



OPEN

DATA DESCRIPTOR

A high-spatial-resolution dataset of human thermal stress indices over South and East Asia

Yechao Yan¹  , Yangyang Xu² & Shuping Yue¹

Thermal stress poses a major public health threat in a warming world, especially to disadvantaged communities. At the population group level, human thermal stress is heavily affected by landscape heterogeneities such as terrain, surface water, and vegetation. High-spatial-resolution thermal-stress indices, containing more detailed spatial information, are greatly needed to characterize the spatial pattern of thermal stress to enable a better understanding of its impacts on public health, tourism, and study and work performance. Here, we present a $0.1^\circ \times 0.1^\circ$ gridded dataset of multiple thermal stress indices derived from the newly available ECMWF ERA5-Land and ERA5 reanalysis products over South and East Asia from 1981 to 2019. This high-spatial-resolution database of human thermal stress indices over South and East Asia (HiTiSEA), which contains the daily mean, maximum, and minimum values of UTCI, MRT, and eight other widely adopted indices, is suitable for both indoor and outdoor applications and allows researchers and practitioners to investigate the spatial and temporal evolution of human thermal stress and its impacts on densely populated regions over South and East Asia at a finer scale.

Background & Summary

Due to the unprecedented scale of climate change, extreme temperature events have become more intense and frequent in many parts of the world over the past few decades^{1–3}. The study of thermal stress or discomfort due to heat or cold extremes has attracted attention worldwide^{4–7}, as thermal stress can have a pronounced negative impact on human health, especially in vulnerable populations such as the elderly, chronically ill and poorer communities^{8–10}.

The level of human thermal stress is determined not only by the ambient air temperature (T_a) but also by a combination of other factors, including solar and thermal radiation (R), wind speed (V_a), relative humidity (RH), personal clothing and activity level. To date, more than 100 indices have been developed to assess and quantify human thermal stress^{11,12}. These indices vary considerably in type, complexity, and capability. Some of them are based on the principles of human thermal exchange, while others are based on empirical relationships obtained by examining human responses to various environmental factors.

Many of the empirical indices, such as the heat index (HI), humidity index (Humidex), net effective temperature (NET) and wind chill temperature (WCT), use only two or three environmental parameters (e.g., T_a and RH) and thus are only applicable to indoor space or outdoor shaded areas. However, some classic indices, due to their simple form and low data input for computation, remain attractive and are widely utilized by national and local weather services¹³.

A few human thermal stress indices, such as the universal thermal climate index (UTCI), the standard effective temperature (SET), and the physiological equivalent temperature (PET), consider more meteorological factors, allowing them to be used in both indoor and outdoor conditions. The UTCI, which is the focus of our study, is a state-of-the-art thermal stress indicator based on heat budget models of the human body and its surrounding environments¹⁴. The UTCI takes into account a suite of relevant meteorological variables (air temperature and humidity, wind speed, and longwave and shortwave radiant heat fluxes) as well as personal factors such as physical activity level and adaptive clothing behaviour, making it applicable in a variety of climates, seasons and spatial scales^{15–17}.

¹School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing, Jiangsu, China.

²Department of Atmospheric Sciences, College of Geosciences, Texas A&M University, College Station, Texas, USA.

✉e-mail: yanyechao@nuist.edu.cn

Thermal Indices	Full Name of the Indices	Air Temperature	Air Humidity	Wind Speed	Radiation
UTCI	universal thermal climate index	T_a	e	V_a	R
indoor UTCI	UTCI for indoor environment	T_a	e		
outdoor shaded UTCI	UTCI for outdoor shaded space	T_a	e	V_a	
MRT	mean radiant temperature				R
ESI	environment stress index	T_a	RH		SR
HI	heat index	T_a	RH		
Humidex	humidity index	T_a	e		
WBGT	wet-bulb globe temperature	T_a	e		
WBT	wet bulb temperature	T_a	RH		
WCT	wind chill temperature	T_a		V_a	
AT	apparent temperature	T_a	e	V_a	
NET	net effective temperature	T_a	RH	V_a	

Table 1. Thermal indices and their input variables. Note: T_a , e , and RH represent the air temperature, water vapour pressure, and relative humidity, respectively. V_a stands for the 10-metre wind speed, with the exception of the NET (Eq. 12), which requires an input of wind speed at 1.2 m above the ground. R stands for the radiation variables, including direct, diffuse, and reflected solar radiation, as well as upward and downward thermal radiation, while SR represents the solar radiation, which includes both the direct and diffuse solar radiation reaching the horizontal surface of the Earth. The indoor UTCI, outdoor shaded UTCI, and UTCI, which take 2, 3, and 4 parameters, respectively, are applicable to indoor, outdoor shaded, and outdoor unshaded environments. All indices are with a unit expressed in °C.

Using the key variables from ERA5-Land reanalysis, along with direct solar radiation from ERA5, this paper presents a higher-spatial-resolution ($0.1^\circ \times 0.1^\circ$) gridded dataset with multiple thermal-stress indices. This newly developed dataset, called the High-spatial-resolution Thermal-stress Indices over South and East Asia (HiTiSEA), contains daily maximum, minimum, and mean values of the indoor and outdoor UTCI (including shaded and unshaded outdoor environments), as well as the mean radiant temperature (MRT) and eight other widely used empirical indices, as listed in Table 1, from 1981 to 2019 for the area of South and East Asia, a region making up more than half of the world's population, many of which are vulnerable to the impacts of extreme thermal stress. Another reason for the limited spatial coverage is due to data access issues, as observed meteorological data used in this study for validation are only available in this area.

Compared to the existing $0.25^\circ \times 0.25^\circ$ spatial-resolution ERA5-HEAT (Human thErMAl comforT) product¹⁸ released by the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Human Discomfort Indices (HDIs, also with a spatial resolution of $0.25^\circ \times 0.25^\circ$) computed from the Global Land Data Assimilation System (GLDAS) by Mistry¹⁹, the new features of the HiTiSEA dataset include the following:

- (i) It features a higher spatial resolution ($0.1^\circ \times 0.1^\circ$, but smaller spatial coverage) based on ERA5-Land reanalysis;
- (ii) It contains 3 types of UTCI indices (UTCI, indoor UTCI, and outdoor shaded UTCI), MRT metrics, and eight other empirical thermal indices that allow applications for indoor, outdoor shaded, and outdoor unshaded environments;
- (iii) It provides comprehensive validation based on thousands of weather stations over South and East Asia (including bias and root mean square error for each index at each station released as part of the dataset), which enables users to further evaluate and select some indices over the others and conduct bias correction if needed;
- (iv) It shares freely available Python scripts that allow users to calculate the UTCI and its variants, as well as other thermal indices for any part of the world.

With a finer spatial resolution and a wider applicability to both indoor and outdoor conditions, this multiple-index dataset is a valuable resource for health authorities and scientists to study the evolution of the thermal environment and identify high-risk areas where people are exposed to potential heat or cold stress. Tourism professionals will find it useful in evaluating thermal comfort conditions and defining the most appropriate time for specific recreational activities. These data can also be used by researchers and policy makers to assess the costs of extreme thermal stress on the economy through reduced labour productivity. Moreover, this newly developed dataset can help researchers estimate the energy demand required to meet residential heating or cooling needs, especially in India, Bangladesh, and China, where large gaps exist²⁰.

Methods

Data source. A complete set of meteorological data, including air temperature and humidity, wind speed, and shortwave and longwave radiation fluxes, is required for computation of the thermal indices included in the HiTiSEA dataset. While various reanalysis products, such as the Global Land Data Assimilation System Version 2 (GLDAS-2) developed jointly by the National Aeronautics and Space Administration (NASA) and National Oceanic and Atmospheric Administration (NOAA), the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) produced by NASA, the Japanese 55-year Reanalysis (JRA-55) released by

Variable	Description	Units	Source Dataset
T_a	The temperature of the air at 2 m above the ground	K	ERA5-Land
T_d	Dewpoint temperature at 2 m above the ground	K	ERA5-Land
u	Eastward component of the 10 m wind	m s^{-1}	ERA5-Land
v	Northward component of the 10 m wind	m s^{-1}	ERA5-Land
$ssrd$	Surface solar radiation downwards: the amount of shortwave radiation (both direct and diffused) that reaches a horizontal plane at the surface	J m^{-2}	ERA5-Land
ssr	Surface net solar radiation, the amount of shortwave radiation (both direct and diffuse) that reaches a horizontal plane at the surface minus the amount reflected at the surface	J m^{-2}	ERA5-Land
$strd$	Surface thermal radiation downwards: the amount of thermal (longwave) radiation emitted by the atmosphere and clouds that reaches a horizontal plane at the surface	J m^{-2}	ERA5-Land
str	Surface net thermal radiation: the difference between downward and upward thermal radiation passing through a horizontal plane at the surface	J m^{-2}	ERA5-Land
$fdir$	Direct solar radiation at the surface: the amount of direct shortwave radiation passing through a horizontal plane at the Earth's surface, which is equal to the $ssrd$ but excluding the diffused solar radiation.	J m^{-2}	ERA5

Table 2. Variables from ERA5-Land and ERA5 to compute MRT and UTCI.

the Japan Meteorological Agency (JMA), etc., are currently available, the ERA5 and ERA5-Land datasets developed by the ECMWF are chosen for use in the present study, as other reanalysis products have either (i) coarser spatial resolutions (e.g., the GLDAS-2, MERRA-2, and JRA-55 provide gridded meteorological variables with a horizontal resolution of $0.25^\circ \times 0.25^\circ$, $0.625^\circ \times 0.5^\circ$, and $1.25^\circ \times 1.25^\circ$ in longitude/latitude, respectively) or (ii) incomplete meteorological variables (e.g., direct solar radiation is not available in other reanalysis products).

ERA5 is the fifth-generation atmospheric reanalysis product recently released by the ECMWF. ERA5 is generated using the latest version of the Integrated Forecasting System and modern parameterizations technique, with a horizontal resolution of $0.25^\circ \times 0.25^\circ$, a temporal resolution of 1 h, and a vertical resolution of 137 levels from the surface up to a height of 80 km^{21,22}. By rerunning the land component of the ERA5 climate reanalysis, the ECMWF has developed a state-of-the-art reanalysis dataset called ERA5-Land, which covers the land surface of the entire globe with a horizontal resolution of $0.1^\circ \times 0.1^\circ$ and a temporal resolution of 1 h. Using “lapse rate correction”, ERA5 air temperature, air humidity and pressure used to run ERA5-Land are corrected to account for the altitude difference between the grid of the forcing and the higher-resolution grid of ERA5-Land²³.

Due to storage limitations, HiTiSEA version 1 presented in this study spanned the period from 1981 to 2019, covering the area of East Asia and South Asia (65°E to 155°E and 3°N to 58°N). To compute the MRT and UTCI, hourly meteorological variables (Table 2) were retrieved from ERA5-Land, with the exception of $fdir$ (direct solar radiation at the surface), which is only available in ERA5. Since the variable $fdir$ has a coarser resolution, it is regridded from $0.25^\circ \times 0.25^\circ$ to $0.1^\circ \times 0.1^\circ$ using nearest-neighbour interpolation to match the other variables. The nearest-neighbour method is used due to its advantage in preserving the values of the original data²⁴. Other resampling methods, such as bilinear and cubic convolution, can increase uncertainties by altering or even distorting the grid values of the original data. Furthermore, the accumulated radiation values in the original dataset of ERA5-Land (J m^{-2} , as in Table 2) are transformed to hourly values. Note that the convention for accumulations used in ERA5-Land differs from that for ERA5²⁵.

Data processing procedure. Figure 1 shows the procedure for processing the ERA5-Land and ERA5 reanalysis data and producing the multi-thermal-index dataset. The procedure includes the following five steps: (1) extracting the variable of direct solar radiation from ERA5 and regridding it from 0.25° to 0.1° ; (2) extracting other radiation variables from ERA5-Land and converting the accumulated radiations to hourly accumulated values; (3) computing the radiation variables, expressed in W/m^2 , in the MRT formula (Eq. 1); (4) calculating the MRT; (5) computing the indoor and outdoor UTCI as well as other empirical thermal indices on an hourly basis; and (6) performing summary statistics for these hourly indices and archiving the HiTiSEA dataset with daily mean, maximum and minimum values.

Calculation of MRT. The MRT is defined as the effective temperature of an imaginary enclosure in which the radiant heat transfer from the human body equals the actual radiant heat transfer in the real nonuniform enclosure²⁶. MRT is the key parameter used to compute UTCI. It is used to assess the impact of radiation fluxes on the energy balance of human bodies, which is not accounted for in indices such as T_w . In operational human biometeorology, fluxes are related to an upright standing or walking person²⁷. Since a resolution of $0.1^\circ \times 0.1^\circ$ or approximately 10 km is insufficient to capture the details of individual persons' surrounding environment, an unshaded plain is assumed with solid angles (f_a) of the land surface and the sky both set to 0.5. Then, the MRT for the outdoor environment is given by Weihs *et al.*²⁸.

$$MRT = \left\{ \frac{1}{\sigma} \left[\frac{\alpha_k}{\varepsilon_p} (f_p \cdot I_{sw} + f_a \cdot D_{sw} + f_a \cdot R_{sw}) + f_a \cdot (D_{lw} + U_{lw}) \right] \right\}^{0.25} - 273.5 \quad (1)$$

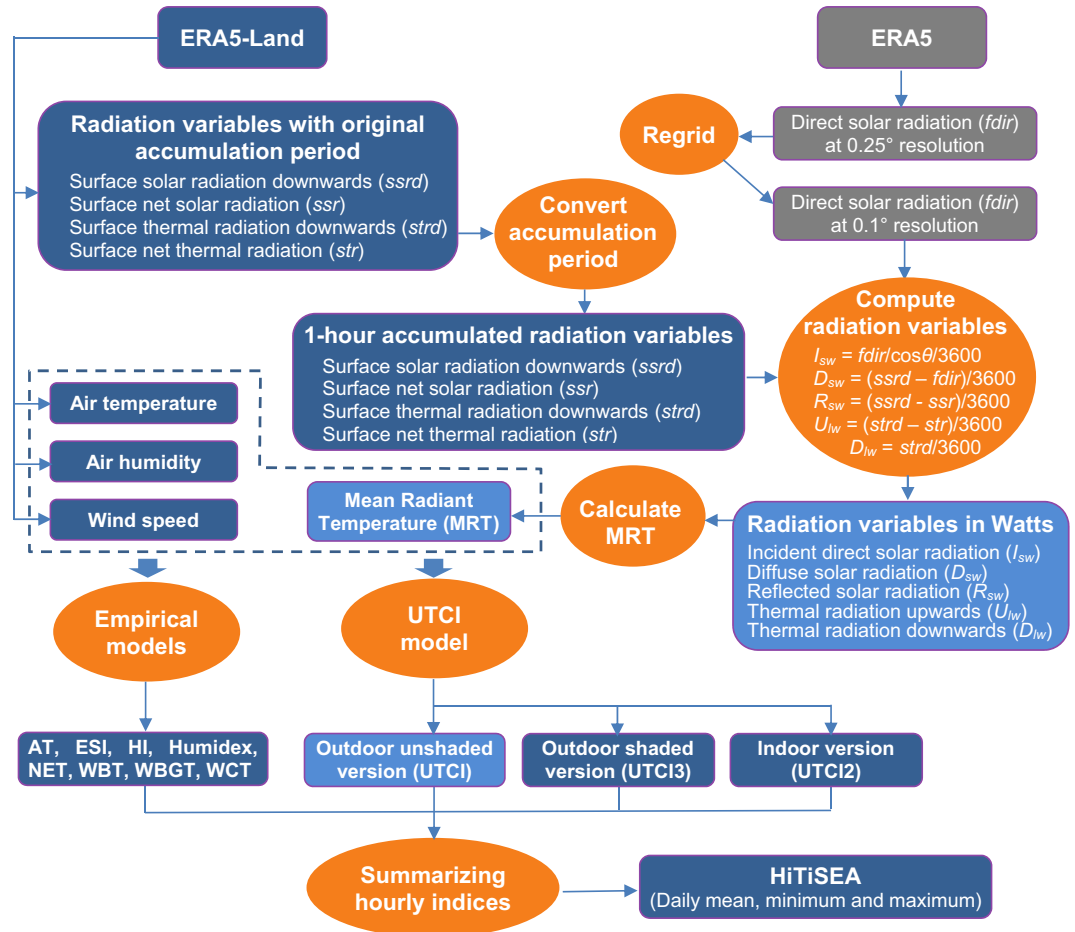


Fig. 1 Schematic of the workflow to generate the HiTiSEA product.

where MRT is the mean radiant temperature ($^{\circ}C$), σ is the Stefan Boltzmann constant ($5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$), α_k is the absorption coefficient of the typical human body for shortwave radiation (here assuming standard value 0.7), and ε_p is the emissivity coefficient of the human body (here assuming standard value 0.97). I_{sw} , D_{sw} , R_{sw} , D_{lw} , and U_{lw} , all expressed in W/m^2 and calculated following the equations in Fig. 1, are the anisotropic incident (I_{sw}) direct shortwave radiation flux, isotropic diffuse (D_{sw}) shortwave radiation flux, surface reflected (R_{sw}) short-wave radiation flux, downwelling (D_{lw}) longwave radiation fluxes and upwelling (U_{lw}) longwave radiation fluxes, respectively.

The projected area factor (f_p) accounts for the directional dependence and is a function of the solar zenith angle. For a rotationally symmetric standing human body, f_p can be estimated using the following formula^{29,30}:

$$f_p = 0.308 \cdot \cos \left\{ \left(\frac{\pi}{2} - \theta \right) \cdot \left[1 - \frac{\left(90 - \frac{180}{\pi} \theta \right)^2}{48402} \right] \right\} \quad (2)$$

where θ is the solar zenith angle (in radians). The cosine of the solar zenith angle can be calculated following Woan³¹:

$$\cos \theta = \sin \delta \sin \varphi + \cos \delta \cos \varphi \cos h \quad (3)$$

where φ is the geographical latitude, δ is the solar declination angle as a function of a given date of the year and h is the hour angle in local solar time. The latter two parameters, i.e., δ and h , are calculated following Spencer³² and NOAA³³.

Considering that I_{sw} can be overestimated at sunset and sunrise times (note that it is computed by dividing $fdir$ by $\cos \theta$, which is close to 0 during those twilight periods), the average $\cos \theta$ between the beginning of the forecast time and the end of the forecast step (1-hour interval in this case) is used instead of the exact endpoint of the forecast time. A detailed description for calculating the average $\cos \theta$ can be found in Di Napoli *et al.*³⁴.

Calculation of UTCI. The UTCI is defined as an equivalent ambient temperature (in the unit of $^{\circ}C$) of a reference environment that produces the same physiological response of a typical person as in the actual environment¹⁴. Calculation of physiological response to meteorological inputs is based on an advanced multinode

thermoregulation model (consisting of 12 body elements with a total of 187 tissue nodes) coupled with an adaptive clothing model considering behavioural changes in clothing insulation related to the actual thermal environment^{15,16}. The reference environment¹⁴ is defined as a condition with calm air (a 10-m wind speed of 0.5 m/s), where the mean radiant temperature equals the air temperature, a 50% relative humidity is used for $T_a \leq 29^\circ\text{C}$, and a water vapour pressure $e = 20$ hPa is used for $T_a > 29^\circ\text{C}$, where an average person walks at 4 km/h, generating a metabolic rate of 135 W/m^2 .

Due to our need to produce a climate dataset with high spatial and temporal resolutions, calculating the UTCI by repeatedly running the thermoregulation model is not practical. In this study, a 6th-order polynomial regression function given by Bröde *et al.*³⁵ is used to calculate the outdoor, unshaded UTCI. The simple form of the function is written as follows (with the full equation in the code release):

$$UTCI = T_a + f(T_a, V_a, e, MRT - T_a) \quad (4)$$

where T_a is the 2-metre air temperature, V_a is the 10-metre wind speed (m/s), e is the water vapour pressure (hPa), and MRT is the mean radiant temperature ($^\circ\text{C}$).

To compute the outdoor shaded UTCI, MRT is set equal to the air temperature, thus ignoring the radiation flux's contribution to thermal comfort. To compute indoor UTCI, in addition to MRT , V_a is also set to the reference values of 0.5 m/s, thus further ignoring the ambient wind speed's contribution to thermal comfort.

Calculation of other empirical thermal indices. *Apparent Temperature.* The apparent temperature (AT) is defined as the temperature at the reference humidity level, producing the same amount of discomfort as that experienced under the current ambient temperature, humidity, and solar radiation³⁶. Two forms are in use by the Australian Bureau of Meteorology: one includes radiation and one does not. The AT index used here is based on a mathematical model of an adult walking outdoors in the shade³⁷ and thus does not include radiation:

$$AT = T_a + 0.33 \times e - 0.7V_a - 4 \quad (5)$$

where AT is the apparent temperature ($^\circ\text{C}$), T_a is the air temperature ($^\circ\text{C}$), e is the water vapour pressure (hPa) and V_a is the 10-m wind speed (m/s).

Environment Stress Index. The environmental stress index (ESI) was introduced by Moran *et al.*³⁸ in 2001 as a substitute for the wet bulb globe temperature (WBGT), which was hard to use due to the required measurements of nonconventional meteorological variables such as the wet-bulb temperature and global temperature. The ESI, which was validated by using large databases and was found to be highly correlated with the WBGT³⁹, is calculated as³⁸:

$$ESI = 0.63T - 0.03RH + 0.002SR + 0.0054 \times T \times RH - \frac{0.073}{0.1 + SR} \quad (6)$$

where T is the air temperature ($^\circ\text{C}$), RH is the relative humidity (%), and SR is the amount of solar radiation (both direct and diffused, in W/m^2) that reaches a horizontal plane of the Earth's surface.

Heat Index. The heat index (HI) is widely used across the United States. It is a measure of how hot it feels when relative humidity is factored in along with the air temperature. The original heat index is a hot-weather version of AT that involves a collection of equations and a large number of input parameters⁴⁰. To arrive at an equation that uses more conventional independent variables, a regression equation was obtained by Rothfus⁴¹ through multiple regression analysis based on the data from Steadman's table:

$$\begin{aligned} HI = & -42.379 + 2.04901523 \times T_a + 10.14333127 \times RH \\ & -0.22475541 \times T_a \times RH - 0.00683783 \times T_a^2 - 0.05481717 \times RH^2 \\ & + 0.00122874 \times T_a^2 \times RH + 0.00085282 \times T_a \times RH^2 \\ & - 0.00000199 \times T_a^2 \times RH^2 \end{aligned} \quad (7)$$

where HI is the heat index (in $^\circ\text{F}$), T_a is the temperature (in $^\circ\text{F}$) and RH is the relative humidity (in %).

If the RH is less than 13% and the temperature is between 80 and 112 $^\circ\text{F}$, then the following adjustment is subtracted from HI :

$$Adj = \frac{13 - RH}{4} \times \sqrt{\frac{17 - |T_a - 95|}{17}} \quad (8)$$

On the other hand, if the RH is greater than 85% and the temperature is between 80 and 87 $^\circ\text{F}$, then the following adjustment is added to HI :

$$Adj = \frac{RH - 85}{10} \times \frac{87 - T_a}{5} \quad (9)$$

The Rothfus regression is not suitable when the HI is below 80 $^\circ\text{F}$. In those cases, a simpler formula, provided by the National Oceanic and Atmospheric Administration⁴², is applied to produce values consistent with Steadman's results:

$$HI = 0.5 \times [T_a + 61 + (T_a - 68) \times 1.2 + RH \times 0.094] \quad (10)$$

Humidex. The Humidex (short for humidity index) is an index developed by Canadian meteorologists⁴³ to describe how hot the weather feels to the average person. By combining the effects of air temperature (T_a , in °C) and water vapour pressure (e , in hPa), the Humidex (in °C) is calculated as follows:

$$\text{Humidex} = T_a + 0.5555 \times (e - 10) \quad (11)$$

Net Effective Temperature. The net effective temperature (NET) was originally established in 1923 by Houghton and Yaglou⁴⁴ to estimate the relative effects of air temperature and humidity on body comfort. It was amended, based on laboratory experiments, by Missenard⁴⁵ using the empirical relationship between the identical state of the organism's thermoregulatory capacity (warm and cold perception) and differing temperature and humidity of the surrounding environment. However, Missenard's formula seemed exclusively appropriate for hot weather conditions. Further modifications included the effect of winds and extended its use to cold conditions^{46,47}. The resulting formula takes the following form:

$$\text{NET} = 37 - \frac{37 - T_a}{0.68 - 0.0014 \times RH + \frac{1}{1.76 + 1.4 \times V_a^{0.75}}} - 0.29 \times T_a \times (1 - 0.01 \times RH) \quad (12)$$

where NET is the net effective temperature (°C), T_a is the air temperature (°C), RH is the relative humidity (%) and V_a is the wind speed (m/s) at a height of 1.2 m, which is approximated by applying a typical logarithmic wind profile approach:

$$V_a = V_{z_r} \frac{\log(Z/Z_0)}{\log(Z_r/Z_0)} \quad (13)$$

where Z is the height (m) of the centre of the body element above ground (i.e., 1.2 m in this case), V_{z_r} (m) is the wind speed at a reference height of the meteorological measurement (i.e., 10 m), and z_0 (m) is the roughness length, assumed to be 0.01 m¹⁶.

Wet-Bulb Globe Temperature. The wet-bulb globe temperature (WBGT), developed in the 1950s by the US Navy as part of a study on heat-related injuries during military training, is one of the most widely used heat stress indices throughout the world. The WBGT is a composite temperature in which the natural wet-bulb temperature T_w (°C), the black globe temperature T_g (°C), and the dry-bulb temperature T_d (°C) are added up with different weightings according to their importance⁴⁸:

$$\text{WBGT} = 0.7 \times T_w + 0.2 \times T_g + 0.1 \times T_d \quad (14)$$

Due to the lack of T_w and T_g in ERA5-Land, the WBGT is calculated using a simplified equation, given by the Australian Bureau of Meteorology⁴⁹, as follows:

$$\text{WBGT} = 0.567 \times T_a + 0.393 \times e + 3.94 \quad (15)$$

where T_a is the air temperature (°C) and e is the water vapour pressure (hPa).

This simplified equation, which only takes the air temperature and the water vapour pressure into consideration, is only applicable for indoor environments.

Wet Bulb Temperature. The wet bulb temperature (WBT) is the lowest temperature that can be reached under current ambient conditions through the evaporation of water. At 100% relative humidity, the WBT is equal to the air temperature, while at a lower humidity, it is lower than the air temperature due to the effect of evaporative cooling. In practice, WBT is measured using a wet-bulb thermometer. In this paper, WBT is approximated using Stull's formula⁵⁰:

$$\begin{aligned} \text{WBT} = & T_a \times \text{atan}[0.151977(RH + 8.313659)^{0.5}] + \text{atan}(T_a + RH) - \text{atan}(RH - 1.676331) \\ & + 0.00391838(RH)^{1.5} \text{atan}(0.023101 \times RH) - 4.686035 \end{aligned} \quad (16)$$

where WBT is the wet bulb temperature (°C), T_a is the air temperature or dry bulb temperature (°C) and RH is the relative humidity (%). The approximation is valid for relative humidity ranging from 5% to 99% and air temperature from -20°C to 50°C.

Wind Chill Temperature. The wind chill index (WCI), developed in the 1940s and revised by weather services in the USA and Canada, expresses the enhancement of heat loss in cold climates from exposed body parts due to wind. In the present study, the WCT was calculated using a multiple regression formula developed by the Joint Action Group for Temperature Indices⁵¹. The following formula provides the equivalent temperature (what the temperature feels like to the human body when the cooling effect of wind is taken into account) as an output:

$$\text{WCT} = 13.12 + 0.6215 \times T_a - 11.37 \times V_a^{0.16} + 0.3965 \times T_a \times V_a^{0.16} \quad (17)$$

where T_a is the air temperature (in °C) and V_a is the 10-m wind speed (in km/h).

Data Records

The geographically gridded dataset consists of daily mean, maximum and minimum values of the following thermal indices at a $0.1^\circ \times 0.1^\circ$ spatial resolution: (1) the universal thermal climate index for the unshaded outdoor environment (UTCI); (2) the universal thermal climate index for shaded outdoor space (outdoor shaded UTCI); (3) the universal thermal climate index for the indoor environment (indoor UTCI); (4) the apparent temperature (AT); (5) the environmental stress index (ESI); (6) the heat index (HI); (7) the Humidex; (8) the mean radiant temperature (MRT); (9) the net effective temperature (NET); (10) the wet bulb temperature (WBT); (11) the wet bulb globe temperature (WBGT); and (12) the wind chill temperature (WCT).

The dataset spans the period from January 3, 1981, to December 31, 2019, covering the area of South and East Asia (65° – 155° E, 3° – 58° N). Individual thermal stress indices were aggregated into a single NetCDF file on a daily basis. Each daily file is named as follows:

HiTiSEA_YYYY-MM-DD.nc

where “YYYY-MM-DD” represents the date of the daily file.

The variables are named in the following format: Index_mean, Index_max and Index_min. For example, the variables for the daily mean, maximum and minimum of UTCI are named UTCI_mean, UTCI_max and UTCI_min, respectively. For each variable, grid cells with no data are filled with the value -32767 .

This newly developed dataset⁵², with a total volume of 450 GB, contains 14242 daily NetCDF files that are archived by year and compressed into tar.gz files to save storage space. The dataset and its metadata are freely available at the figshare repository (<https://doi.org/10.6084/m9.figshare.c.5196296>).

Technical Validation

We select nine indices in our dataset (Table 3) for comparison, which do not require radiation data for computation. They were compared against the corresponding indices computed from observed meteorological data obtained from the China Meteorological Data Service Center (CMDSC)⁵³ through a portal located at Nanjing University of Information Science & Technology (NUIST)⁵⁴. The observed data in 2018 have a temporal resolution of 3 h, including the air temperature (T_a), dew-point temperature (T_d), 10-metre wind speed (V_a), and surface air pressure (P). Meteorological records with missing or incomplete values (missing any of the above 4 meteorological variables) were excluded, and 1281 stations were finally used for validation.

Table 3 shows that the RMSE values for daily mean, maximum, and minimum indoor UTCI are 1.6°C , 1.9°C , and 2.2°C , respectively, with 81% of the stations presenting an RMSE for daily mean lower than 2°C (Fig. 2 upper left), making this index ideal for indoor thermal stress assessment. In comparison, the outdoor shaded UTCI shows higher RMSE values, with approximately 30% of the stations having an RMSE for daily mean less than 2°C and 71% having an RMSE below 3°C . Stations with RMSE values greater than 5°C , as depicted in Fig. 2 (upper right), are mostly located in higher-latitude areas and a few coastal areas where the wind speed is significantly affected by local factors. As depicted in Fig. 2 (lower row), both the estimated indoor UTCI and outdoor shaded UTCI are overall negatively biased, with more stations exhibiting negative bias and fewer stations, most of which are located north of the line of latitude 40°N , exhibiting positive bias.

Among the empirical thermal indices with 2 climate parameters, the WBGT shows the highest accuracy, with RMSE values ranging from 1.1 to 1.6°C , followed by the WBT ranging from 1.3 to 1.9°C . HI and the Humidex, which also take air temperature and air humidity as input variables, present RMSE values no more than 2.5°C and 2.7°C , respectively. The WCT with input variables of air temperature and wind speed, however, shows the lowest accuracy, with RMSE values varying between 3.1°C and 4.8°C . For the 3-parameter empirical thermal indices, the average RMSE values for daily mean, maximum and minimum AT are found to be 2.0°C , 2.3°C , and 2.7°C , respectively, and the RMSE values for NET are all above 2.7°C but no more than 3.6°C .

Almost all indices listed in Table 3 are slightly biased towards negative values, which suggests that compared to the observed results, these thermal-stress indices are underestimated in most cases. While on average, the bias for estimation of daily maximum WCT can be as large as -2.5°C , the biases for most indices are within -1°C .

The other three indices, i.e., the MRT, the outdoor unshaded UTCI, and the ESI, which require radiation for calculation, were also evaluated against the corresponding indices computed from observations but with a much smaller sample. This is because hourly radiation data are not open to the public and are difficult to acquire. While commonly observed meteorological variables (i.e., T_a , T_d , V_a , P , etc.) are all available at the 1281 stations with a time step of 3 h, and only 8 of them provided daily radiation observations for 2018 to registered users on CMDSC's website. The observed radiation data include daily values of global radiation, direct solar radiation, diffuse solar radiation, reflected solar radiation, maximum global radiation flux, the time when maximum global radiation flux occurs, etc. To assimilate the two sets of observations with different time steps, we rounded the time when the maximum global radiation flux occurred to the nearest 3-hour synoptic time (00:00, 03:00, 06:00 UTC, etc.). By doing so, we paired the maximum global radiation flux with the commonly observed meteorological data. After removing incomplete records, these paired-up observations have a size of 2220 hourly records, as listed in Table 4 for each station. They were then fed into the BioKlima 2.6 software package⁵⁵ to calculate the MRT and the outdoor unshaded UTCI for those specific records. These observational results were used to validate the corresponding hourly MRT and UTCI from which the daily maximum, minimum, and mean MRT and UTCI were derived in our dataset. Similarly, the paired-up radiation data and other meteorological data were input into Eq. (6) to compute the observational ESI for validation of the corresponding ESI in our dataset.

Compared to the existing ERA5-HEAT product, which has an RMSE of $5.2 \pm 2.5^\circ\text{C}$ ¹⁸, this newly developed outdoor unshaded UTCI, due to the use of the enhanced resolution of ERA5-Land, exhibits improved accuracy with an average RMSE of 4.5°C , ranging from 2.9°C to 6.9°C (Table 4). However, using finer resolution radiation data from ERA5-Land does not seem to have a significant effect on the accuracy of the MRT, which has an average

Thermal Indices	Daily Mean		Daily Maximum		Daily Minimum	
	RMSE	Bias	RMSE	Bias	RMSE	Bias
indoor UTCI	1.6	-0.4	1.9	-0.7	2.2	-0.3
outdoor shaded UTCI	2.7	-0.9	3.1	-1.2	3.7	-0.7
HI	2.0	-0.6	2.4	-0.9	2.5	-0.4
Humidex	1.9	-0.6	2.3	-0.8	2.7	-0.5
WBGT	1.1	-0.4	1.3	-0.5	1.6	-0.3
WBT	1.3	-0.3	1.4	-0.4	1.9	-0.3
WCT	3.1	-1.7	4.8	-2.5	3.3	-1.3
AT	2.0	-0.7	2.3	-0.9	2.7	-0.7
NET	2.7	-0.3	3.3	-0.7	3.6	0.2

Table 3. Summary table of accuracy, in terms of RMSE (°C) and bias (°C), obtained by comparing the indices computed from ERA5-Land reanalysis and weather station observations. This table only lists the indices that do not require radiation as data input.

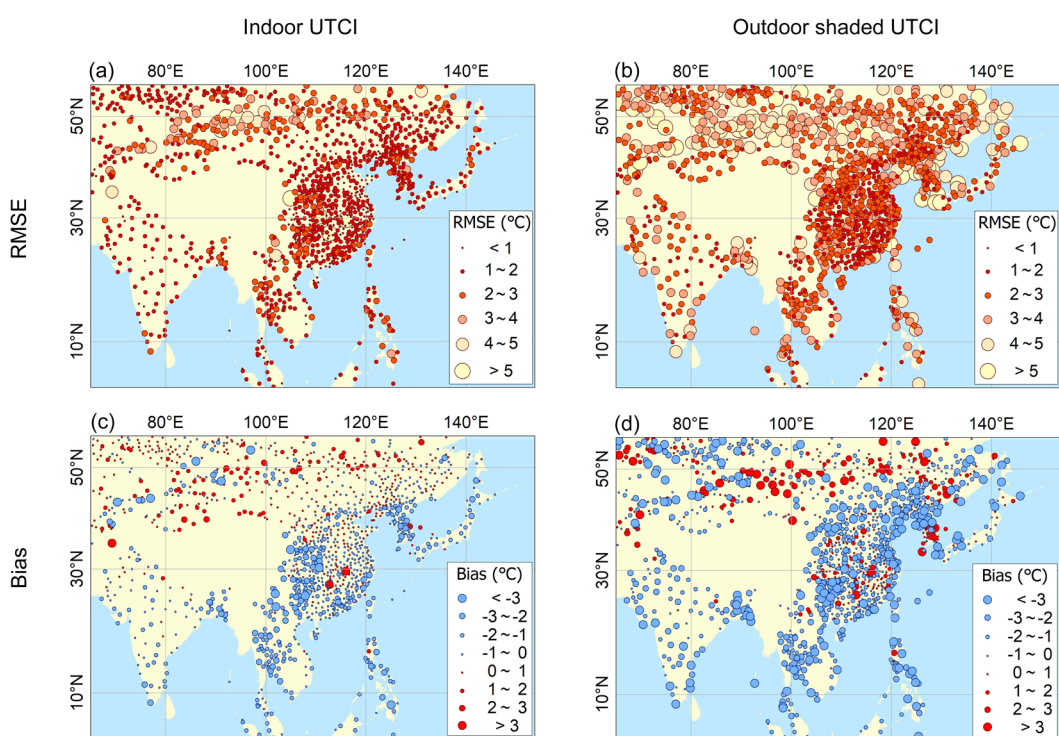


Fig. 2 Spatial distribution of values of RMSE and bias for daily mean indoor UTCI (left column) and outdoor shaded UTCI (right column) computed from ERA5-Land.

RMSE of 9.5 °C with a range of 7.1 °C to 12.1 °C, compared to the MRT (with an RMSE of 8.6 ± 2.5 °C) released along with the UTCI in the ERA5-HEAT product. This is partly because the direct solar radiation, which is the most important radiation variable in determining the MRT, is derived from ERA5, not ERA5-Land. Another reason that leads to the low accuracy of the MRT might be due to the small number of radiation stations used for validation (Table 4). In contrast, the ESI shows strong consistency with the observational result (RMSE values at 7 out of 8 stations are all below 2 °C), which suggests that the outdoor thermal-stress indicator of the ESI is not as sensitive to the change in solar radiation as the UTCI.

Concerning the biases of the three indices listed in Table 4, while the MRT exhibits strong positive biases and the ESI shows slight negative biases at all stations, the UTCI, however, has inconsistent results, with 6 stations positively biased and 2 negatively biased.

Because the accuracy of weather forecasts varies throughout the year, the reliability of this dataset differs in different seasons. Generally, the dataset has a better performance in warm periods and summer monsoon seasons than on cold winter days (Figs. 3 and 4). This is especially true for those indices that include the variables of wind speed or radiation. For example, the RMSE for daily mean values of the outdoor shaded UTCI ranges from the lowest value of 1.9 °C in August to the highest value of 3.5 °C in January. The accuracy of the WCT, which uses wind speed and air temperature for calculation, shows the strongest seasonal effect, with the RMSE for daily

Station ID	Station Name	Longitude	Latitude	Number of Records	MRT		UTCI		ESI	
					RMSE	Bias	RMSE	Bias	RMSE	Bias
54511	Beijing	116.47	39.80	230	10.1	8.1	5.4	3.8	1.0	-0.1
54342	Shenyang	123.52	41.73	283	8.7	4.3	4.5	0.1	1.6	-0.2
50953	Harbin	126.57	45.93	282	11.1	8.0	5.5	2.9	1.5	-0.3
58362	Baoshan	121.45	31.40	289	7.4	3.3	3.2	-0.5	1.2	-0.7
57494	Wuhan	114.05	30.60	284	9.8	5.4	3.8	0.7	1.6	-0.4
59287	Guangzhou	113.48	23.22	288	7.1	3.6	2.9	0.5	1.5	-1.0
56187	Wenjiang	103.87	30.75	289	9.9	2.2	3.9	0.9	1.9	-1.3
51463	Urumqi	87.65	43.78	275	12.1	1.6	6.9	-0.8	3.2	-0.4

Table 4. Average RMSE values ($^{\circ}\text{C}$) and biases ($^{\circ}\text{C}$) of the MRT, UTCI, and ESI for stations that have both radiation data and other commonly observed meteorological data for 2018.

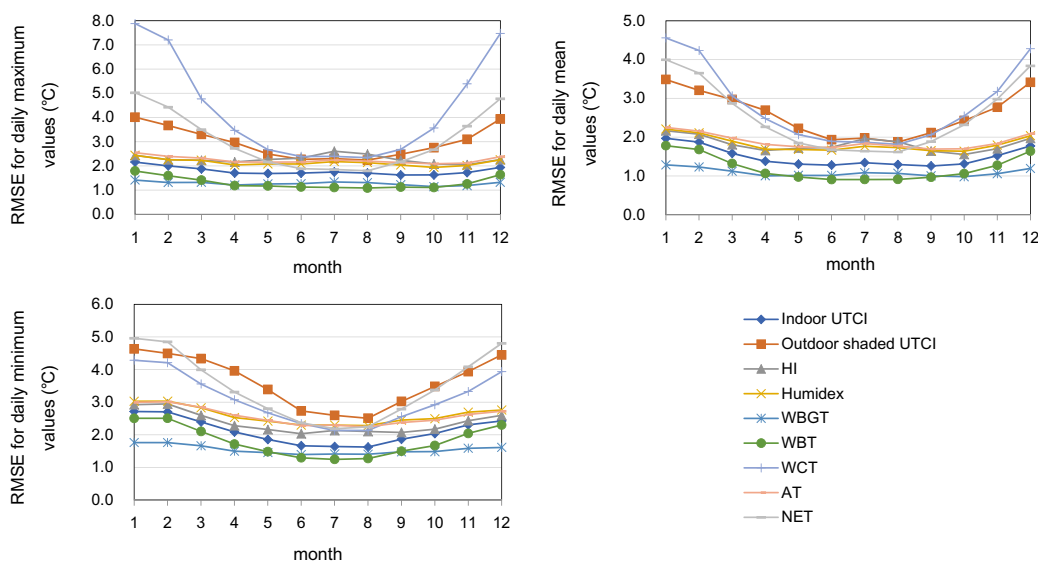


Fig. 3 Average monthly RMSE values for daily maximum (upper left), minimum (lower left), and mean (upper right) thermal-stress indices. This figure only includes the nine indices that don't require radiation as data input.

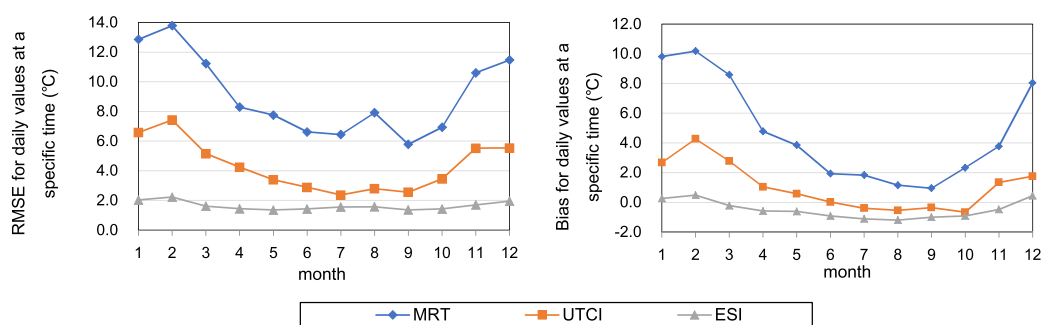


Fig. 4 Average monthly RMSE values (left) and biases (right) for daily values of the MRT, UTCI, and ESI at specific time of the day when maximum global radiation flux occurs.

maximum values varying between 2.4°C and 7.9°C . The accuracy of AT and the other two-variable indices with air temperature and humidity as inputs (i.e., the indoor UTCI, HI, Humidex, WBGT, and WBT), however, exhibits a slight seasonal effect, with RMSE values for the daily mean, maximum and minimum ranging from 1.0°C to 2.3°C , 1.1°C to 2.6°C and 1.3°C to 3.0°C , respectively, in the validation year.

As seen from Figs. 4 and 5, while most of the indices are negatively biased across all seasons, the MRT is positively biased throughout the year, especially in cold winter months. The UTCI is biased towards positive values most of the year except for July to October.

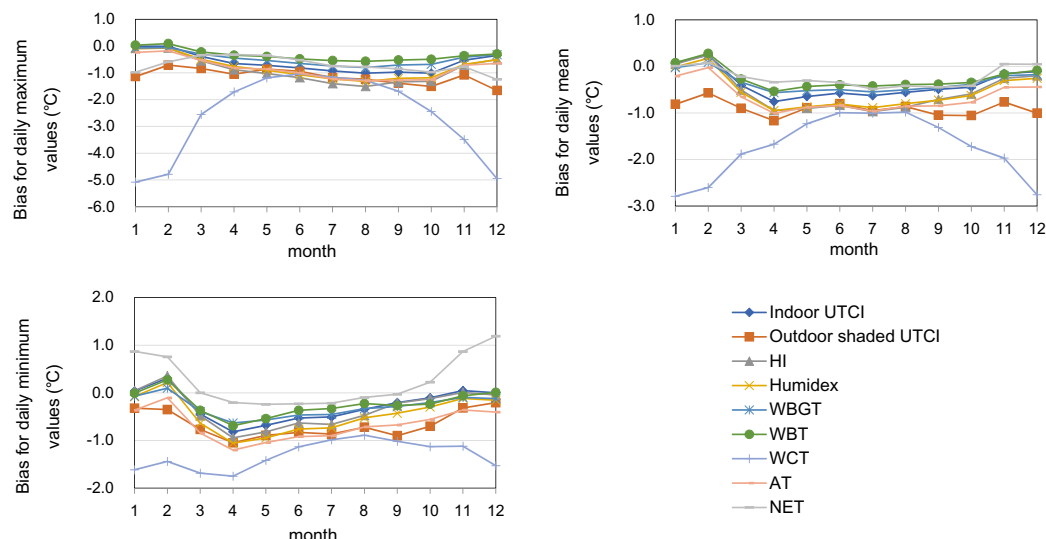


Fig. 5 Average biases for daily maximum (upper left), minimum (lower left), and mean (upper right) thermal-stress indices. This figure only includes the nine indices that don't require radiation as data input.

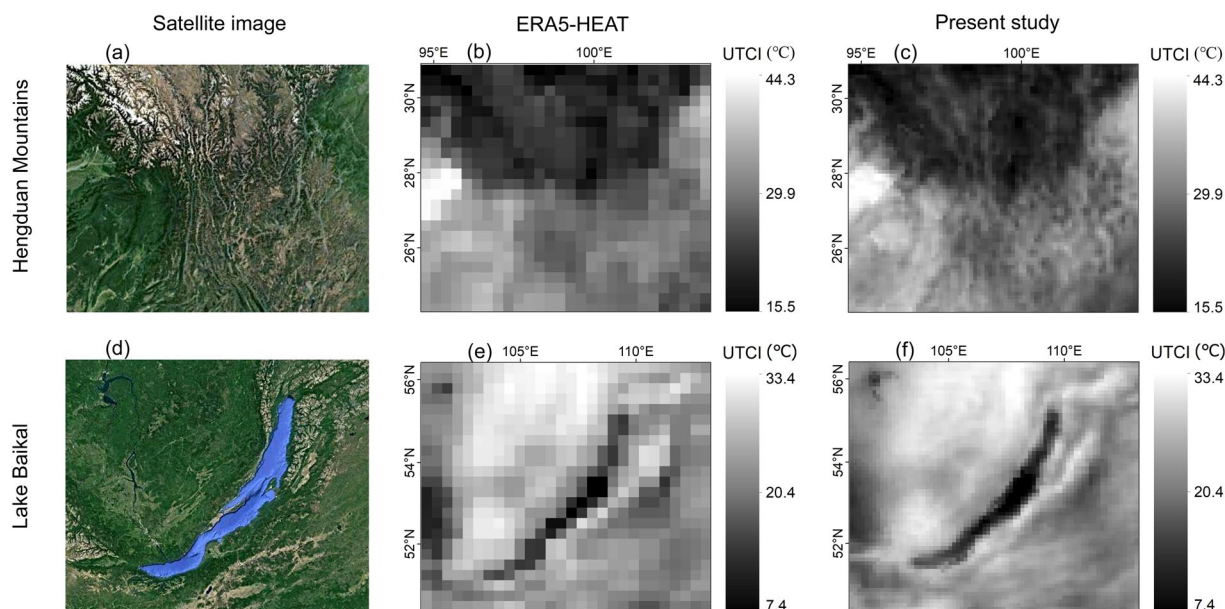


Fig. 6 The satellite images from Google Earth for the regions of Hengduan Mountains (upper left) and Lake Baikal (lower left), and the distributions of daily maximum UTCI from ERA5-HEAT (middle) and the present study (right) on 2018-07-20.

To enable users to learn more about the seasonal effects of dataset accuracy at individual weather stations, we created text-formatted validation files (archived and named “validation.tar.gz”, available at the abovementioned repository) in which the monthly and yearly summaries of RMSE and bias at each station, as well as their locations, are included. With these data, users can reduce uncertainties by examining the verification results at stations located in their study areas.

Usage Notes

In comparison to the existing $0.25^\circ \times 0.25^\circ$ spatial resolution thermal-index product¹⁸, this dataset provides more details on studying the spatial variation of heat/cold stress. As seen from the upper images in Fig. 6, the $0.1^\circ \times 0.1^\circ$ gridded UTCI allows us to quantify the difference between the human thermal stress in longitudinal valleys and their associated mountain ridges in Southwest China. The lower images of Fig. 6 show that while the spatial contrast of UTCI near Lake Baikal is blurred in the $0.25^\circ \times 0.25^\circ$ gridded product (downloaded from the Copernicus Climate Data Store implemented by ECMWF), more detailed information, especially along the lakeshore, is visible in our new dataset.

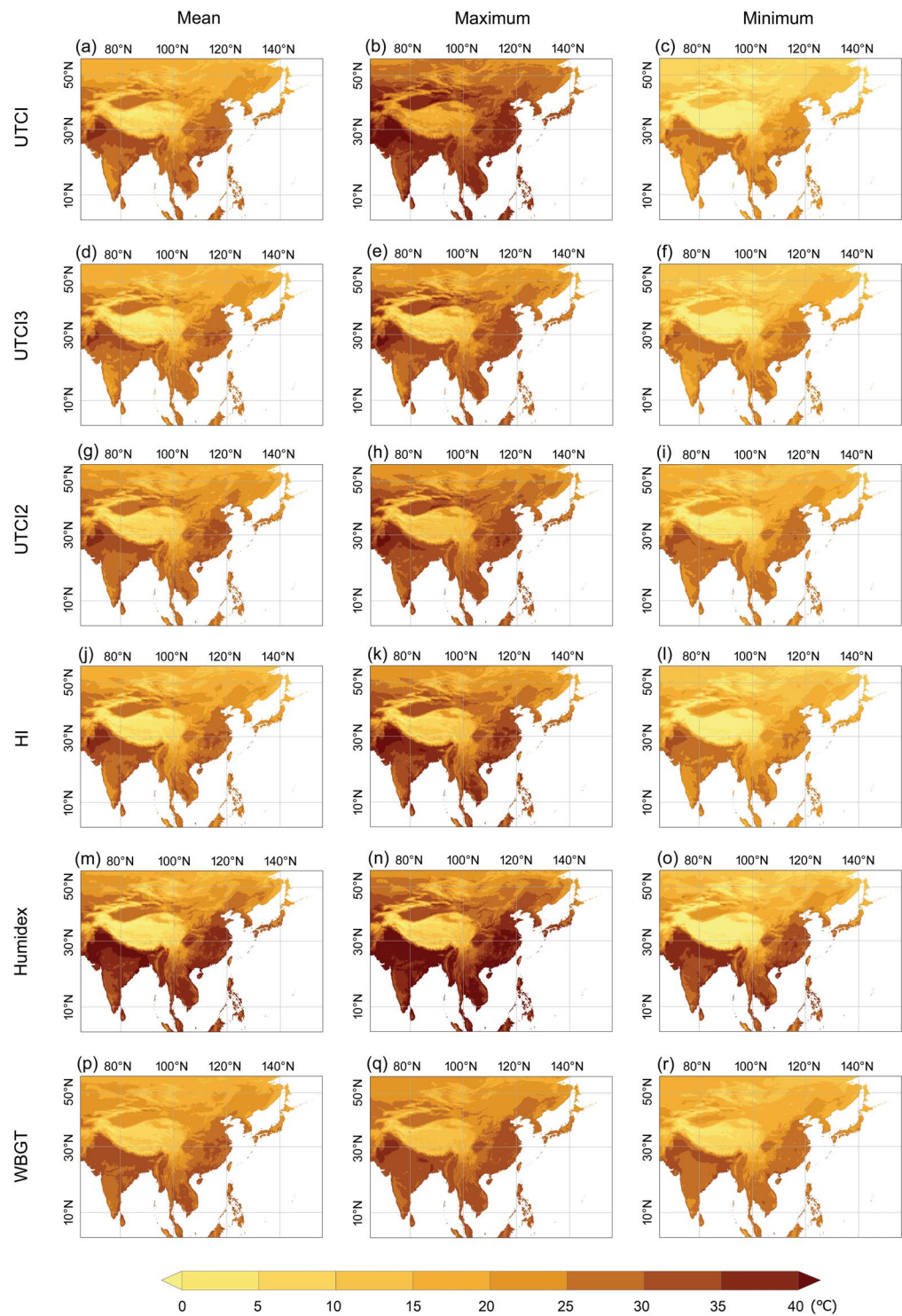


Fig. 7 Averaged daily mean (left column), maximum (middle column), and minimum (right column) of the thermal indices for July during the period of 1981 - 2019. Only select indices suitable for hot conditions are illustrated. UTCI2 refers to the indoor UTCI, which uses two parameters, and UTCI3 stands for the outdoor shaded UTCI, which takes three parameters for the calculation.

Combined with heat- or cold-related morbidity and mortality, this dataset can be used to identify thermal stress thresholds for the general population or specific groups working indoors or outdoors. This dataset can also serve to assess the thermal comfort conditions required for tourism activities directly exposed under the sun or in the shade.

Although all thermal indices used in this study are temperature equivalents expressed in degrees Celsius (note that a conversion from Fahrenheit to Celsius for the index of HI is performed) and share a similar spatial pattern

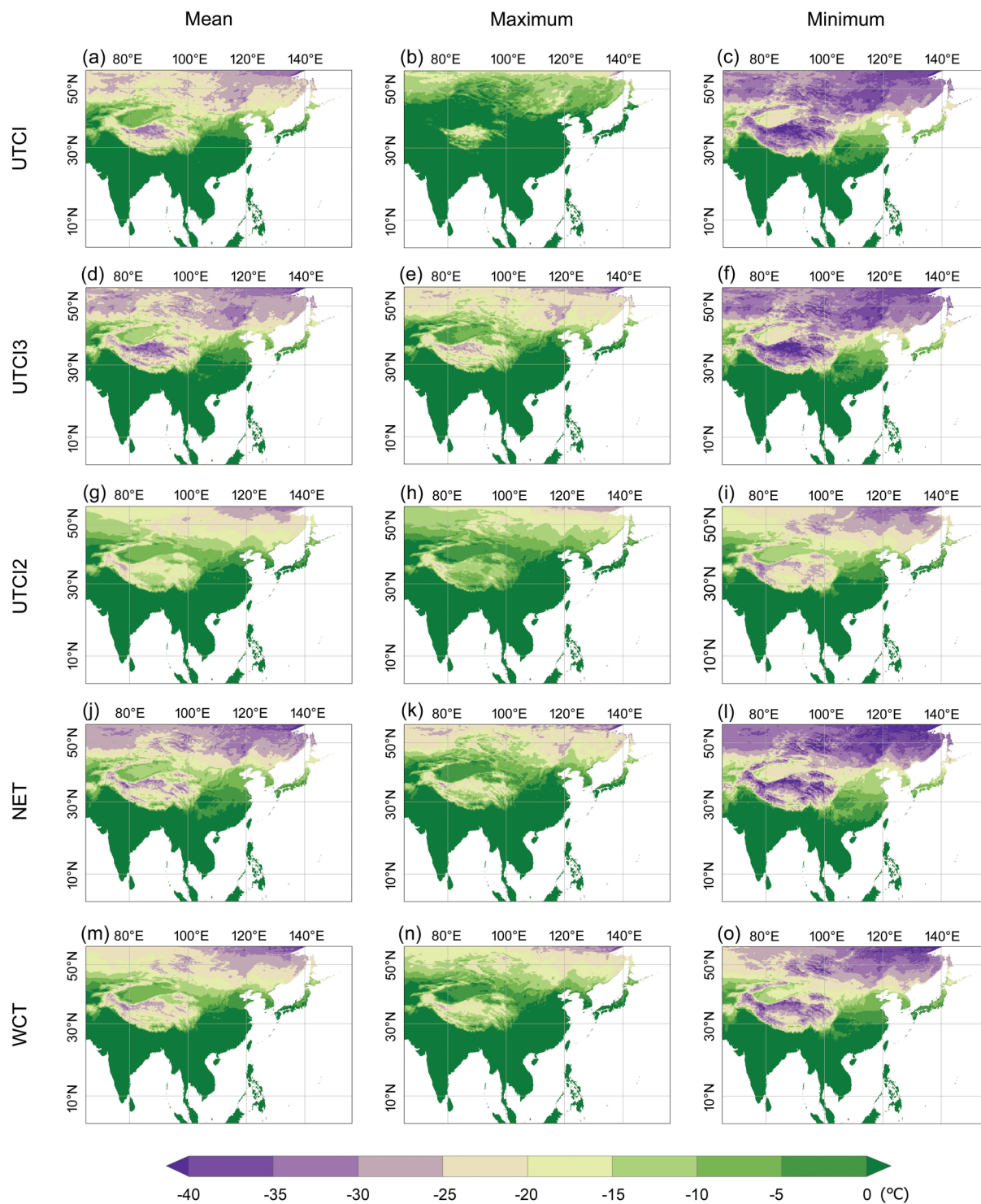


Fig. 8 Averaged daily mean (left column), maximum (middle column), and minimum (right column) of the thermal indices for January during the period 1981 - 2019. Only essential indices suitable for cold conditions are illustrated. UTCI2 refers to the indoor UTCI, which uses two parameters, and UTCI3 stands for the outdoor shaded UTCI, which takes three parameters for the calculation.

(Figs. 7 and 8), it is worth noting that each index is associated with a particular assessment scale. For example, UTCI values between 32 °C and 38 °C are categorized as “strong heat stress”³⁵, whereas for Humidex, a similar sensation would range from 40 °C to 45 °C⁴³. A comprehensive description of assessment scales with defined thresholds for commonly used thermal indices was provided by Blazejczyk *et al.*¹¹.

Another important note is that while the UTCI can be applied in all climates and all seasons throughout the year, the use of the other indices is often restricted to specific conditions. For example, two-variable indices (with

air temperature and humidity as inputs), such as the indoor UTCI, HI, Humidex, WBGT and WBT, are suitable for use in indoor conditions, while three-variable indices, such as the outdoor shaded UTCI, AT and NET, can be applied in an outdoor shaded environment, as the effect of wind speed is accounted for.

While this dataset shows higher accuracy in flat areas (e.g., the Indo-Gangetic Plain and the lowland plains in eastern China, as shown in Fig. 2), its accuracy degrades in areas with heterogeneous landscapes, especially in mountainous areas (e.g., western mountainous areas of China), with strong orographic effects and coastal zones affected by the mixed-pixel problem (e.g., areas along the coastline of the Korean Peninsula where land and water coexist within specific grid cells). Researchers and practitioners interested in those regions might have to pay more attention, as thermal-stress indices may vary substantially due to complex topography or land-water contrasts at a subgrid scale.

Code availability

All codes for calculating the indoor and outdoor UTCI, MRT, and other empirical thermal indices, written in Python (3.8) using `cdsapi` (0.3.1), `numpy` (1.19.2), `pandas` (1.1.3), `netCDF4` (1.5.4), and `scipy` (1.5.3) libraries, were developed on Linux (CentOS 6.10) and can be easily adapted to Windows and other platforms. The codes are freely available at the abovementioned repository⁵².

Received: 13 November 2020; Accepted: 2 August 2021;

Published online: 01 September 2021

References

- Della-Marta, P. M. *et al.* Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability. *Clim. Dyn.* **29**, 251–275 (2007).
- Steffen, W., Hughes, L. & Perkins, S. *Heatwaves: hotter, longer, more often* <https://www.climatecouncil.org.au/heatwaves-report> (The Climate Council of Australia, 2014).
- Xu, Y. *et al.* Substantial increase in the joint occurrence and human exposure of heatwave and high-PM hazards over South Asia in the mid-21st century. *AGU Adv.* **1**, e2019AV000103 (2020).
- Gao, C., Kuklane, K., Östergren, P. O. & Kjellström, T. Occupational heat stress assessment and protective strategies in the context of climate change. *Int. J. Biometeorol.* **62**, 359–371 (2018).
- Di Napoli, C., Pappenberger, F. & Cloke, H. L. Assessing Heat-Related Health Risk in Europe via the Universal Thermal Climate Index (UTCI). *Int. J. Biometeorol.* **62**, 1155–1165 (2018).
- Desai, M. S. & Dhorde, A. G. Trends in thermal discomfort indices over western coastal cities of India. *Theor. Appl. Climatol.* **131**, 1305–1321 (2018).
- Kong, Q., Zheng, J., Fowler, H. J., Ge, Q. & Xi, J. Climate change and summer thermal comfort in China. *Theor. Appl. Climatol.* **137**, 1077–1088 (2019).
- Åström, D. O., Forsberg, B., Ebi, K. & Rocklöv, J. Attributing mortality from extreme temperatures to climate change in Stockholm, Sweden. *Nat. Clim. Chang.* **3**, 1050–1054 (2013).
- Azhar, G. S. *et al.* Heat-related mortality in India: Excess all-cause mortality associated with the 2010 Ahmedabad heat wave. *PLoS One* **9**, e91831 (2014).
- Mora, C. *et al.* Global risk of deadly heat. *Nat. Clim. Chang.* **7**, 501–506 (2017).
- Blazejczyk, K., Epstein, Y., Jendritzky, G., Staiger, H. & Tinz, B. Comparison of UTCI to selected thermal indices. *Int. J. Biometeorol.* **56**, 515–535 (2012).
- De Freitas, C. R. & Grigorieva, E. A. A comprehensive catalogue and classification of human thermal climate indices. *Int. J. Biometeorol.* **59**, 109–120 (2015).
- Giannaros, T. M., Lagouvardos, K., Kotroni, V. & Matzarakis, A. Operational forecasting of human-biometeorological conditions. *Int. J. Biometeorol.* **62**, 1339–1343 (2018).
- Jendritzky, G., de Dear, R. & Havenith, G. UTCI—Why another thermal index? *Int. J. Biometeorol.* **56**, 421–428 (2012).
- Fiala, D., Havenith, G., Bröde, P., Kampmann, B. & Jendritzky, G. UTCI-Fiala multi-node model of human heat transfer and temperature regulation. *Int. J. Biometeorol.* **56**, 429–441 (2012).
- Havenith, G. *et al.* The UTCI-clothing model. *Int. J. Biometeorol.* **56**, 461–470 (2012).
- McGregor, G. R. Special issue: Universal Thermal Comfort Index (UTCI). *Int. J. Biometeorol.* **56**, 419 (2012).
- 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. *Geosci. Data J.* <https://doi.org/10.1002/gdj3.102> (2020).
- Mistry, M.N. A high spatiotemporal resolution global gridded dataset of historical human discomfort indices. *Atmosphere* <https://doi.org/10.3390/atmos11080835> (2020).
- Sustainable Energy for All (SEforALL). *Chilling Prospects: Tracking Sustainable Cooling for All* <https://www.seforall.org/chilling-prospects-2020> (2020).
- Hersbach, H. *et al.* The ERA5 global reanalysis. *Quart. J. Roy. Met. Soc.* **146**, 1999–2049 (2020).
- Hersbach, H. *et al.* ERA5 hourly data on single levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* <https://doi.org/10.24381/cds.adbb2d47> (2018).
- Muñoz Sabater, J. ERA5-Land hourly data from 1981 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* <https://doi.org/10.24381/cds.e2161bac> (2019).
- Environmental Systems Research Institute (ESRI). *ArcGIS 10.3.1 for Desktop Online Help* <https://desktop.arcgis.com/en/arcmap/10.3/tools/data-management-toolbox/resample.htm> (2020).
- Copernicus Climate Change Service (C3S) Climate Data Store (CDS). *ERA5-Land: data documentation* <https://confluence.ecmwf.int/display/CKB/ERA5-Land%3A+data+documentation> (2019).
- American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). *2001 ASHRAE Handbook: Fundamentals* (ASHRAE, 2001).
- Matzarakis, A., Rutz, F. & Mayer, H. Modelling radiation fluxes in simple and complex environments—application of the RayMan model. *Int. J. Biometeorol.* **51**, 323–334 (2007).
- Weih, P. *et al.* The uncertainty of UTCI due to uncertainties in the determination of radiation fluxes derived from measured and observed meteorological data. *Int. J. Biometeorol.* **56**, 537–555 (2012).
- Jendritzky, G. Bioklimatische Bewertungsgrundlage der Räume am Beispiel von mesoskaligen Bioklimakarten. In: *Methodik zur räumlichen Bewertung der thermischen Komponente im Bioklima des Menschen* Vol. **114** (ed. Schirmer, H.) (Akademie für Raumforschung und Landesplanung, 1990).
- Leroyer, S., Bélaïr, S., Spacek, L. & Gultepe, I. Modelling of radiation-based thermal stress indicators for urban numerical weather prediction. *Urban Clim.* **25**, 64–81 (2018).

31. Woan, G. Astrophysics. In: *The Cambridge handbook of physics formulas* (Cambridge University Press, 2000).
32. Spencer, J. W. Fourier series representation of the position of the sun. *Search* **2**, 162–172 (1971).
33. National Oceanic and Atmospheric Administration (NOAA). NOAA Global Vegetation Index User's Guide APPENDIX L: software to calculate relative azimuth from third generation weekly composite GVI date <http://www2.ncdc.noaa.gov/docs/gviug/html/l/app-l.htm> (1997).
34. Di Napoli, C., Hogan, R. J. & Pappenberger, F. Mean radiant temperature from global-scale numerical weather prediction models. *Int. J. Biometeorol.* **64**, 1233–1245 (2020).
35. Bröde, P. *et al.* Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). *Int. J. Biometeorol.* **56**, 481–494 (2012).
36. Steadman, R. G. A universal scale of apparent temperature. *J. Appl. Meteorol. Climatol.* **23**, 1674–1687 (1984).
37. Steadman, R. G. Norms of apparent temperature in Australia. *Aust. Met. Mag.* **43**, 1–16 (1994).
38. Moran, D. S. *et al.* An environmental stress index (ESI) as a substitute for the wet bulb globe temperature (WBGT). *J. Therm. Biol.* **26**, 427–431 (2001).
39. Moran, D. S. & Epstein, Y. Evaluation of the environmental stress index (ESI) for hot/dry and hot/wet climates. *Ind. Health* **44**, 399–403 (2006).
40. Steadman, R. G. The assessment of sultriness. part I: a temperature-humidity index based on human physiology and clothing science. *J. Appl. Meteor.* **18**, 861–873 (1979).
41. Rothfusz, L.P. The heat index equation. *National Weather Service Technical Attachment*. Report No. SR 90–23 (1990).
42. National Oceanic and Atmospheric Administration (NOAA). *The Heat Index Equation* https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml (2020).
43. Masterson, J. & Richardson, F.A. *Humidex: a method of quantifying human discomfort due to excessive heat and humidity* (Environment Canada, 1979).
44. Houghton, F. C. & Yaglou, C. P. Determining equal comfort lines. *J. Am. Soc. Heat. & Vent. Engrs.* **29**, 165–176 (1923).
45. Missenard, F.A. Température effective d'une atmosphère Généralisation température résultante d'un milieu. In: *Encyclopédie Industrielle et Commerciale, Etude physiologique et technique de la ventilation* (Librerie de l'Enseignement Technique, 1933).
46. Landsberg HE. *The assessment of human bioclimate: a limited review of physical parameters*. Technical Note No. 123, WMO-No. 331 (World Meteorological Organization, 1972).
47. Hentschel, G. A human biometeorology classification of climate for large and local scales. In: *Proceeding of WMO/HMO/UNEP symposium on climate and human health*. WCPA-No.1 (World Meteorological Organization, 1987).
48. Yaglou, C. P. & Minard, D. Control of heat casualties at military training centers. *AMA Arch. Ind. Health* **16**, 302–316 (1957).
49. Australian Bureau of Meteorology. *Thermal comfort observations* http://bom.gov.au/info/thermal_stress/ (2020).
50. Stull, R. Wet-bulb temperature from relative humidity and air temperature. *J. Appl. Meteorol. Climatol.* **50**, 2267–2269 (2011).
51. Office of the Federal Coordinator for Meteorological services and supporting research (OFCM). *Report on Wind Chill Temperature and extreme heat indices: evaluation and improvement projects*. Report No. FCM-R19-2003 (U.S. Office of the Federal Coordinator for Meteorological Services and Supporting Research, 2003).
52. Yan, Y., Xu, Y. & Yue, S. A high-spatial-resolution dataset of human thermal stress indices over South and East Asia. *figshare* <https://doi.org/10.6084/m9.figshare.c.5196296> (2021).
53. China Meteorological Data Service Center (CMDSC) <https://data.cma.cn/en/> (CMDSC, 2020).
54. National Demonstration Center for Experimental Atmospheric Science and Environmental Meteorology Education. <https://etcme.nuist.edu.cn/> (NUIST, 2020).
55. Blazejczyk, K. BioKlima—Universal tool for bioclimatic and thermophysiological studies <https://www.igipz.pan.pl/BioKlima-zgik.html> (2010).

Acknowledgements

The first author was funded by Jiangsu Overseas Visiting Scholar Program for University Prominent Young & Middle-aged Teachers and Presidents. Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing. ERA5-Land and ERA5 are made available at the Copernicus Climate Change Service^{22,23}.

Author contributions

Yechao Yan conceived the research idea and wrote the paper with Yangyang Xu. Yechao Yan wrote the Python scripts that generated the datasets. Shuping Yue produced the figures.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.Y.

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

© The Author(s) 2021