Supporting Information

Climate Change will Accelerate the High-End Risk of Compound Drought and Heatwave Events

Kumar Puran Tripathy[1], Sourav Mukherjee[1], *Ashok Kumar Mishra[1], Michael E. Mann[2], A. Park Williams[3,4]

[1]Glenn Department of Civil Engineering, Clemson University, South Carolina, USA.

[2] Department of Earth & Environmental Science, University of Pennsylvania, PA, USA.

[3]Department of Geography, University of California, Los Angeles, CA 90095

[4]Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10096

Corresponding author: Ashok Kumar Mishra (ashokm@g.clemson.edu)

Appendix A

A1. Nonstationary Bias Correction

The daily precipitation and daily temperature (T_{max} and T_{min}) GCM datasets need to be bias corrected to obtain more robust and reliable estimates for the future. The present paper used the non-stationary bias correction technique – updated nonstationary CDF matching (CNCDFm) method, developed by Miao et al. (2016)[1]. This method is an improvement over the traditional quantile mapping approach to correct the GCM outputs. The CNCDFm method, which is a combination of EDCDFm (Equidistant CDF matching) and equiratio CDFm[2], treats temperature and precipitation data separately, thus avoiding the problem of obtaining unreasonably high values of daily precipitation values.

The bias-correction for temperature data can be mathematically written as,

$$\tilde{x}_{m-p.adjust} = x_{m-p} + F_{o-c}^{-1}\left(F_{m-p}\left(x_{m-p}\right)\right) - F_{m-c}^{-1}(F_{m-p}(x_{m-p})) \tag{1}$$

The bias correction for precipitation data is as follows,

$$\tilde{x}_{m-p.adjust} = \begin{cases} I(x) & if\ I(x) > 0 \\ g(x) & if\ I(x) < 0 \\ 0 & if\ x_{m-p} = 0 \end{cases} \tag{2}$$

$$I(x) = x_{m-p} + F_{o-c}^{-1}\left(F_{m-p}\left(x_{m-p}\right)\right) - F_{m-c}^{-1}(F_{m-p}(x_{m-p})) \tag{3}$$

and,

$$I(x) = x_{m-p} \times \frac{F_{o-c}^{-1}\left(F_{m-p}\left(x_{m-p}\right)\right)}{F_{m-c}^{-1}(F_{m-p}(x_{m-p}))} \tag{4}$$

Where, x is the meteorological variable (temperature or precipitation) of interest for observation (o) or model (m) corresponding to historical observation period or current climate (c), or for projected future period (p). F_{m-p} indicates the CDF of the model for future projection (Socio-economic Pathways), and F_{o-c}^{-1} and F_{m-c}^{-1} indicate the inverse CDF functions (quantile mapping functions) for observation and model scenarios for the recent observed period (1982-2019) respectively.

A2. Estimation of CDHW severity

The estimation of CDHW severity in this paper is followed from the previous work by Mukherjee and Mishra (2020)[3]. The severity of a particular CDHW day is calculated by taking the product of the daily standardized values of maximum temperature (standardized w.r.t. the interquartile range of Tmax for the summer period of that year) and the sc-PDSI value corresponding to that week ($scPDSI_w$). Then, the severity of a CDHW event is calculated by taking the cumulative sum of daily severity for each of the CDHW days corresponding to that CDHW event, which is given in the equation as follows,

$$CDHWs_i = \sum_{d=1}^{d=D_i}\left(\left(-1 \times scPDSI_{w,i}\right) \times \left(\frac{T\,max_{d,i} - T_{25p}}{T_{75p} - T_{25p}}\right)\right); D_i \geq 3, d \in w \qquad (5)$$

Where, D_i is the number of CDHW days in the i^{th} CDHW event, d indicates the heatwave days falling inside the drought week (w). $T_{\max d,i}$ indicates the maximum temperature for that day, T_{25p} and T_{75p} are the 25^{th} and 75^{th} percentile of daily T_{max} during the summer period, respectively. It is important to know that the CDHW severity is dimensionless as it is the product of scPDSI (a standardized entity) and standardized maximum temperature.

Then the mean annual CDHW severity ('severity' in the main text) is calculated by taking the average of events within the year. For example – if 3 CDHW events are observed with the magnitude of severities 30, 40, and 50, then the mean annual CDHW severity is 40.

$$Severity = \begin{cases} \frac{\left(\sum_{i=1}^{i=N} CDHWs_i\right)}{N} & if N \neq 0 \\ 0 & if N = 0 \end{cases} \tag{6}$$

Where, N is the total number of CDHW events observed for that particular year.

A3. Calculation of area of grids

The areas of the rectangular grids gradually decreases as we move from the equator to the higher latitudes in both North and South direction. Therefore, an area correction is applied when calculations involve the area of grids, for example, in calculating global mean of CDHW characteristics and areal thresholds in Figure 1 (main manuscript). The present study calculates the area of each rectangular $2^o \times 2^o$ latitude-longitude grid, surrounded by $[lon_1, lon_2, lat_1, lat_2]$ using the following mathematical expression (https://www.pmel.noaa.gov/maillists/tmap/ferret_users/fu_2004/msg00023.html),

$$A = \left(\frac{\pi \times R^2}{180}\right) \times |\sin(lat_1) - \sin(lat_2)| \times |lon_1 - lon_2| \tag{7}$$

Where, $R = 6400\ kms$ is the equatorial radius of the earth.

In this paper, we have a total of 3347 global land area grids. After calculating the area of each grid, the area of equatorial grid is divided to the area of every grid, which gives an area correction factor (γ) such that $\gamma \epsilon (0,1]$. γ is equal to one for the equatorial grids since they have the maximum area and for higher latitudes γ decreases.

A4. Trend Analysis

The interannual trend present in the different CDHW characteristics are evaluated using Sen's slope estimator. Sen's slope estimate can be easily evaluated using the *scipy* library (https://docs.scipy.org/doc/scipy-1.7.0/reference/) of *python 3.0*. The linear trends present are tested for significance based on Mann-Kendall's trend test. The detail descriptions of Sen's Slope estimation and Mann-Kendall test for significance are given in the following subsections.

A2.1. Calculation of Sen's Slope Estimator

Theil-Sen, (1968)[4] developed a non-parametric method to find the true slope (change per unit time) present in univariate time series.

1. Slopes of the trend in a sample of N pairs of data are calculated as follows:

$$Q_k = \frac{y_i - y_j}{x_i - x_j} \text{ for } k = 1, 2, \dots, N \tag{8}$$

Where, (x_i, y_i) *and* (x_j, y_j) are the data pairs out of the N pairs of data given that $j > i$.

2. Sen's slope estimate is calculated as the median of the N values of Q_i as follows:

$$Q_{med} = \begin{cases} Q_{\left[\frac{N+1}{2}\right]}, & if \quad N \quad is \quad odd \\ \frac{1}{2}\left(Q_{\left[\frac{N}{2}\right]} + Q_{\left[\frac{N+2}{2}\right]}\right), & if \quad N \quad is \quad even \end{cases} \tag{9}$$

A2.2. Mann-Kendall's test

Mann-Kendall's (MK) test[5] is a non-parametric test that statistically assesses if there is a monotonic upward or downward trend in the variable of interest over time. The MK test checks whether to reject the null hypothesis H_o and accept the alternative hypothesis H_a,

where, $H_o : No\ Monotonic\ trend\ present\ in\ the\ series$ vs.

$H_a : There\ is\ a\ monotonic\ trend\ present\ in\ the\ series$

The Mann-Kendall's test statistic to test for significance is given by,

$$Z_{MK} = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & if \quad S > 0 \\ 0 & if \quad S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & if \quad S < 0 \end{cases} \tag{10}$$

Z_{MK} is obtained from standard normal distribution.

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_i) \tag{11}$$

$$sgn(x_j - x_i) = \begin{cases} +1, if x_j - x_i > 0 \\ 0, if x_j - x_i = 0 \\ -1 if x_j - x_i < 0 \end{cases} \tag{12}$$

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{t=1}^{m} t_i(t_i - 1)(2t_i + 5)}{18} \tag{13}$$

Where, n is the length of the dataset, m is the number of tied groups, and t indicates the total number of ties. Positive (or, negative) values of Z indicate increasing (or, decreasing) trends present in the time series. The null hypothesis (H_o) of no trend is rejected for of $Z_{MK} \geq z_{MK_{(1-\frac{\alpha}{2})}}$. The present study considers a significance level of $\alpha = 0.01$.

A5. Calculation of Areal Thresholds

First, we performed the area corrections of the grids provided in section A3. Before calculating the areal thresholds, we calculated the yearly time series of CDHW frequency, duration, and severity from 1982-2099 (118 years) for all the grids over the global land areas. Then, the area-correction factor (γ) of each grid is multiplied with the respective CDHW characteristics. For this study, we have a total of 3347 grid points. Thus, we have a 118-by-3347 matrix; rows represent the time, and columns represent the grid points.

For the calculation of areal thresholds for any CDHW characteristics, follow these steps:

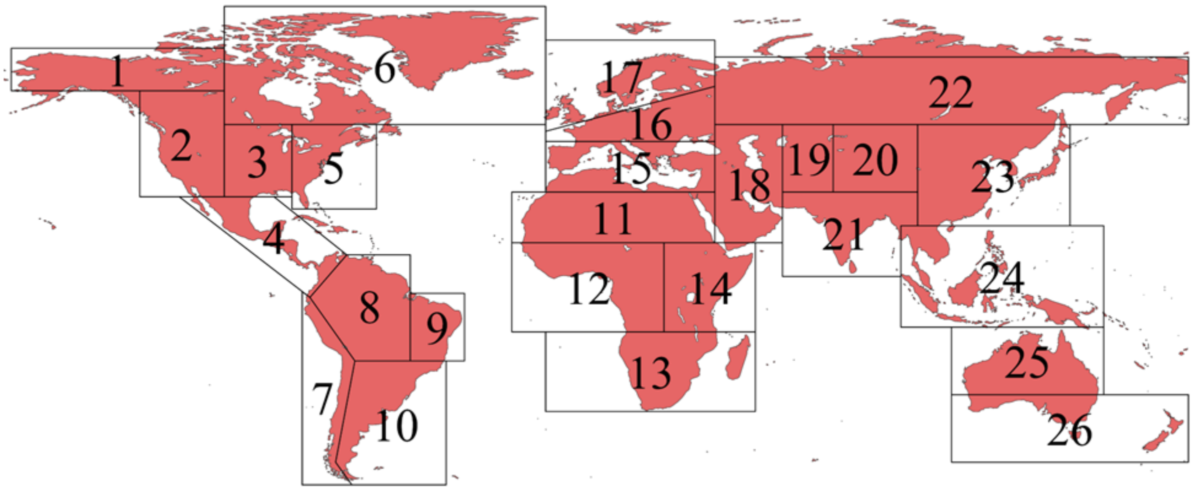
(1) Calculate the 20th, 40th, 60th, and 80th percentiles for the first row (i.e., for the year 1982). Here, the 80th percentile value represents that at least 20% of the global land areas are witnessing CDHW characteristic (frequency, duration, or severity) greater than that value.

(2) Calculate these values for all 118 years. Now, we get a time series for all the four percentiles, plotted in Figure 1 (d-f; main text).

A6. Area vs. frequency curves

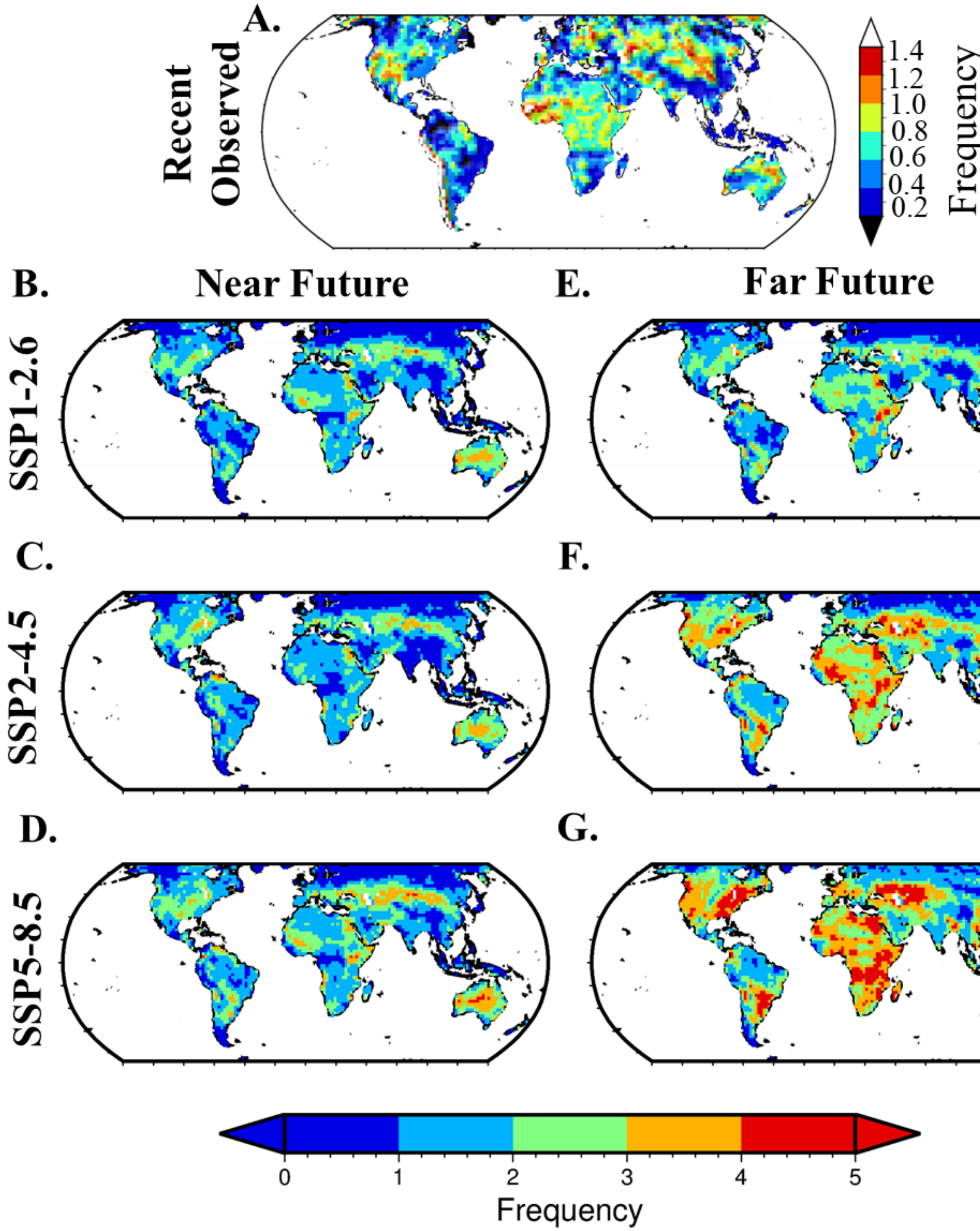
1. For a climate division, the return periods for all the grid points are first segregated into *N* (a minimum value of 50 is recommended) number of bins. This is done using the "*histcount"* command in MATLAB.
2. Then we calculated the number of grid points having RP more than or equal to a particular RP value by taking the sum of histogram frequency.
3. The percentage of area affected is calculated by dividing the total number of grid points, and these values are smoothened using a 7- year running mean.

Figures



1. Alaska (ALA)
2. West North America (WNA)
3. Central North America (CNA)
4. Central America/Mexico (CAM)
5. Canada/Greenland/Iceland (CGI)
6. East North America (ENA)
7. West Coast South America (WSA)
8. Amazon (AMZ)
9. North-East Brazil (NEB)
10. Southeastern South America (SSA)
11. Sahara (SAH)
12. West Africa (WAF)
13. Southern Africa (SAF)
14. East Africa (EAF)
15. Mediterranean (MED)
16. Central Europe (CEU)
17. North Europe (NEU)
18. West Asia (WAS)
19. Central Asia (CAS)
20. Tibetan Plateau (TIB)
21. South Asia (SAS)
22. North Asia (NAS)
23. East Asia (EAS)
24. Southeast Asia (SEA)
25. North Australia (NAU)
26. South Australia/New Zealand (SAU)

Figure S1 *The spatial boundaries of the twenty-six climate divisions based on IPCC AR5 classification.*



155

156 ***Figure S2*** *Spatial map showing the frequency CDHW events for (a) recent observed period (1982-2020), (b) SSP1-2.6 near-future (2021-2057), (c) SSP2-4.5 near-future (2021-2057), (d) SSP5-8.5 near-future (2021-2057), (e) SSP1-2.6 far-future (2058-2095), (f) SSP2-4.5 far-future (2058-2095), (g) SSP5-8.5 far-future (2058-2095).*

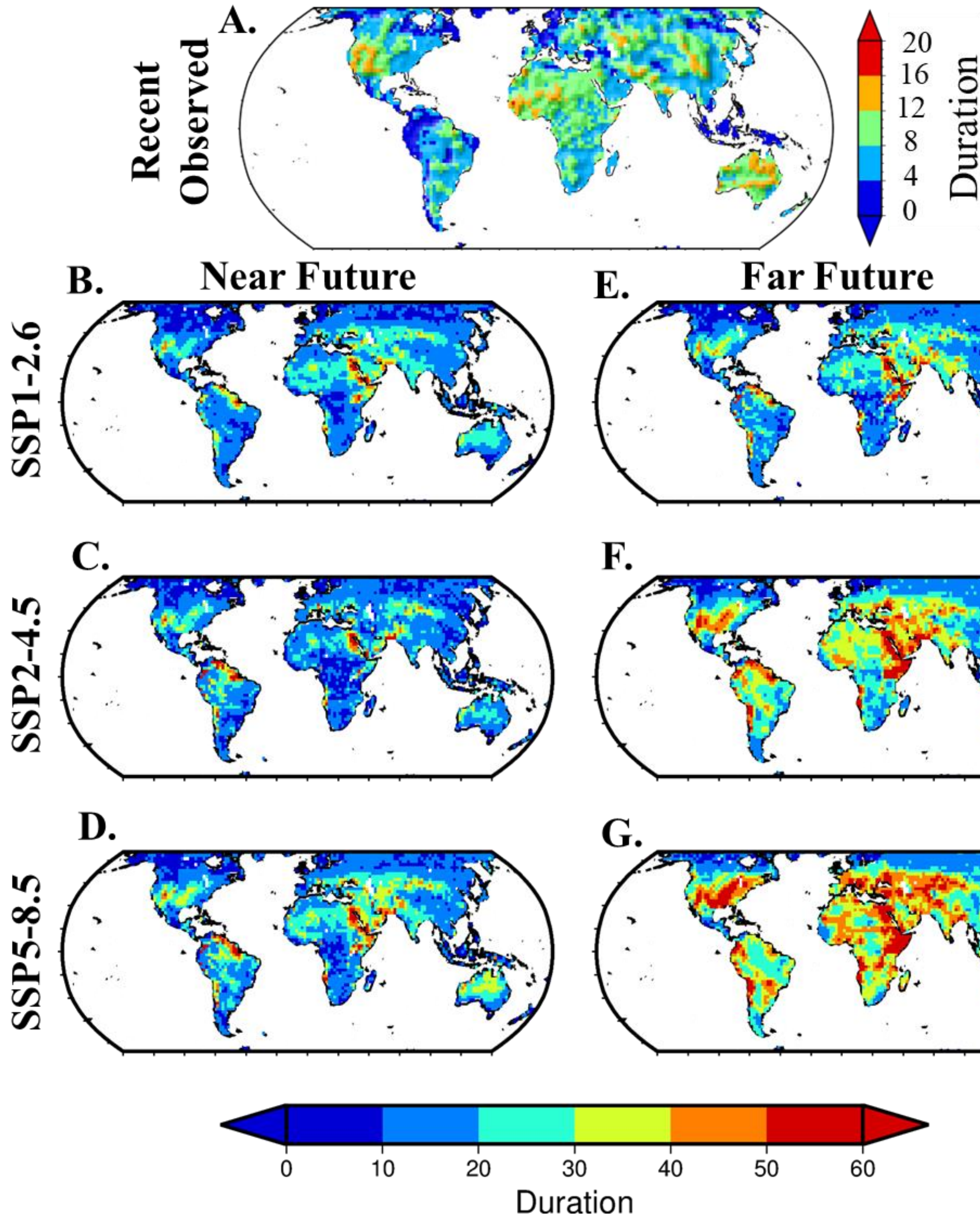


Figure S3 *Spatial map showing the duration of CDHW events for (a) recent observed period (1982-2020), (b) SSP1-2.6 near-future (2021-2057), (c) SSP2-4.5 near-future (2021-2057), (d) SSP5-8.5 near-future (2021-2057), (e) SSP1-2.6 far-future (2058-2095), (f) SSP2-4.5 far-future (2058-2095), (g) SSP5-8.5 far-future (2058-2095).*

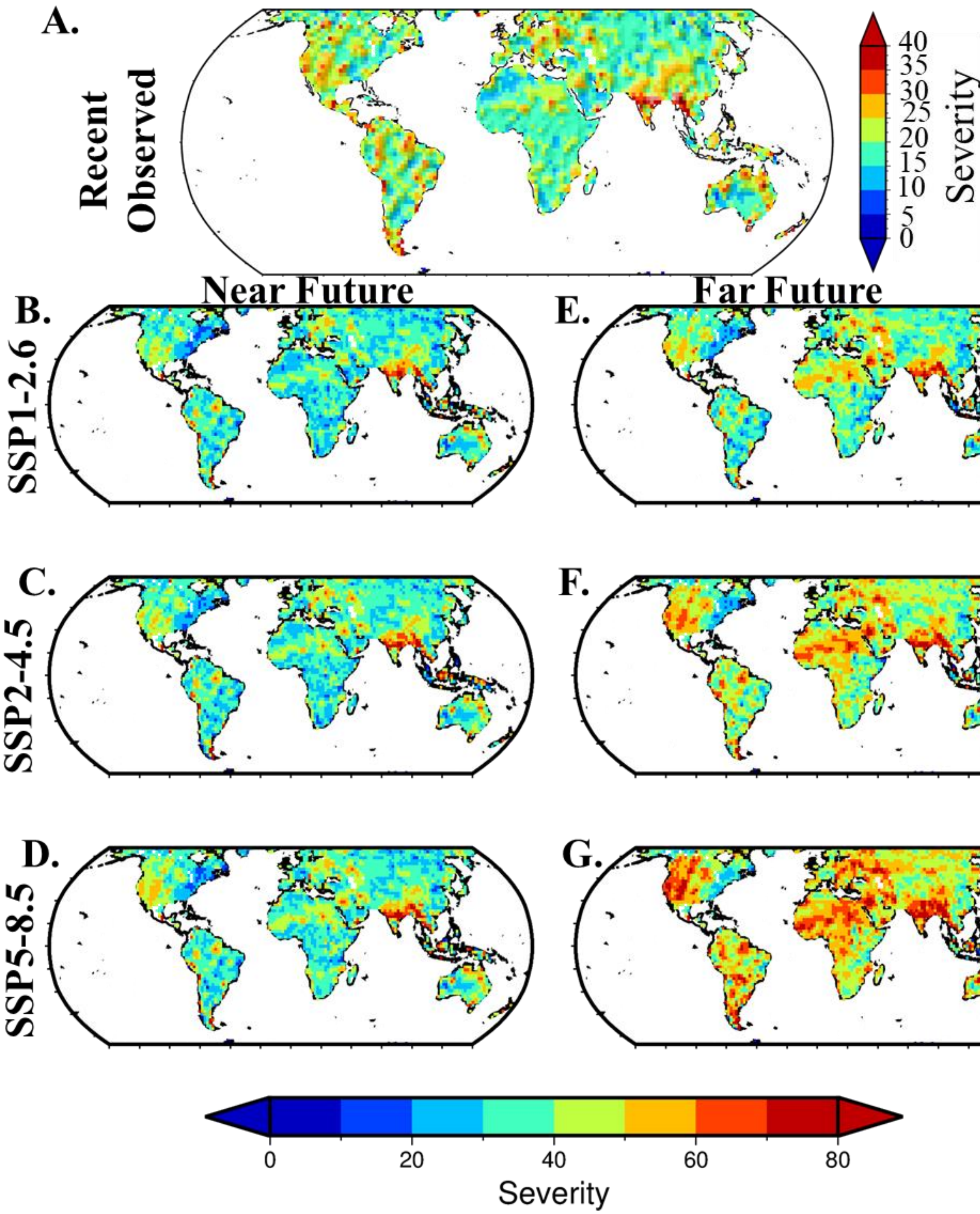


Figure S4 *Spatial map showing the CDHW severity for (a) recent observed period (1982-2020), (b) SSP1-2.6 near-future (2021-2057), (c) SSP2-4.5 near-future (2021-2057), (d) SSP5-8.5 near-future (2021-2057), (e) SSP1-2.6 far-future (2058-2095), (f) SSP2-4.5 far-future (2058-2095), (g) SSP5-8.5 far-future (2058-2095).*

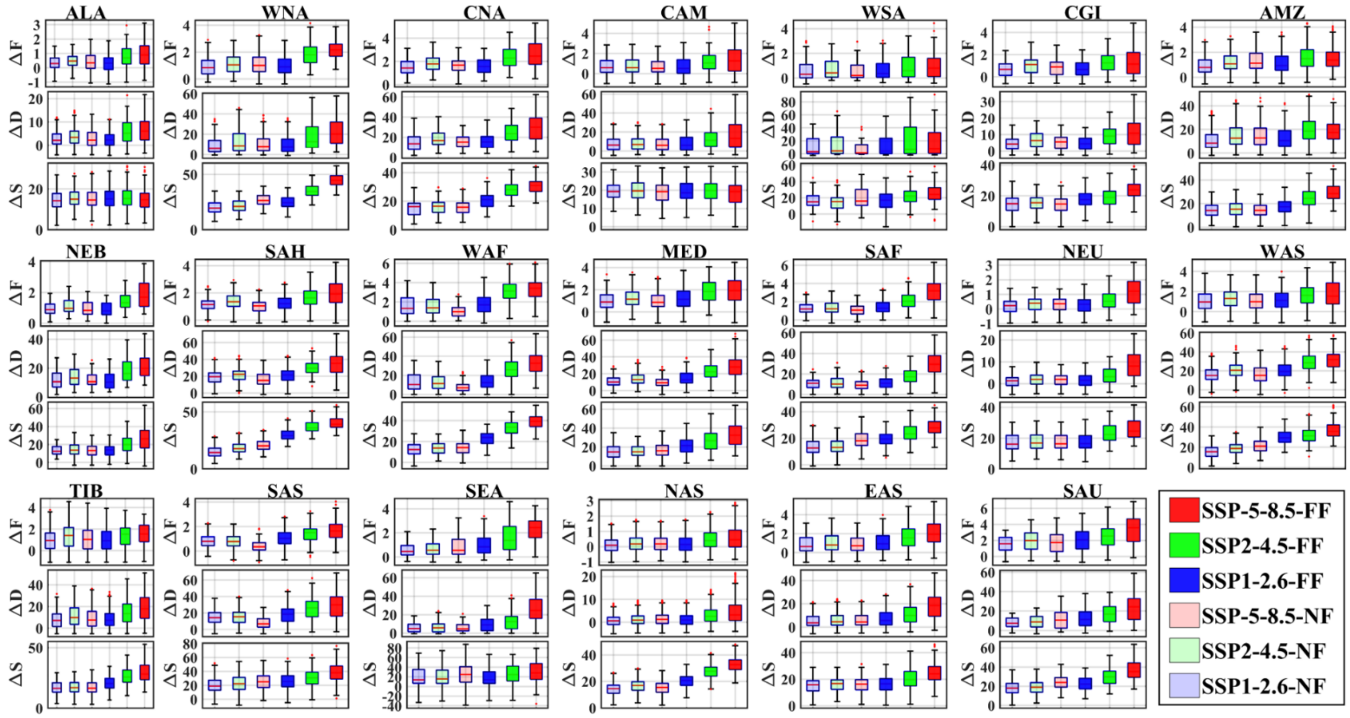


Figure S5*. Boxplots indicate the change in different CDHW metrics relative to recent observed period (top: frequency, middle: duration, and bottom: severity of each 3 by 1 subplot) are shown by SSP1-2.6-NF (light blue), SSP2-4.5-NF (light green), SSP5-8.5-NF (light red), SSP1-2.6-FF (dark blue), SSP2-4.5-FF (dark green), and SSP5-8.5-FF (dark red) for the twenty AR5 climate regions. Note that on the y-axis, the labels ΔF, ΔD, and ΔS represent the change in frequency, change in duration, and change in severity, respectively, relative to the recently observed period (1982 to 2019).*

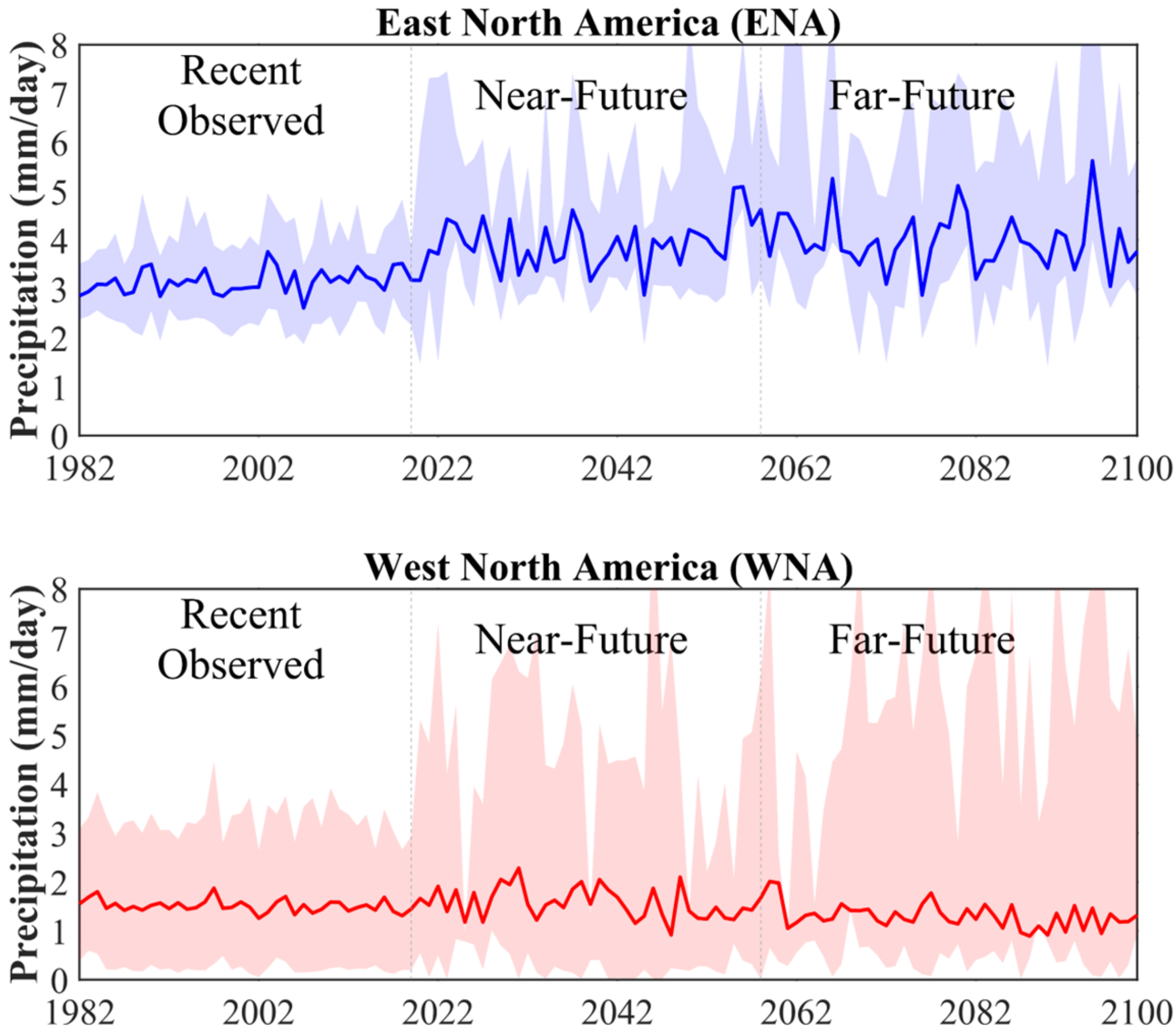


Figure S6: This figure presents the summer mean precipitation trends in East North America and West North America spanning from 1982 to 2100. Precipitation data for the period 1982-2019 is derived from the Global Precipitation Climatology Centre (GPCC) records, while data for the years beyond 2020 is based on the multimodel mean ensemble projections from the Shared Socioeconomic Pathways 5-8.5 (SSP-5-8.5) scenario of the CMIP6 GCMs.

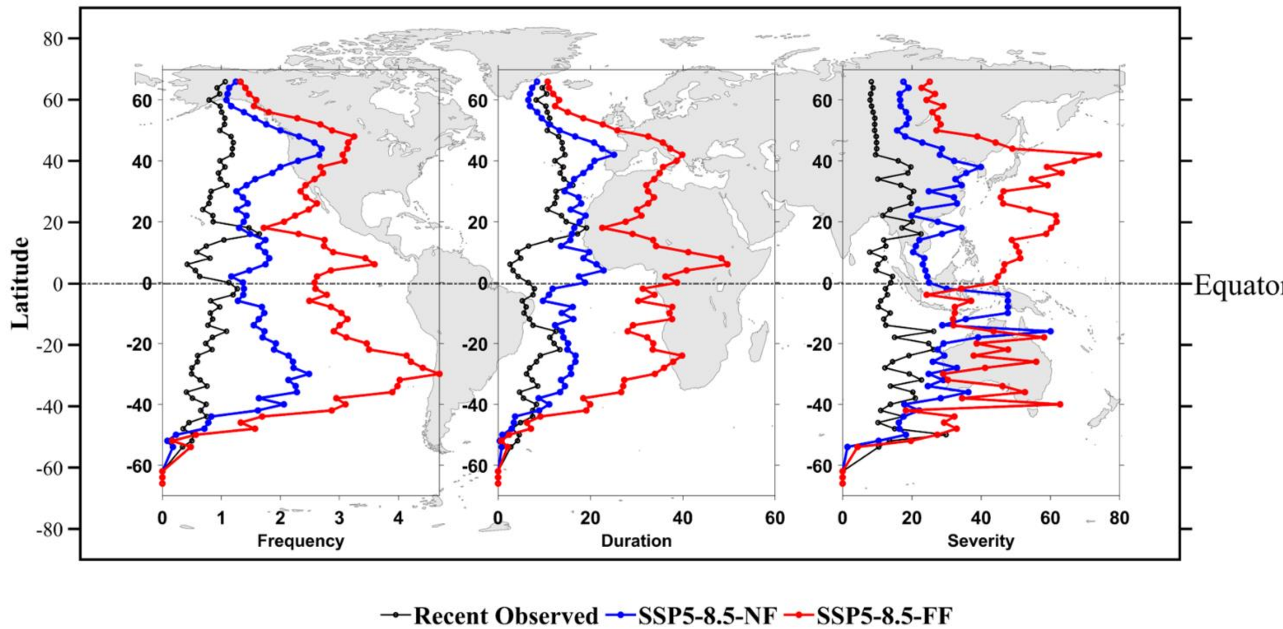


Figure S7. *Latitudinal variations of the CDHW characteristics — Frequency (left), Duration (middle), Severity (right) — are shown for recent observed period (black), SSP5-8.5 near-future (blue), and SSP5-8.5 far-future (red).*

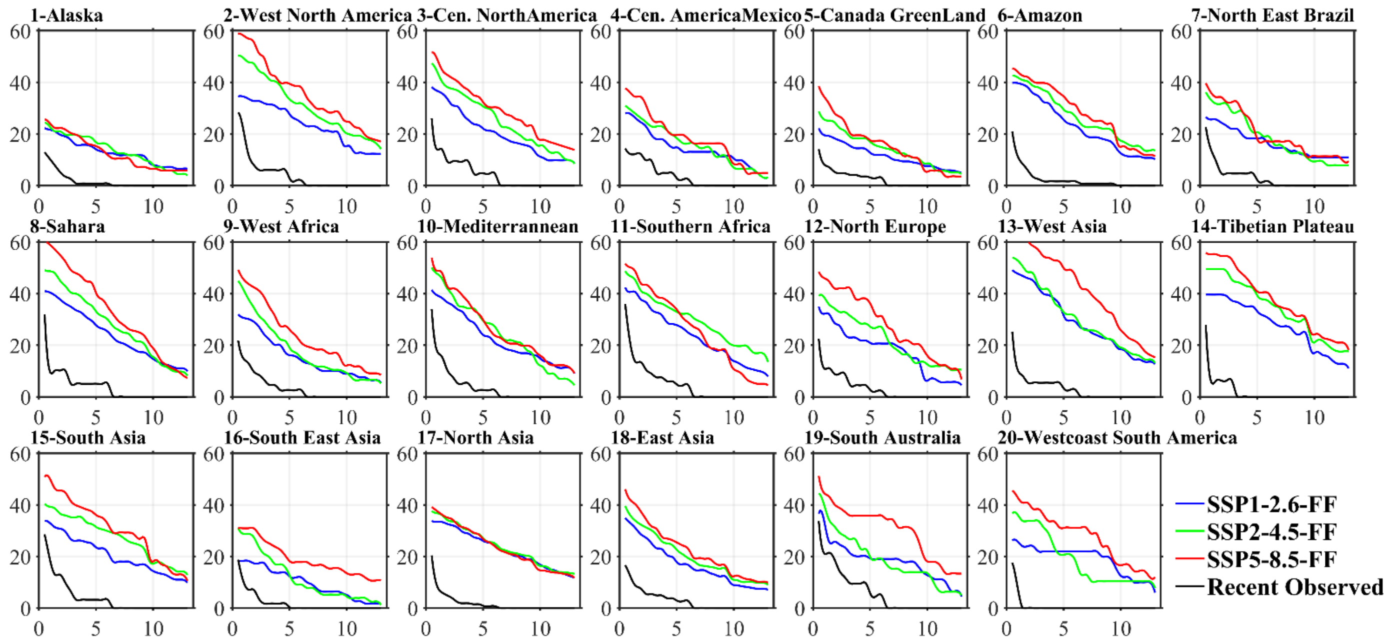


Figure S8. *Percentage of area (y-axis) versus return period (x-axis) curves for different climate regions are shown for the recent observed period (black dotted line), SSP1-2.6 (solid blue line), SSP2-4.5 (solid green line), SSP5-8.5 (solid red line). The curves are smoothened using a 7- year running mean.*

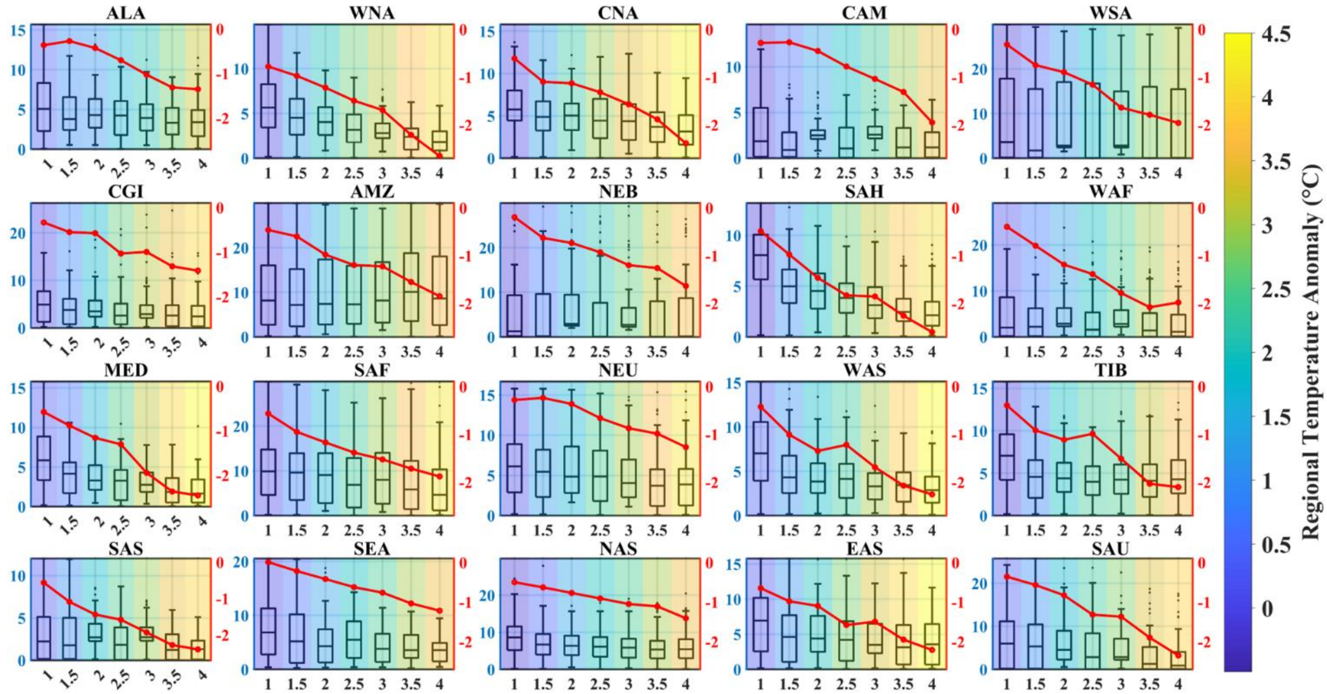


Figure S9. *Boxplots indicating the return periods (y-axis, left) of the CDHW events corresponding to different global mean temperature anomalies (x-axis) for 20 climate regions at different global warming levels. The color patch indicates the regional warming anomaly for a particular global warming anomaly. The red line specifies the sc-PDSI values (y-axis, right) during these periods.*

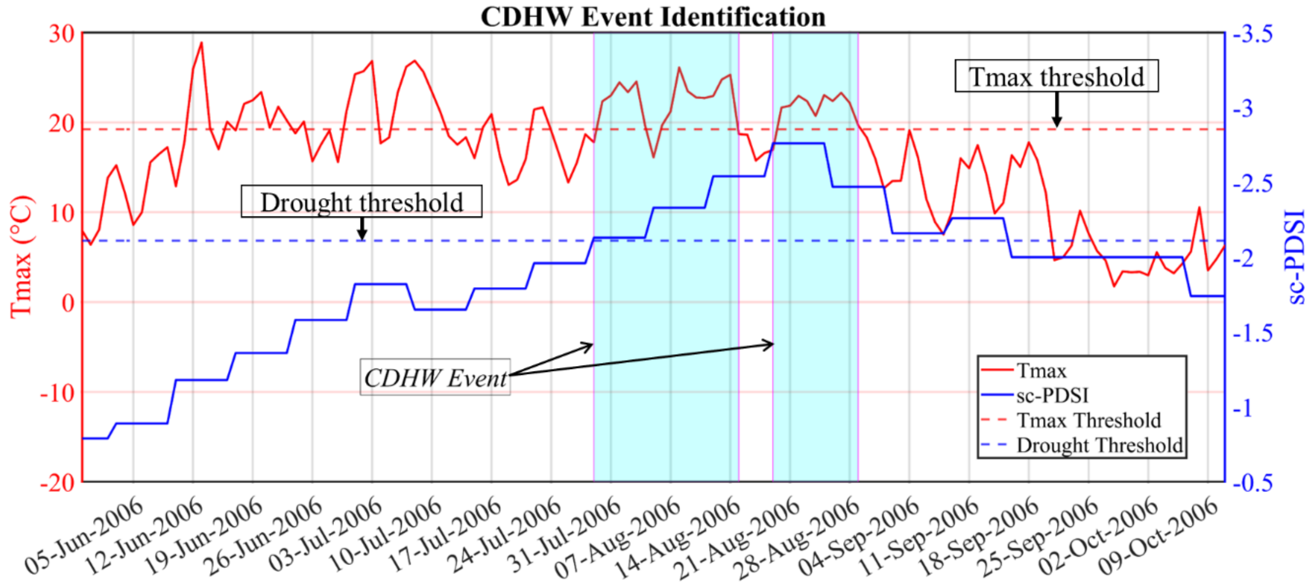


Figure S10. *This figure presents the daily maximum temperature (left y-axis) and weekly sc-PDSI time series (right y-axis) for a specific grid location (Lon=* $25°E$ *and* $Lat = 60°N$*) during the summer of 2006. The drought threshold is defined as the 10th percentile of the weekly sc-PDSI for the recent observed period (1982-2019), while the heatwave threshold corresponds to the 95th percentile of daily maximum temperature during the same period. In this example, two compound drought and heatwave events, illustrated by the shaded yellow areas, are identified when both drought and heatwave conditions simultaneously exceed their respective thresholds.*

Tables

Table S1. List of CMIP6 GCMs considered in the study

Sl No.	GCM Name	Realization	Original Resolution (Lat X Lon)
1	ACCESS-CM2	r1i1p1f1	1.25×1.875
2	ACCESS-ESM1-5		1.25×1.875
3	CanESM5		2.8×2.8
4	MIROC6		1.4×1.4
5	MPI-ESM1-2-HR		0.9375×0.9375
6	MPI-ESM1-2-LR		1.875×1.875
7	MRI-ESM2-0		1.125×1.125
8	NorESM2-MM		0.9×1.25

Table S2. Mann-Kendall test results with $\alpha = 0.01$ level of significance and Sen's slope values for different CDHW characteristics. H=1 (H=0) indicates there is (not) a statistically significant trends present in the series.

Scenarios	CDHW Events			CDHW Days			CDHW Severity		
	Slope	**P-value**	**H**	**Trend**	**P-value**	**H**	**Slope**	**P-value**	**H**
Observation	0.011	$<10^{-9}$	1	0.05	$<10^{-6}$	1	0.17	$<10^{-7}$	1
SSP126 NF	0.01	$<10^{-9}$	1	0.021	$<10^{-9}$	1	0.5	$<10^{-9}$	1
SSP245 NF	0.014	$<10^{-10}$	1	0.151	$<10^{-14}$	1	0.49	$<10^{-9}$	1
SSP585 NF	0.016	$<10^{-14}$	1	0.155	$<10^{-15}$	1	0.396	$<10^{-9}$	1
SSP126 FF	-0.0003	0.18	0	0.03	0.0032	1	0.06	0.16	0
SSP245 FF	0.013	$<10^{-9}$	1	0.12	$<10^{-15}$	1	0.593	$<10^{-10}$	1
SSP585 FF	0.028	$<10^{-14}$	1	0.45	$<10^{-19}$	1	2.53	$<10^{-14}$	1

References

1. Miao, C., Su, L., Sun, Q. & Duan, Q. A nonstationary bias-correction technique to remove bias in GCM simulations. *J. Geophys. Res. Atmospheres* **121**, 5718–5735 (2016).
2. Wang, L. & Chen, W. A CMIP5 multimodel projection of future temperature, precipitation, and climatological drought in China. *Int. J. Climatol.* **34**, 2059–2078 (2014).
3. Mukherjee, S. & Mishra, A. K. Increase in Compound Drought and Heatwaves in a Warming World. *Geophys. Res. Lett.* **48**, e2020GL090617 (2021).
4. Sen, P. K. Estimates of the Regression Coefficient Based on Kendall's Tau. *J. Am. Stat. Assoc.* **63**, 1379–1389 (1968).
5. Mann, H. B. Nonparametric Tests Against Trend. *Econometrica* **13**, 245–259 (1945).