI)^{15,16}. Hadley Extremes database (HadEX)¹⁷ (<http://www.metoffice.gov.uk/hadobs/hadex2/>), for example, provides coarse resolution ETCCDI indices on a 2.5° × 3.75° grid from 1901 to 2010. Hadley Centre/Global Historical Climatology Network (HadGHCND) extremes database (GHCNDEX) compute such indices based on ground-based station data^{18,19}. However, over a relatively short period of time, significant strides have been made in characterizing heatwaves and understanding their impacts¹, some of which call for or propose localized and sector-specific thresholds for identifying heatwaves^{1,2,14,20–22}. GHWR leverages the recent contributions in the literature, as well as the ETCCDI indices, to create a multi-method global heatwave and warm-spell record based on gridded 0.5° × 0.5° temperature data from 1979 to 2017. In more detail, GHWR leverages a plethora of different methods (discussed in Methods) with various settings (see Table 1) to enable future users to compare different heatwave definitions and select the most proper method for the target study.

Methods

A heatwave is generally defined as temperatures rising above a certain threshold (either percentile-based or fixed value) for a number of consecutive days. Despite the strides made in the field, there is no universal consensus for defining heatwaves⁴. A plethora of definitions are available in the literature²³ each focusing on a particular application or target group (e.g. human morbidity and mortality, infrastructure and transportation systems, wildfire, and agriculture sector)^{14,24}. However, all definitions of heatwaves involve at least one form of temperature (daily max, min, or average) and require a predefined threshold to be exceeded for a number of consecutive days¹. The specification of threshold is of highest significance, and is dependent on the application or the target group of study (e.g., plants, humans, energy sector).

Even a slight change in heatwave definition might lead to considerable effects on the estimated impacts²⁵. Some animals are sensitive to a specific range of temperatures, and some agricultural products show a direct response to a particular temperature threshold²⁶. For example, flying foxes (native to Australia) may die from dehydration at air temperatures of above 40 °C, especially when relative

Method	Temperature type	Temperature threshold (°C)	Temporal persistence (days)	Other	Localized threshold?
Constant Threshold	TMAX TMIN TMEAN	35, 36, ..., 45 25, 26, ..., 35 30, 31, ..., 40	2, 3, 4, 5, 6, 7, and 10	—	None
Average temperature + 5 °C	TMAX TMIN TMEAN	Average annual temperature + 5 °C	2, 3, 4, 5, 6, 7 and 10	—	Spatial only
Upper tail percentile	TMAX TMIN TMEAN	70, 75, 80, 85, 90, 95 th percentile of PDF from: • 15 day window centered at calendar day • 21 day window centered at calendar day	2, 3, 4, 5, 6, 7 and 10	—	Temporal and spatial
Summer-derived threshold	TMAX TMIN TMEAN	70, 75, 80, 85, 90, 95 th percentile of PDF from: • June-August • June-September • May-August • May-September	2, 3, 4, 5, 6, 7 and 10	—	Spatial only
EHF	TMEAN	-	3-day average	EHF ≥ 0 EHF ≥ 1	Temporal and spatial
SHI	TMEAN	-	3-day average	SHI ≥ 1	Temporal and spatial

Table 1. Definitions of heatwave/warm-spells used in global heatwave and warm-spell record.

humidity is low². Human's core temperature is 37 °C, on average, and a rise to 38 °C for several hours may lead to heat exhaustion and may impair physical and mental capacities^{27,28}. The temperature threshold for heatwaves may also be selected based on the climatology of a region³, such as the 90th percentile of the local daily temperature distribution, since humans and the natural ecosystem acclimate to the local climate²⁵. Moreover, temperature thresholds could/should differ for individual calendar days. For example, in a region with strong seasonality, identifying warm-spells in winter requires a much lower temperature threshold than that in summer¹.

In the following, we provide a comprehensive, but non-exhaustive, list of definitions from the literature, and introduce a new standardized and probabilistic definition. The below description focuses mainly on the heatwave definitions used in GHWR (also see Table 1). The user should note that GHWR provides a plethora of different methods to detect heatwaves and warm-spells, but not all methods are necessarily appropriate for all regions and or applications. The user should carefully select the proper method based on the region and application of interest. However, a unique feature of GHWR is indeed enabling users to contrast and compare different approaches.

Constant temperature threshold approach

This approach defines a heatwave as an event during which temperatures exceed a constant threshold for an extended period of time^{25,29–35}. In GHWR, the extended period of time is considered 2, 3, 4, 5, 6, 7 and 10 consecutive days during which daily max, min, or average temperature surpasses a wide range of thresholds (35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45 °C for daily maximum temperature, TMAX; 25, 26, 25, 28, 29, 30, 31, 32, 33, 34, 35 °C for daily minimum temperature, TMIN; and 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 °C for daily average temperature, TMEAN). In other words, GHWR provides different global heatwave records using different combinations of temperature threshold and duration for daily max, min and average temperature observations (e.g., all 5-day heatwave events with daily maximum temperature greater than 35 °C).

PDF-derived temperature threshold

Temperature thresholds for heatwaves, in this approach, are assigned according to the local climatology. For example, events during which temperature observations exceed a localized threshold (e.g., 90th percentile) for several consecutive days are defined as heatwaves^{36–38}. GHWR provides heatwave and warm-spell information based on different PDF-derived thresholds using daily max, min and average temperatures.

Average temperature+5degC

This approach considers a spatially localized temperature threshold to delineate heatwave days from non-heatwave ones^{1,39}. This method only considers daily average temperatures with a temperature threshold that is estimated as the mean of long-term daily average temperature observations + 5 °C. Note that this is an outdated approach that neglects the relative climatological features of the region, and might not be applicable everywhere. As a matter of fact, in some regions temperature might not even get 5 °C above average climatological temperature⁴⁰. However, we have included it for cross comparison of different methods.

Upper tail percentile

This method localizes temperature threshold both spatially and temporally. In more detail, temperature thresholds for each calendar day and each grid are specified as an upper tail threshold (75th, 80th, 85th, 90th, 95th percentiles) of the probability distribution function (PDF) constructed from long-term daily temperature record over a window of 15 (21) days centered around the calendar day of interest^{11,12,41–43}.

Temporal localization in this approach enables detection of warm-spells (e.g., unusually warm winter) as well as heatwaves.

Summer-derived threshold approach

This method spatially localizes temperature threshold for heatwaves, and assigns an upper tail threshold (75th, 80th, 85th, 90th, 95th percentiles) of the PDF constructed from long-term daily summer (warm season) temperatures in each grid^{3,44}. This approach is conceptually similar to the PDF-derived threshold but only uses data from summer (warm season), and does not consider temporally localized thresholds. This definition is more appropriate for studies that focus on extreme heat and their impacts on mortality, crop yield, etc.³. GHWR provides heatwave information based on this definition using daily max, min and average observed temperatures, and different duration (2, 3, 4, 5, 6, 7 and 10 consecutive days).

Excess heat factor (EHF)

Excess heat factor (EHF) combines measures of “*excess heat*” and “*heat stress*” to develop an index that can effectively describe the impacts of heatwaves on human mortality and morbidity^{14,21,41}. Excess heat characterizes localized long-term temperature anomaly compared to climatology of the region (grid), whereas heat stress represents short-term temperature anomalies to address thermal acclimation²². EHF is constructed based on a running three-day-averaged TMEAN (daily average temperature), as vulnerable population show highest sensitivity to heatwaves after three days of exposure⁴¹. EHF is specifically designed to capture heatwave severity²².

High day-time heat accumulates if night-time temperatures are high as well. Average of daily maximum and minimum temperature over a three day running period can then represent the accumulated excess heat, when compared to a long-term climatological baseline (reference)⁴¹. The long-term climate reference is set to 95th percentile of the PDF of daily average temperature (hereafter, T_{95}). Positive and consecutive anomaly of a three-day-averaged mean temperature from this reference value is indicative of a heatwave event. Excess heat index is formulated as,

$$EHI_{sig} = (T_i + T_{i+1} + T_{i+2})/3 - T_{95}, \quad (1)$$

in which, T_i denotes mean daily temperature at day i . Heat stress, reflecting short-term acclimation, is then defined as,

$$EHI_{accl} = (T_i + T_{i+1} + T_{i+2})/3 - (T_{i-1} + T_{i-2} + \dots + T_{i-30})/30. \quad (2)$$

Finally, EHF aggregates the impacts of EHI_{sig} and EHI_{accl} to develop a comparative index representing intensity, load, duration and spatial distribution of a heatwave event⁴¹,

$$EHF = EHI_{sig} \times \max(EHI_{accl}, 1). \quad (3)$$

In GHWR, we consider $EHF \geq 0^\circ\text{C}^2$ and $EHF \geq 1^\circ\text{C}^2$ as thresholds for heatwave.

Standardized heat index (SHI)

We introduce standardized heat index (SHI) as a probabilistic and generalized index to describe heatwaves according to their probability of occurrence in long-term records. SHI is inspired by the standardized precipitation index (SPI)⁴⁵ and is spatially and temporally localized. For each grid (region), a temperature PDF is constructed for a window of 15 (and 21) days centered around the calendar day of interest, and the empirical probability for each temperature level is calculated using the Gringorten plotting position⁴⁶

$$p(T_i) = \frac{r_i - 0.44}{n + 0.12}, \quad (4)$$

in which T_i is the i^{th} observed mean daily temperature in the constructed PDF, and r_i is the rank of T_i from the smallest value⁴⁷. SHI for temperature level \hat{T} is then estimated as

$$SHI(\hat{T}) = \Phi^{-1}(p(\hat{T})), \quad (5)$$

where Φ signifies the standard normal distribution. This approach does not require the data to follow normal distribution, and the last step is only used to describe all heatwave information in a standardized scale (e.g., standard normal distribution).

In order to address the persistence of excess heat for an extended period of time, and similar to EHF, we consider three-day-averaged mean daily temperature ($\hat{T} = (T_i + T_{i-1} + T_{i-2})/3$) to develop the SHI index (Eq. 5). We adopt a threshold of $SHI \geq 1$ for identifying heatwaves, which is associated with three-day-averaged mean daily temperature being greater than the $\sim 84^{\text{th}}$ percentile of the long-term temperature PDF ($\Phi(1) = 0.8416$). A threshold of 1 is selected herein to be synergistic with the SPI index (see National Drought Monitor for definition of dry and wet years according to SPI). The user can select other thresholds such as 1.28, 1.5, or 2 which coincide with 90th, 93.3th and 97.7th percentiles, respectively. Given the statistical basis of this index, SHI can be used to identify both coldwaves (cold-spells) and heatwaves (warm-spells), by tuning the threshold. The user can select thresholds that are more appropriate for their own studies.

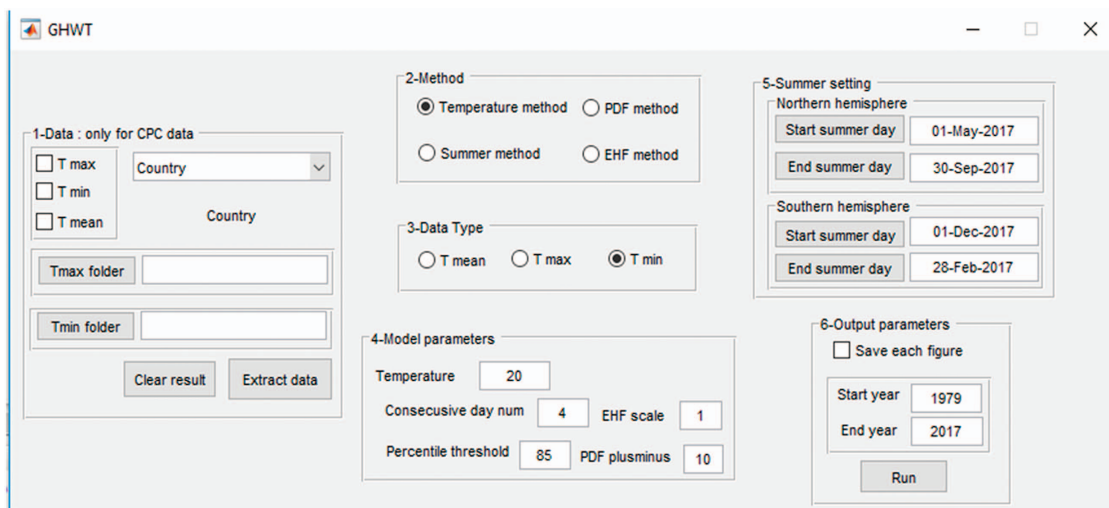


Figure 1. Graphical user interface (GUI) of the global heatwave and warm-spell toolbox, GHWT.

Code Availability

The GHWR source codes are developed in MATLAB and are available to public along with the GHWR data set (under “GHWR_Codes” folder). The source codes are developed based on the CPC gridded global daily temperature data set, but the user can readily execute these codes to generate heatwave record based on any gridded daily source data, and selection of heatwave method and its parameters. The user can alternatively use the global heatwave and warm-spell toolbox, GHWT, as explained in the next section.

Global heatwave and warm-spell toolbox, GHWT

GHWT is a user-friendly and flexible toolbox developed in MATLAB to construct long-term heatwave and warm-spell record with any gridded global daily temperature data source for any country/continent around the globe. The user can select to perform computations using either the graphical user interface (GUI; execute “GHWT.m” in MATLAB) or the “run script” (modify and execute “GHWT_Script_run.m” in MATLAB). Fig. 1 displays the GHWT toolbox GUI. First step is to select temperature data directory and the country (or continent) of interest. This first step is designed for the CPC temperature data source. If the user opts to employ any other source for temperature data, (s)he should first modify the “config.txt” file and execute “extract_data.m” under the “extract other database format” subfolder to pre-process the source data. Subsequent steps can all be performed in the GUI. Next is to select the heatwave/warm-spell identification method, and selection of data type (max, min, or average daily temperature). Then, the user should define the method-specific parameters of the analysis (such as the number of consecutive days for the temperature to be above a certain threshold). If the user selects the “Summer method”, (s)he needs to define the start and end date of summer. The user subsequently selects the start and end year of analysis and execute the program. The GHWT toolbox plots several figures displaying the number of heatwave days and events, first and last heatwave day of year, length of longest heatwave, and magnitude and amplitude of heatwaves for each year. These information will also be saved under “GHWT_results.mat” under the “result” folder.

GHWT is publicly available at https://github.com/mojtabasadeh/Global_Heatwave_and_Warm_Spell_Toolbox under Attribution-ShareAlike 3.0 license (<https://creativecommons.org/licenses/by-sa/3.0/>). Country borders shapefiles used in this toolbox were obtained from <http://thematicmapping.org/>.

Data Records

The global heatwave and warm-spell record (GHWR) is freely available to public at (Data Citation 1).

The GHWR record provides information on heatwaves including intensity, duration, frequency, and spatial extent. The gridded global heatwave record is provided as binary (0/1) and annual summary for each heatwave definition in “netcdf” format. The binary form will categorize each day as 1 if identified as heatwave/warm-spell, and 0 otherwise. The annual summary files for each definition provides detailed information about heatwave/warm-spell including: a. number of days, b. number of events, c. first heatwave occurrence day of year, d. last heatwave occurrence day of year, e. length of longest event, f. length of shortest event, g. magnitude (average of the mean heatwave/warm-spell temperatures of all events), h. amplitude (peak daily temperature in the hottest heatwave as defined by the heatwave with highest magnitude). It is also important to mention that for SHI and EHF methods, we also provide time series of EHF and SHI record for each grid over 1979–2017. Moreover, temporally and spatially localized heatwave/warm-spell thresholds for each method is included in GHWR for future and further analysis. A

tabulated description of heatwave/warm-spell definitions that are used in the GHWR data set is provided in Table 1.

It is noteworthy that all files are compressed to “7z” format to reduce size, and proper MATLAB and Python codes are provided to automatically decompress the selected files to “netcdf” format. The GHWR data set is provided in three general folders namely: Record, Summary, and Threshold; which in turn are divided into two parts (e.g. Record1 and Record2) for facilitating upload to data repository. Each of these general folders then include heatwave information from all methods and possible combinations of parameters as presented in Table 1. A “Data Directory” (“pplx” format) provides guidance to find any specific file. A “readme.txt” file is also available along with the data set that explains the decompressing process.

Technical Validation

We analyze three historical heatwave events from around the globe that led to high societal impacts using the GHWR data set. We provide evidence that the different methodologies employed to create the data set not only capture the intensity and duration of these events, but also provide important insights about the spatial extent of these events. Moreover, GHWR’s length of record (1979–2017) enables users to analyze frequency of a heatwave or warm-spell event and, placing it in a climatological perspective. GHWR also provides temporally and spatially localized heatwave/warm-spell temperature thresholds for different methods, which can in turn be used for identification of future heatwave events given temperature predictions.

Chicago heatwave (July 1995)

The heatwave of July 1995 that affected most of the Midwest US, were unprecedented in many senses. Extreme day- and night-time temperatures persisted over an extended 48-hour period, was exacerbated by high relative humidity, and left remarkable societal impacts in Chicago, IL⁴⁸. Due to the excessive heat and humidity, hospital admission increased 35% for the elderly (65 years and older)²⁴. Multiple studies^{49,50} attributed at least 600 excess deaths to the extreme heat during and right after this event. It is noteworthy that while extreme day-time apparent temperature (an index of the combined effect of temperature and humidity on humans) in the Chicago heatwave is not particularly exceptional, the persistence of high night-time apparent temperature was unprecedented in the modern history (with probability of occurrence of less than 0.1%,⁴⁸). Night-time lower temperatures are necessary to relieve the day-time excess heat, absence of which significantly exacerbate the impacts of heatwaves².

Figure 2 shows the frequency of heatwave days in July of 1995 in the US (with a focus on Chicago, IL), based on different definitions/methods. Figure 2a displays frequency of heatwave days defined based on the maximum daily temperatures above 35 °C with a duration of at least four consecutive days. While Midwest US, and specifically Chicago, experienced around 5 heatwave days in July of 1995 according to this definition, heatwave days were much more frequent in South and Southwest US. Indeed, the entire month of July, 1995 (31 days) is categorized as heatwave days for some regions in the west of Arizona and southeast of California in this category. This highlights the need to define a local heatwave threshold for each region. Figure 2b displays frequency of heatwave days in July, 1995 for the entire US with a temperature threshold of 90th percentile of the long-term TMAX distribution for each grid and each calendar day, with a temporal persistence of at least four days. It is noticeable that the geographical extent of the Midwest US heatwave expands to north of Chicago in this definition (compare to Fig. 2a). Moreover, frequency of heatwave days in the south and southeast US also significantly decreases in this definition. However, still some parts of Arizona show more frequent heatwave days compared to the Midwest US where the Chicago heatwave had left more significant impacts. This is due to the persistence of high day- and night-time temperature in the Midwest US. Figure 2c shows frequency of heatwave days with minimum (night-time) temperature above 90th percentile of the long-term TMIN distribution for at least four successive days. Indeed, persistence of high day- (Fig. 2b) and night-time (Fig. 2c) temperatures explain the high impact of this event. Figure 2 clearly demonstrates the advantages of defining local thresholds for heatwaves compared to that of the constant temperature threshold approach¹.

Note that the choice of heatwave duration, as well as other heatwave parameters, is merely to demonstrate different features of GHWR. The user can select any heatwave setting that fits the target study from hundreds of combinations available in GHWR.

European heatwave (summer 2003)

The mega-heatwave of July and August, 2003 in Europe was the hottest period that western Europe has experienced since at least 1500 AD⁵¹. Driven by the anticyclonic (high atmospheric pressure) conditions and unusual soil moisture deficit⁵², the 2003 European heatwave claimed more than 70,000 lives, hitting France the hardest^{53,54}. Similar to other deadly heatwaves⁵⁵, night-time temperature during the European 2003 heatwave were significantly higher than normal⁵⁶. Recent studies have provided considerable evidence that the rate of increase in night-time temperature is even higher than its day-time counterpart, which exacerbates the negative societal and environmental impacts of such events².

Figure 3a displays the frequency of heatwave days over July and August, 2003 in western Europe, as defined by minimum daily temperature exceeding 90th percentile of long-term distribution of TMIN for each day and each grid for at least four consecutive days. Some locations, interestingly, endure around 45

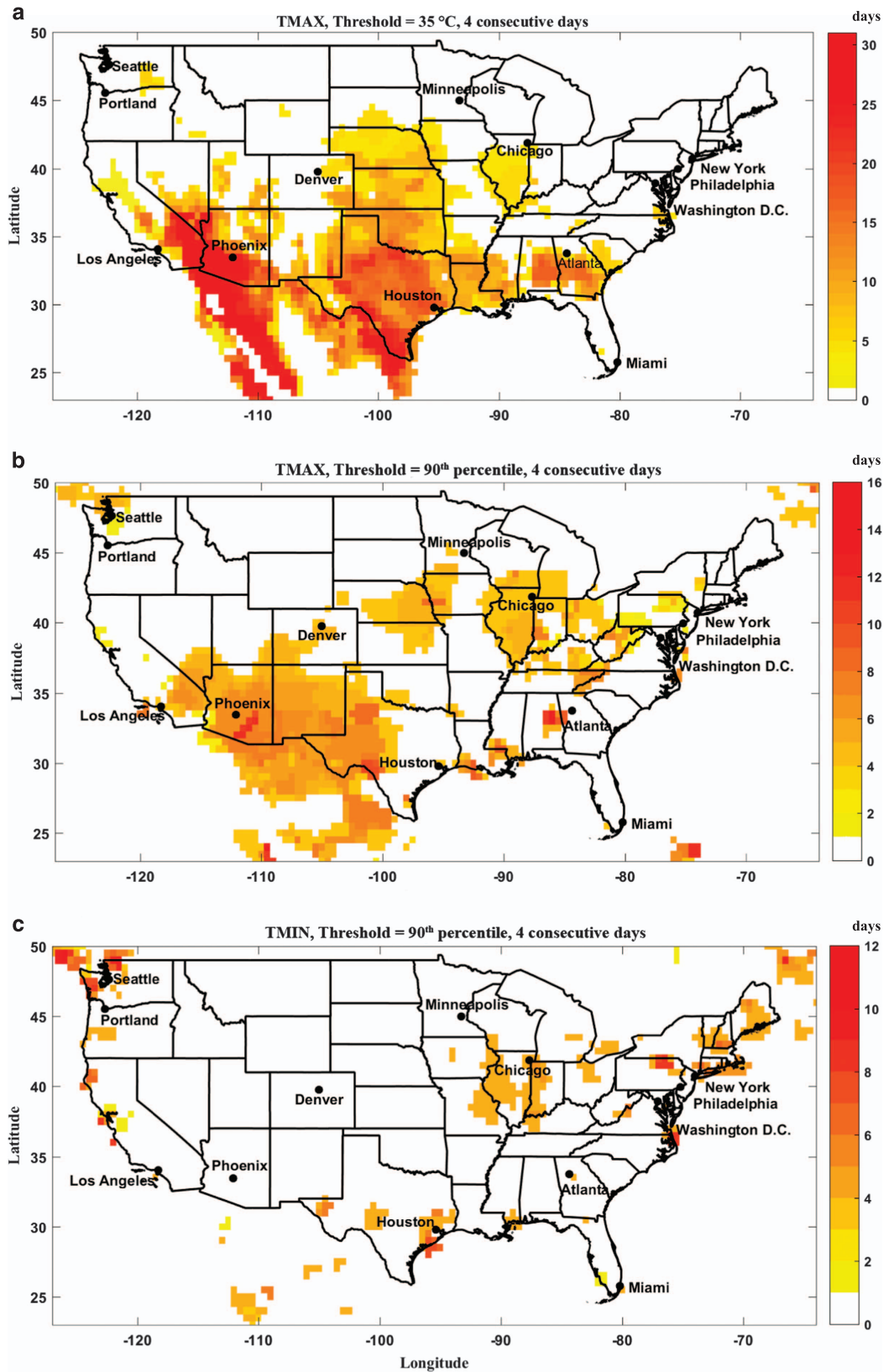


Figure 2. Frequency of heatwave days in the month of July, 1995 using different heatwave definitions. (a) Daily maximum temperature above 35 °C, (b) Daily maximum temperature above 90th percentile of the long-term TMAX climatology, and (c) Daily minimum temperature above 90th percentile of the long-term TMIN climatology, for at least four consecutive days.

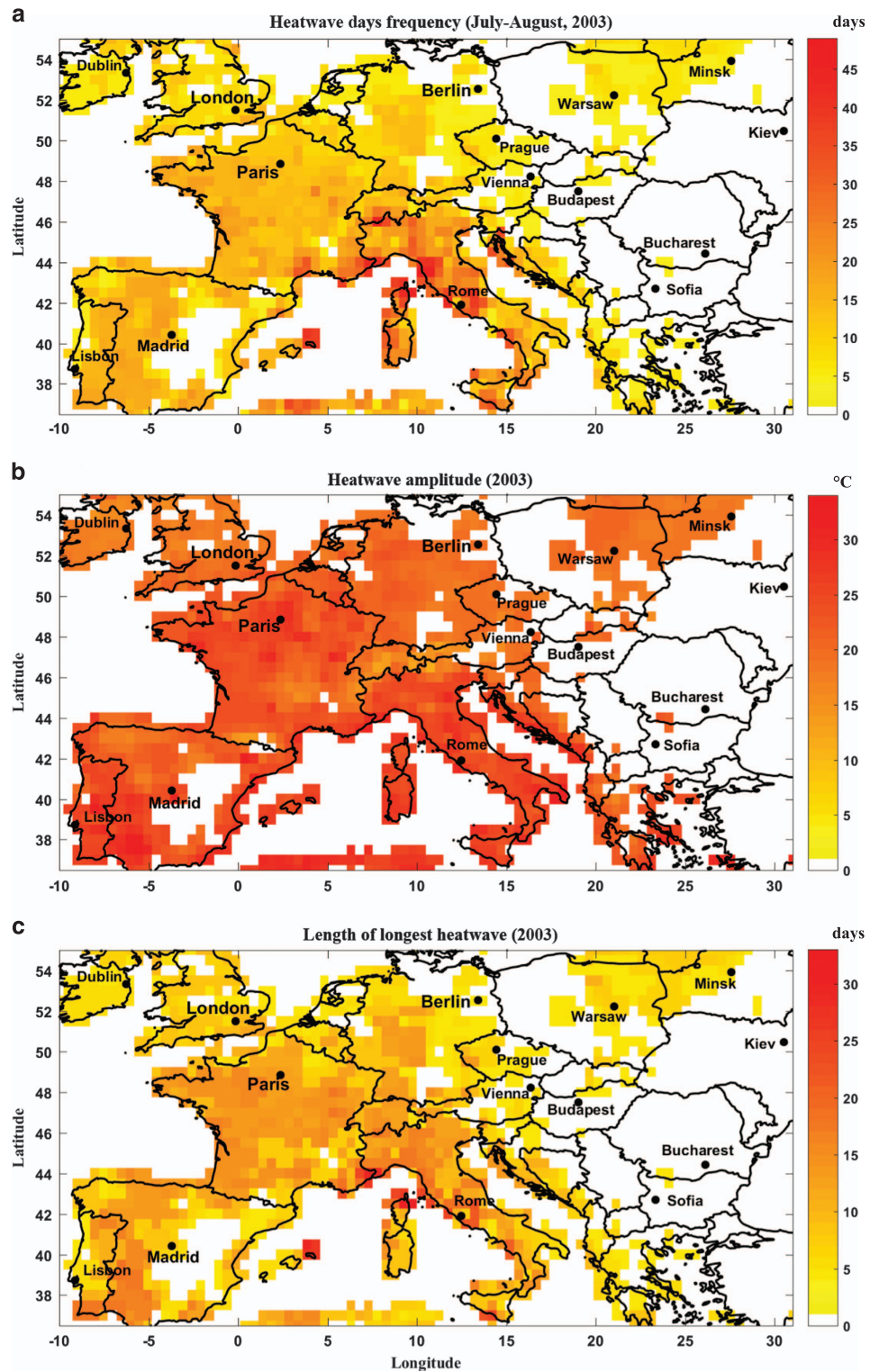


Figure 3. The 2003 European heatwave characteristics. (a) Frequency of heatwave days in July-August, 2003, (b) Heatwave amplitude (based on TMIN) in 2003, and (c) Length of the longest heatwave in 2003. Heatwave is defined as persistence of daily minimum temperature above 90th percentile of TMIN climatology for at least four consecutive days.

heatwave days in a two months period. Following a similar heatwave definition, Fig. 3b displays the heatwave amplitude for daily minimum temperature (as defined by the peak daily TMIN in the hottest heatwave), which exceeds a night-time temperature of 30 °C in an extended region around Paris, France. Finally, Fig. 3c shows length of the longest heatwave event in 2003. Similarly, heatwaves are defined as

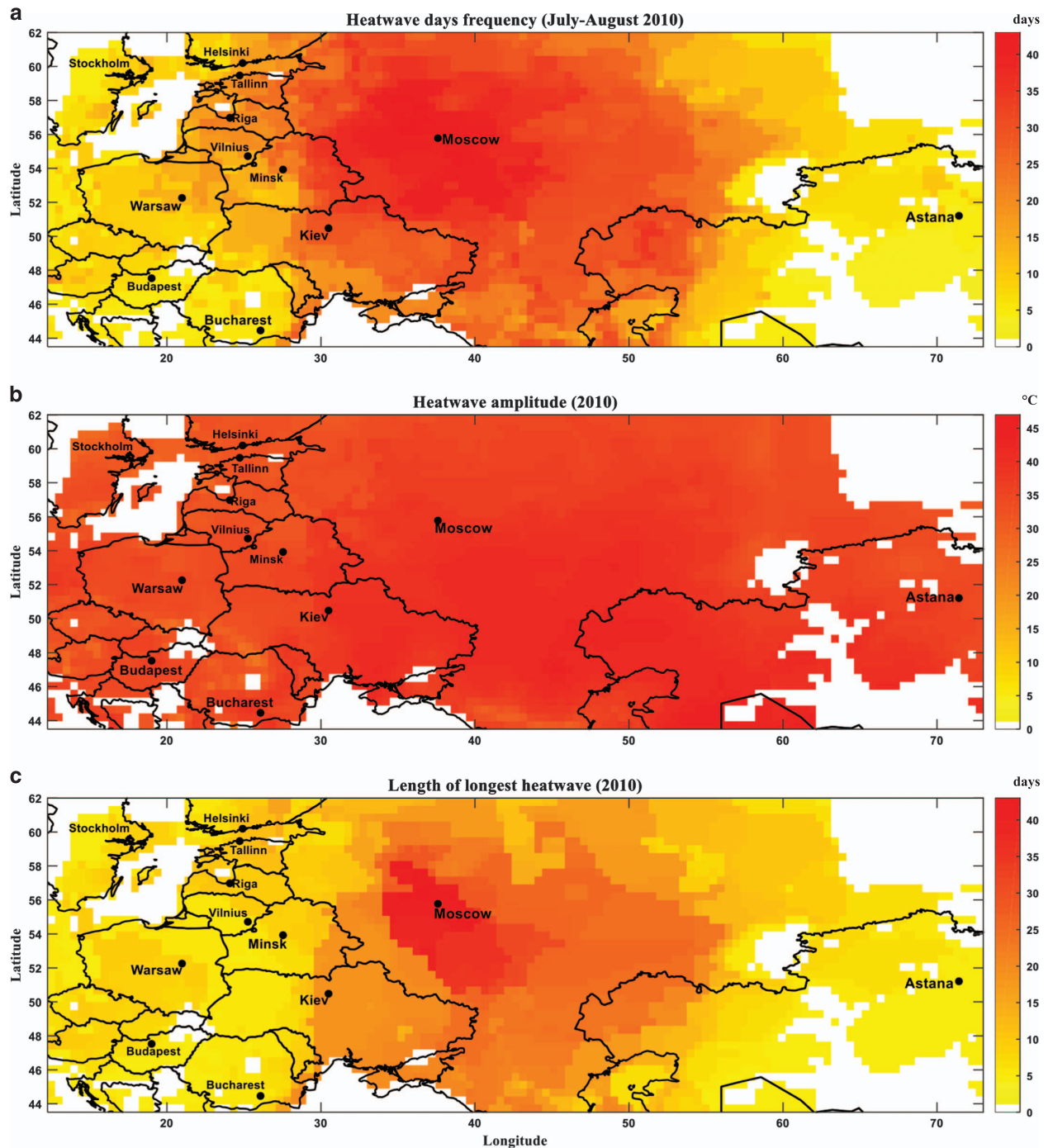


Figure 4. The 2010 Russian heatwave characteristics. (a) Frequency of heatwave days in July-August, 2010, (b) Heatwave amplitude (based on TMAX) in 2010, and (c) Length of longest heatwave in 2010. Heatwave is defined as persistence of daily maximum temperature above 90th percentile of TMAX climatology for at least four consecutive days.

those with daily minimum temperature exceeding 90th percentile of TMIN climatology for at least four days. Most locations in Europe experience around 10 consecutive heatwave days in 2003, which extends to 30+ days in some regions. The spatiotemporal extent of this heatwave was indeed unprecedented, only to be surpassed by the Russian 2010 heatwave.

Russian heatwave (summer 2010)

The Russian heatwave in summer of 2010 was exceptional in many senses. It exceeded the amplitude and spatial extent of the European 2003 heatwave¹², and persisted for over one month⁵⁷. Indeed, July 2010 was by far the warmest on record in western Russia⁵⁸. Similar to the European heatwave of 2003, the

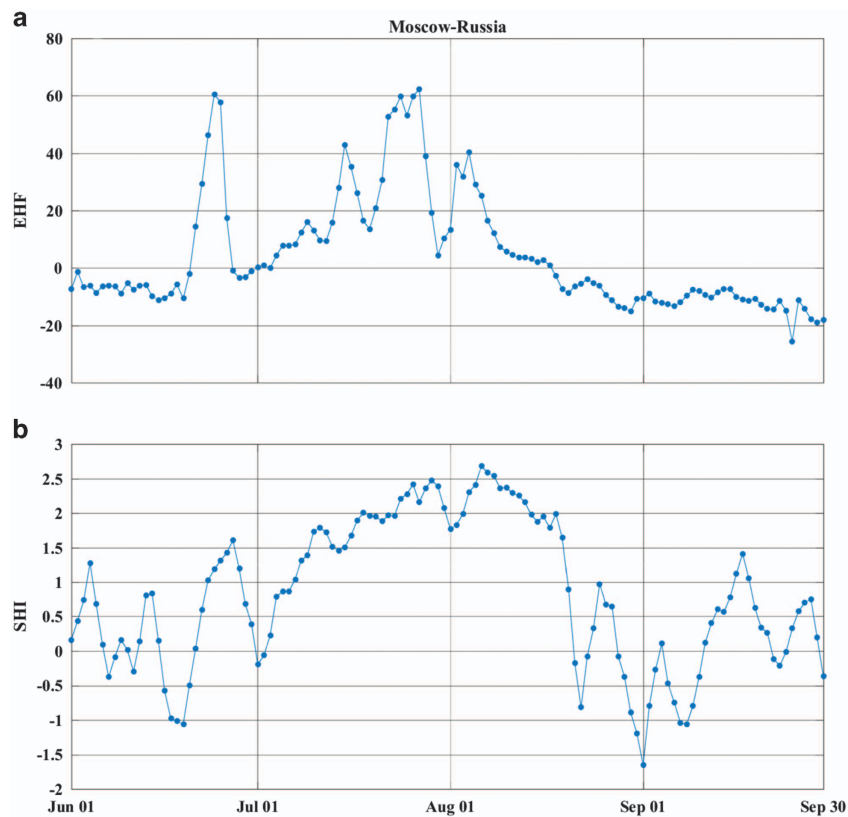


Figure 5. Heatwave intensity for Moscow, Russia, during June-September 2010. (a) Time series of excess heat factor (EHF), and (b) Time series of standardized heat index (SHI).

Russian 2010 heatwave was driven by a blocking anticyclone and exacerbated by the soil moisture-temperature feedback⁵⁹. The soil moisture deficit in western Russia in 2010 alone increased the odds of a severe heatwave sixfold⁵⁹. This event broke several records, and daily maximum temperatures persisted around the record levels for an extended period of time⁶⁰. This mega heatwave claimed some 56,000 lives⁵⁷, left negative impacts on the agriculture¹, and contributed to 500 wildfires¹, which in turn provoked air pollution⁵².

Figure 4a displays frequency of heatwave days in western Russia in July and August, 2010. Heatwave is defined by daily maximum temperature exceeding 90th percentile of the long-term TMAX climatology for at least four consecutive days for each location and calendar day. A large region surrounding Moscow, Russia, experienced 40+ heatwave days in two months. Figure 4b shows heatwave amplitude for daily maximum temperature soaring 45+ °C for most of western Russia. Finally, length of the longest heatwave event is portrayed in Fig. 4c, showing a continuous 40+ heatwave days in a vast area around Moscow. Temporal extension of heatwaves exacerbates the heatwave impacts by restricting relief periods needed for the human and environmental systems to recover.

Figure 5 displays daily time series of EHF (a) and SHI (b) values for Moscow, Russia, during June-September, 2010. Both approaches capture the intensity and temporal extension of the 2010 Russian heatwave over Moscow, Russia. However, they show different behavior in late July and mid September. In late July as average daily temperature drops from ~32 °C to ~25 °C, EHF drastically reduces from ~60 °C² to ~4 °C², whereas SHI shows a more moderate drop from ~2.5 (~99th percentile) to ~1.75 (~96th percentile). The different response of these two metrics is rooted in their definitions, as EHF considers acclimation to recent weather conditions (mean daily temperature of the previous 30 days) whereas SHI does not include such an effect. SHI is closer in definition to that of the traditional heatwave literature. SHI and EHF also show a different behavior in mid September, when SHI shows a heatwave/warm-spell event whereas EHF does not. Indeed, in mid September average daily temperature rises to ~17 °C which equates to the ~92th percentile of the TMEAN climatology (15 day window centered at September 15). This shows that September 15, 2010 is unusually hot for its historical record, but since EHF is only focused on human mortality and morbidity, it would not recognize this event as warm-spell. A deeper discussion is required to recognize the similarities and differences between SHI and EHF, which is beyond the scope of this paper. It is noteworthy that SHI is a probabilistic measure with clear association with the temperature distribution, and can readily be interpreted as different percentiles of the

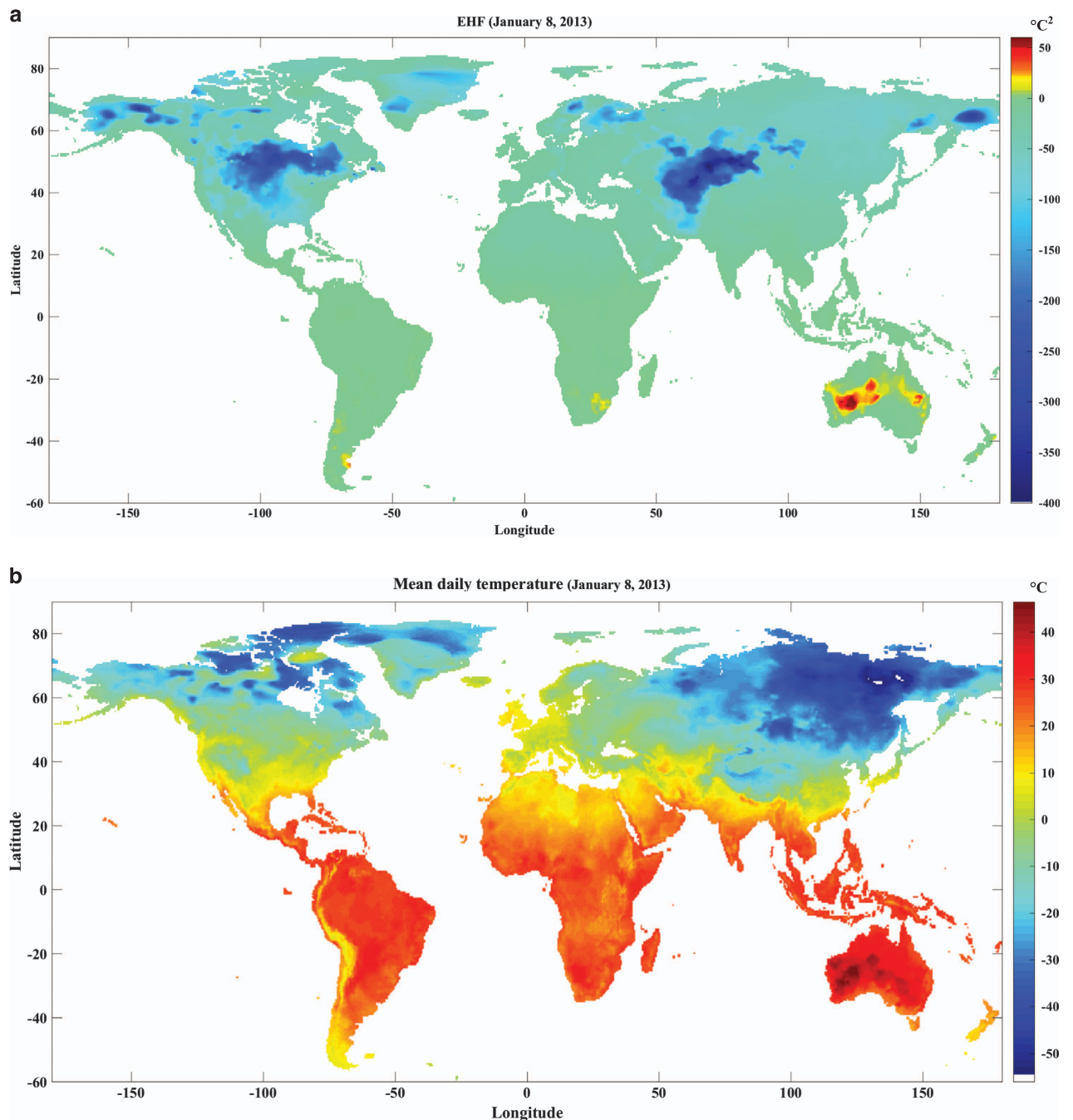


Figure 6. Global temperature and heatwave situation on January 8, 2013. (a) Excess heat factor (EHF), and (b) Mean daily temperature.

long-term mean daily temperature PDF. Moreover, SHI can also identify cold-spells as they fall in the lower tail of the long-term TMEAN PDF.

Global heatwave/warm-spell maps

Figure 6a displays the EHF record for January 8, 2013 over the entire globe. January 2013 coincides with the “angry summer” in Australia, and was the hottest months on record in Australia. “Angry summer” broke the all-time TMAX record for several dozen stations across Australia⁶¹. EHF correctly captures the severity of the “angry summer” heatwave on January 8, 2013 for western and central Australia (EHF reaching 50 $^{\circ}\text{C}^2$). The mean daily temperatures soar above 45 $^{\circ}\text{C}$ on this day (Fig. 6b) for many parts of Australia, specially in the western and central regions.

Figure 7a shows SHI record for June 20, 2015 over the entire globe. Karachi, Pakistan, experienced an extreme heatwave in June, 2015, which along with power outage and water distribution system failure,

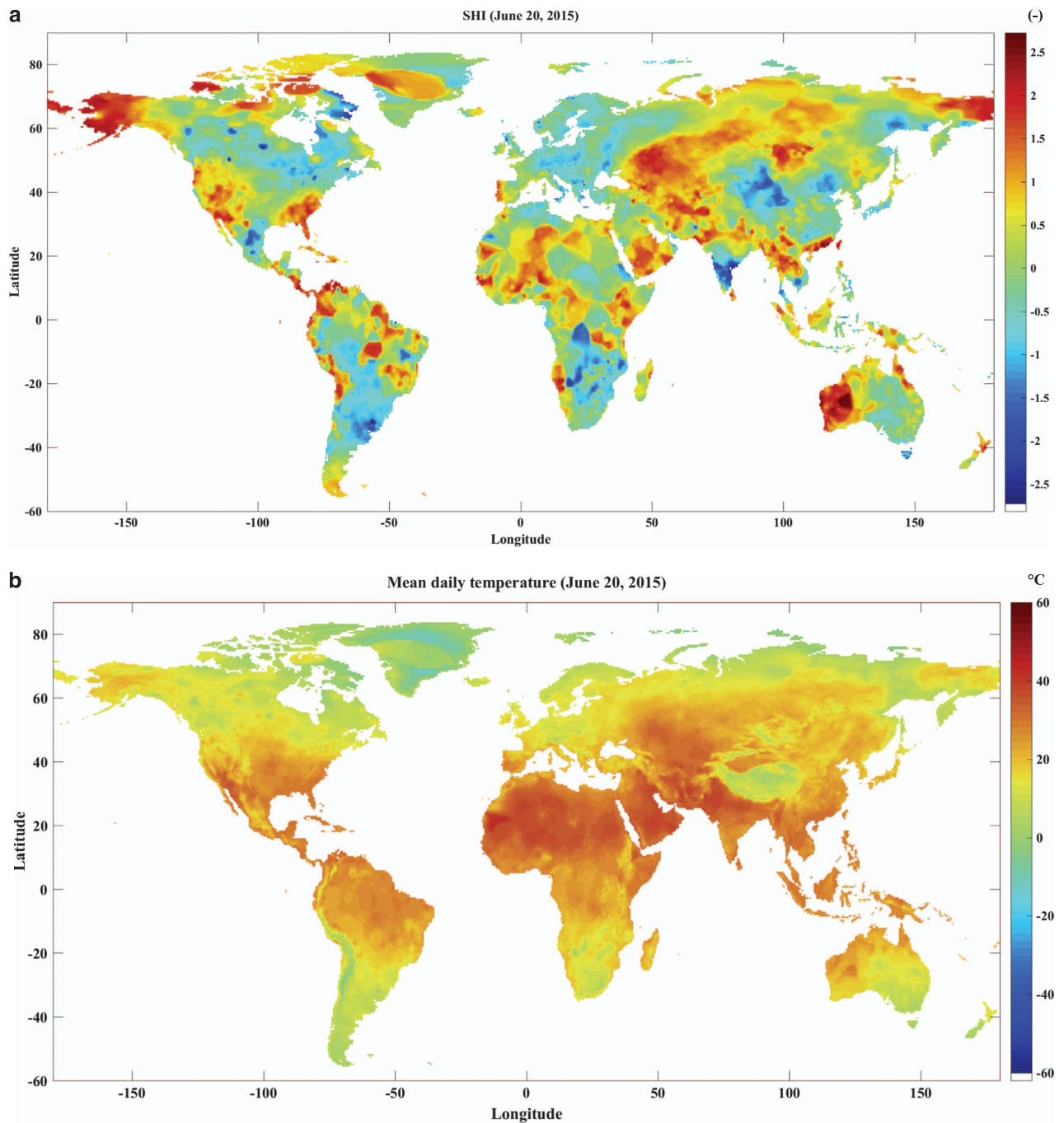


Figure 7. Global temperature and heatwave situation on June 20, 2015. (a) Standardized heat index (SHI), and (b) Mean daily temperature.

imposed a death toll of 700+ people⁶². This figure captures the Karachi's heatwave with SHI > 2 (daily mean temperature above 97th percentile of TMEAN climatology). Figure 7a also shows heatwaves over many regions in the northern hemisphere and warm-spells in the southern hemisphere.

Usage Notes

In this paper, we introduce a multi-method global heatwave and warm-spell record, GHWR. We leverage different definitions of heatwaves from the literature, and introduce a probabilistic index namely standardized heatwave index (SHI), that provides spatially and temporally consistent information on heatwaves and warm-spells in a standardized scale. This data set employs the CPC global air temperature data at the surface level with a spatial resolution of 0.5°, and identifies heatwave and warm-spells using different definitions. The GHWR data provides heatwaves and warm-spells information based on daily max, min and average temperature levels.

One important feature of GHWR is its focus on warm-spells alongside heatwaves. Heatwaves with direct and immediate human impacts have been widely analyzed in the literature¹, whereas warm-spells are yet to be more rigorously studied. Warm-spells also deeply impact human lives and livelihood. For example, warm-spells have induced earlier spring snow melt, which in turn increase western US forest wildfire activity⁶³. Another prominent contribution of the GHWR data set is enabling spatial analysis of heatwaves at the global scale. Note that GHWR, as well as any other global product, often does not account for the urban heat islands (UHIs), unless the UHI effect is reflected in the input data. Inner cities might be 1 to 3 °C hotter than the open areas, a phenomenon that is referred to as urban heat island⁶⁴. This is due to a combination of ‘increased heat storage and reduced water retention capacities of the land surface’⁶⁵. UHIs’ major influence is on increasing the night-time temperature, which along with large exposure of vulnerable groups enhance the risks of heat waves².

Finally, governments and the public should take a more proactive approach to reduce the impacts of heatwaves⁶⁶. The sustained 2013 England heatwave is a good example of an extreme event for which preparation and proper response reduced mortality rate to much less than expected⁶⁷. Early warning systems⁶⁸ and science-informed strategies to help the vulnerable population^{50,69} are essential to reduce the impacts of heatwaves. However, improving our understanding of heatwave features, developing warning systems and adaption strategies all rely on data. We hope that GHWR facilitates research and data analysis on heatwaves and their characteristics.

Caveats

The GHWR data set is constructed based on the CPC gridded global daily temperature record, and its accuracy largely depends on the source data. The CPC gridded temperature data set is constructed using a combination of two large data sets of observation stations around the globe, namely Global Historical Climatology Network version 2 (GCHN)^{70,71} and Climate Anomaly Monitoring System (CAMS)⁷². GCHN and CAMS together provide some 10,978 stations from around the globe (although not consistent through time and space), measured temperatures of which are interpolated to uniform global grids with a spatial resolution of 0.5° × 0.5° using an anomaly interpolation method, which is superior to other interpolation techniques in face of sparse station network coverage and temporally missing data⁷³. Moreover, temperatures are topographically adjusted using spatiotemporally varying temperature lapse rate estimated from observation-based global reanalysis temperature fields. This data set is comprehensively vetted and contrasted against other products such as Parameter elevation Regression on Independent Slopes Model (PRISM), National Climate Data Center (NCDC) Climate Division Data, Climate Research Unit (CRU) Climatology, Global Climate Data Assimilation System (CDAS) Reanalysis II and ERA40 reanalysis data sets, and rendered a “reasonably good” performance⁷³.

Note that GCHN and CAMS stations have a good coverage over the U.S., Europe, Russia and China; and less coverage over central part of Southern America, majority of Africa, and central Australia. Spatiotemporal coverage of observation stations used in CPC data set is displayed in Fig. 2 of Fan and Van den Dool⁷³ and can also be accessed through the Climate Prediction Center website. Accuracy of the gridded temperature data set is largely dependent on the spatiotemporal density of observation stations, and hence care is advised when using the GHWR product for areas with very few gauges. However, if lack of accuracy in the source data is manifested in terms of (uniform) bias, it would not impact the percentile-based methods used for identifying heatwaves/warm-spells. Finally, the presented GHWR toolbox allows integrating other data sources for analyzing heatwaves and warm spells.

References

- Perkins, S. E. A review on the scientific understanding of heatwaves—their measurement, driving mechanisms, and changes at the global scale. *Atmospheric Research* **164**, 242–267 (2015).
- Steffen, W., Hughes, L. & Perkins, S. *Heatwaves: Hotter, Longer, More Often*. Climate Council of Australia (2014).
- Mazdiyasi, O. *et al.* Increasing probability of mortality during Indian heat waves. *Science Advances* **3**, e1700066 (2017).
- Gao, J. *et al.* Impact of extreme high temperature on mortality and regional level definition of heat wave: a multi-city study in China. *Science of the Total Environment* **505**, 535–544 (2015).
- Chhetri, P., Hashemi, A., Basic, F., Manzoni, A. & Jayatilleke, G. Bushfire, heat wave and flooding case studies from Australia. *Report from the International Panel of the Weather project funded by the European Commission’s 7th framework programme* (2012).
- Shukla, S., Safeeq, M., AghaKouchak, A., Guan, K. & Funk, C. Temperature impacts on the water year 2014 drought in California. *Geophysical Research Letters* **42**, 4384–4393 (2015).
- Vahedifard, F., AghaKouchak, A. & Robinson, J. D. Drought threatens California’s levees. *Science* **349**, 799–799 (2015).
- Shaposhnikov, D. *et al.* Mortality related to air pollution with the Moscow heat wave and wildfire of 2010. *Epidemiology (Cambridge, Mass.)* **25**, 359 (2014).
- Ciais, P. *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* **437**, 529–533 (2005).
- Lobell, D. B. & Field, C. B. Global scale climate-crop yield relationships and the impacts of recent warming. *Environmental Research Letters* **2**, 014002 (2007).
- Stefanon, M., D’Andrea, F. & Drobinski, P. Heatwave classification over Europe and the Mediterranean region. *Environmental Research Letters* **7**, 014023 (2012).
- Russo, S. *et al.* Magnitude of extreme heat waves in present climate and their projection in a warming world. *Journal of Geophysical Research: Atmospheres* **119** (2014).
- Panda, D. K., AghaKouchak, A. & Ambast, S. K. Increasing heat waves and warm spells in India, observed from a multispect framework. *Journal of Geophysical Research: Atmospheres* **122**, 3837–3858 (2017).
- Perkins, S. & Alexander, L. On the measurement of heat waves. *Journal of Climate* **26**, 4500–4517 (2013).

15. Klein Tank, A. & Können, G. Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *Journal of Climate* **16**, 3665–3680 (2003).
16. Zhang, X. *et al.* Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisciplinary Reviews: Climate Change* **2**, 851–870 (2011).
17. Donat, M. *et al.* Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: the hadex2 dataset. *Journal of Geophysical Research: Atmospheres* **118**, 2098–2118 (2013).
18. Caesar, J. *et al.* Changes in temperature and precipitation extremes over the indo-pacific region from 1971 to 2005. *International Journal of Climatology* **31**, 791–801 (2011).
19. Donat, M. G. *et al.* Global land-based datasets for monitoring climatic extremes. *Bulletin of the American Meteorological Society* **94**, 997–1006 (2013).
20. Alexander, L. V. *et al.* Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.* **111**, D05109, <https://doi.org/10.1029/2005JD006290> (2006).
21. Perkins, S., Alexander, L. & Nairn, J. Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophysical Research Letters* **39** (2012).
22. Nairn, J. R. & Fawcett, R. J. The excess heat factor: a metric for heatwave intensity and its use in classifying heatwave severity. *International Journal of Environmental Research and Public Health* **12**, 227–253 (2014).
23. Kent, S. T., McClure, L. A., Zaitchik, B. F., Smith, T. T. & Gohlke, J. M. Heat waves and health outcomes in Alabama (USA): the importance of heat wave definition. *Environmental Health Perspectives* **122**, 151 (2014).
24. Semenza, J. C., McCullough, J. E., Flanders, W. D., McGeehin, M. A. & Lumpkin, J. R. Excess hospital admissions during the July 1995 heat wave in Chicago. *American Journal of Preventive Medicine* **16**, 269–277 (1999).
25. Tong, S., Wang, X. Y. & Barnett, A. G. Assessment of heat-related health impacts in Brisbane, Australia: comparison of different heatwave definitions. *PLoS One* **5**, e12155 (2010).
26. Peterson, T. C. *et al.* Monitoring and understanding changes in heat waves, cold waves, floods, and droughts in the United States: state of knowledge. *Bulletin of the American Meteorological Society* **94**, 821–834 (2013).
27. Parsons, K. *Human Thermal Environments: The Effects of Hot, Moderate, and Cold Environments on Human Health, Comfort, and Performance*. CRC press (2014).
28. Berry, H. L., Bowen, K. & Kjellstrom, T. Climate change and mental health: a causal pathways framework. *International Journal of Public Health* **55**, 123–132 (2010).
29. Hutter, H.-P., Moshammer, H., Wallner, P., Leitner, B. & Kundi, M. Heatwaves in Vienna: effects on mortality. *Wiener Klinische Wochenschrift* **119**, 223–227 (2007).
30. Hertel, S., Le Tertre, A., Jöckel, K.-H. & Hoffmann, B. Quantification of the heat wave effect on cause-specific mortality in Essen, Germany. *European Journal of Epidemiology* **24**, 407–414 (2009).
31. Huang, W., Kan, H. & Kovats, S. The impact of the 2003 heat wave on mortality in Shanghai, China. *Science of the Total Environment* **408**, 2418–2420 (2010).
32. Nitschke, M. *et al.* Impact of two recent extreme heat episodes on morbidity and mortality in Adelaide, south Australia: a case-series analysis. *Environmental Health* **10**, 42 (2011).
33. Wang, X. Y. *et al.* The impact of heatwaves on mortality and emergency hospital admissions from non-external causes in Brisbane, Australia. *Occupational and Environmental Medicine* **69**, 163–169 (2012).
34. Williams, S. *et al.* The impact of summer temperatures and heatwaves on mortality and morbidity in Perth, Australia 1994–2008. *Environment International* **40**, 33–38 (2012).
35. Sun, X. *et al.* Heat wave impact on mortality in Pudong new area, China in 2013. *Science of The Total Environment* **493**, 789–794 (2014).
36. Karl, T. R., Nicholls, N. & Ghazi, A. Clivar/gcos/wmo workshop on indices and indicators for climate extremes workshop summary. In *Weather and Climate Extremes 3–7* (Springer, 1999).
37. Peterson, T. *et al.* *Report on the Activities of the Working Group on Climate Change Detection and Related Rapporteurs*. World Meteorological Organization Geneva, (2001).
38. Fischer, E. M. & Schär, C. Future changes in daily summer temperature variability: driving processes and role for temperature extremes. *Climate Dynamics* **33**, 917 (2009).
39. Frich, P. *et al.* Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research* **19**, 193–212 (2002).
40. Perkins, S. E. Biases and model agreement in projections of climate extremes over the tropical Pacific. *Earth Interactions* **15**, 1–36 (2011).
41. Nairn, J. R. & Fawcett, R. G. Defining Heatwaves: Heatwave Defined as a Heat-impact Event Servicing All Community and Business Sectors in Australia (Centre for Australian Weather and Climate Research, 2013).
42. Vautard, R. *et al.* The simulation of European heat waves from an ensemble of regional climate models within the euro-cordex project. *Climate Dynamics* **41**, 2555–2575 (2013).
43. Schoetter, R., Cattiaux, J. & Douville, H. Changes of western European heat wave characteristics projected by the CMIP5 ensemble. *Climate Dynamics* **45**, 1601–1616 (2015).
44. Basagaña, X. *et al.* Heat waves and cause-specific mortality at all ages. *Epidemiology* **22**, 765–772 (2011).
45. Guttman, N. B. Accepting the standardized precipitation index: a calculation algorithm. *JAWRA Journal of the American Water Resources Association* **35**, 311–322 (1999).
46. Gringorten, I. I. A plotting rule for extreme probability paper. *Journal of Geophysical Research* **68**, 813–814 (1963).
47. Farahmand, A. & AghaKouchak, A. A generalized framework for deriving nonparametric standardized drought indicators. *Advances in Water Resources* **76**, 140–145 (2015).
48. Karl, T. R. & Knight, R. W. The 1995 Chicago heat wave: How likely is a recurrence? *Bulletin of the American Meteorological Society* **78**, 1107–1119 (1997).
49. Dematte, J. E. *et al.* Near-fatal heat stroke during the 1995 heat wave in Chicago. *Annals of Internal Medicine* **129**, 173–181 (1998).
50. Semenza, J. C. *et al.* Heat-related deaths during the July 1995 heat wave in Chicago. *New England Journal of Medicine* **335**, 84–90 (1996).
51. Stott, P. A., Stone, D. A. & Allen, M. R. Human contribution to the European heatwave of 2003. *Nature* **432**, 610–614 (2004).
52. García-Herrera, R., Díaz, J., Trigo, R., Luterbacher, J. & Fischer, E. A review of the European summer heat wave of 2003. *Critical Reviews in Environmental Science and Technology* **40**, 267–306 (2010).
53. Robine, J.-M. *et al.* Death toll exceeded 70,000 in Europe during the summer of 2003. *Comptes Rendus Biologies* **331**, 171–178 (2008).
54. Coumou, D. & Robinson, A. Historic and future increase in the global land area affected by monthly heat extremes. *Environmental Research Letters* **8**, 034018 (2013).
55. Trigo, R. M., García-Herrera, R., Díaz, J., Trigo, I. F. & Valente, M. A. How exceptional was the early August 2003 heatwave in France? *Geophys. Res. Lett.* **32**, L10701, <https://doi.org/10.1029/2005GL022410> (2005).

56. Black, E., Blackburn, M., Harrison, G., Hoskins, B. & Methven, J. Factors contributing to the summer 2003 European heatwave. *Weather* **59**, 217–223 (2004).
57. McMichael, A. J. & Lindgren, E. Climate change: present and future risks to health, and necessary responses. *Journal of Internal Medicine* **270**, 401–413 (2011).
58. Otto, F. E., Massey, N., Oldenborgh, G., Jones, R. & Allen, M. Reconciling two approaches to attribution of the 2010 Russian heat wave. *Geophysical Research Letters* **39** (2012).
59. Hauser, M., Orth, R. & Seneviratne, S. I. Role of soil moisture versus recent climate change for the 2010 heat wave in western Russia. *Geophysical Research Letters* **43**, 2819–2826 (2016).
60. Dole, R. *et al.* Was there a basis for anticipating the 2010 Russian heat wave? *Geophysical Research Letters* **38** (2011).
61. Steffen, W. L. *The Angry Summer*. Climate Commission Secretariat, Department of Climate Change and Energy Efficiency, (2013).
62. Masood, I., Majid, Z., Sohail, S., Zia, A. & Raza, S. The deadly heat wave of Pakistan, June 2015. *The International Journal of Occupational and Environmental Medicine* **6**, 672–247 (2015).
63. Westerling, A. L., Hidalgo, H. G., Cayan, D. R. & Swetnam, T. W. Warming and earlier spring increase western US forest wildfire activity. *science* **313**, 940–943 (2006).
64. Wolf, T. & McGregor, G. The development of a heat wave vulnerability index for London, United Kingdom. *Weather and Climate Extremes* **1**, 59–68 (2013).
65. Ramamurthy, P., Li, D. & Bou-Zeid, E. High-resolution simulation of heatwave events in New York city. *Theoretical and Applied Climatology* **128**, 89–102 (2017).
66. Kovats, R. S. & Kristie, L. E. Heatwaves and public health in Europe. *European Journal of Public Health* **16**, 592–599 (2006).
67. Green, H. K., Andrews, N., Armstrong, B., Bickler, G. & Pebody, R. Mortality during the 2013 heatwave in England-how did it compare to previous heatwaves? a retrospective observational study. *Environmental Research* **147**, 343–349 (2016).
68. Lowe, D., Ebi, K. L. & Forsberg, B. Heatwave early warning systems and adaptation advice to reduce human health consequences of heatwaves. *International Journal of Environmental Research and Public Health* **8**, 4623–4648 (2011).
69. Vandentorren, S. *et al.* August 2003 heat wave in France: risk factors for death of elderly people living at home. *The European Journal of Public Health* **16**, 583–591 (2006).
70. Peterson, T. C. & Vose, R. S. An overview of the global historical climatology network temperature database. *Bulletin of the American Meteorological Society* **78**, 2837–2849 (1997).
71. Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. & Houston, T. G. An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology* **29**, 897–910 (2012).
72. Ropelewski, C., Janowiak, J. & Halpert, M. *The climate anomaly monitoring system (CAMS)*. Climate Analysis Center, NWS, NOAA: Washington, DC, (1984).
73. Fan, Y. & Van den Dool, H. A global monthly land surface air temperature analysis for 1948-present. *J. Geophys. Res.* **113**, D01103, <https://doi.org/10.1029/2007JD008470> (2008).

Data Citation

1. Raei, E., Nikoo, M.R., AghaKouchak, A., Mazdiyasi, O. & Sadegh, M. figshare <https://doi.org/10.6084/m9.figshare.c.4004668> (2018).

Acknowledgements

This study was partially supported by National Science Foundation (NSF) grant (CMMI-1635797), National Oceanic and Atmospheric Administration (NOAA) grant (NA14OAR4310222) and California Energy Commission grant (500-15-005).

Author Contributions

E.R., M.N., A.A. and M.S. conceived the study. E.R. created the data set; M.S. and O.M. double-checked the data for accuracy. M.S. wrote the first draft of the manuscript, and all authors contributed to finalizing the paper. All authors reviewed the manuscript and provided input.

Additional Information

Competing interests: The authors declare no competing interests.

How to cite this article: Raei, E. *et al.* GHWR, a multi-method global heatwave and warm-spell record and toolbox. *Sci. Data.* 5:180206 doi: 10.1038/sdata.2018.206 (2018).

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 Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files made available in this article.

© The Author(s) 2018