nature geoscience

Supplementary information

https://doi.org/10.1038/s41561-025-01737-w

Accelerating increase in the duration of heatwaves under global warming

In the format provided by the authors and unedited

Supplementary Text 1: Theoretical prototypes for heat wave duration probability distributions

No-memory model

In this model, a temperature anomaly T_{his} under a historical configuration is simply given by

$$T_{his}(t) = \sigma_{his} r(t),$$

where σ_{his} is the temperature anomaly standard deviation under historical conditions, and r is a Gaussianly distributed random number with zero mean and unit variance. Under only a shift in the mean ΔT, a future temperature anomaly (with respect to a historical climatology) will be given by

$$T_{fut}(t) = \Delta T + \sigma_{his} r(t).$$

Both T_{his} and T_{fut} are normally distributed with standard deviation σ_{his} and means 0 and ΔT respectively, and are day-to-day uncorrelated.

Under this configuration, the probability of having n consecutive T values above a critical threshold T_c used to define a heat wave (i.e., a heat wave duration of n days) is given by a series of Bernoulli trials with a probability p of success. Here, p is established from the threshold used to define a heat wave; in our study a heat wave is defined for values above the 90th percentile of the historical temperature anomaly distribution, so $p_{his} = 0.1$. The resulting heat wave duration probability distribution $PDF(t)$ is exponentially distributed (see Figure 1 in the main text and Supplementary Figure 8) and is given by

$$PDF(t) = \frac{1}{t_L} exp(\frac{-t}{t_L}), \qquad \text{(S1)}$$

where t is the heat wave duration and $t_L = -\frac{1}{log(p)}$ is the duration cutoff-scale in this case. Note the normalization constant uses the continuum approximation, but the exponential is exact. The probability of exceedance (or "success" in the Bernoulli trials context) in the historical (p_{his}) and global warming (p_{fut}) cases are given by

$$p_{his} = 1 - \frac{1}{2}\left(1 + erf\left(\frac{T_c}{\sqrt{2}\sigma_{his}}\right)\right) = 0.1,$$
$$p_{fut} = 1 - \frac{1}{2}\left(1 + erf\left(\frac{T_c - \Delta T}{\sqrt{2}\sigma_{his}}\right)\right), \qquad \text{(S2)}$$

where erf is the error function. The dependence of p_{fut} on $\frac{\Delta T}{\sigma_{his}}$ is illustrated in Supplementary Figure 9.

The duration t_q corresponding to the $100q$ percentile is given by

$$t_q = -t_L log(1-q) = \frac{log(1-q)}{log(p)}.$$

This implies that the ratio between future and historical heat wave duration percentiles is given by

$$\frac{t_q^{fut}}{t_q^{his}} = \frac{t_L^{fut}}{t_L^{his}} = \frac{log(p_{his})}{log(p_{fut})}, \qquad \text{(S3)}$$

and is independent of percentile.

As p_{fut} is a function of $\frac{\Delta T}{\sigma_{his}}$, it follows that $\frac{t_q^{fut}}{t_q^{his}}$ will also be a function of $\frac{\Delta T}{\sigma_{his}}$. Note that the analytical solution implies that the t_L ratio (or t_{99} ratio) continue increasing (Supplementary Figure 10) even at the regime where the future exceedance p_{fut} saturates (Supplementary Figure 9). This behavior is also found for the autocorrelated model (see below; Supplementary Figure 11), suggesting that it may also hold for future projections. Empirically we find that the $\frac{t_q^{fut}}{t_q^{his}}$ dependence on $\frac{\Delta T}{\sigma_{his}}$ can be reasonably approximated by a quadratic expansion over the range of interest (Supplementary Fig. 10).

Autocorrelated model

This model adds a layer of complexity by considering a decorrelation timescale t_D for the temperature anomaly generated. In the historical case, the evolution of a temperature anomaly $T = T_{his}$ is governed by

$$\frac{dT}{dt} = \frac{-T}{t_D} + \sqrt{\frac{2}{t_D}}\sigma_{his}\xi,$$

where ξ is a Gaussian white noise process. This type of model is generically referred to as an Ornstein-Uhlenbeck (OU) process. Similarly as in the no-memory model, T is distributed following a Gaussian distribution with mean 0 and standard deviation σ_{his}, but now the time series are day-to-day autocorrelated with a timescale t_D (t_D is the timescale where the autocorrelation coefficient first drops below e^{-1}). Under a shift in mean temperature ΔT, the future anomalies $T = T_{fut}$ (with respect to the historical baseline) follows a similar equation

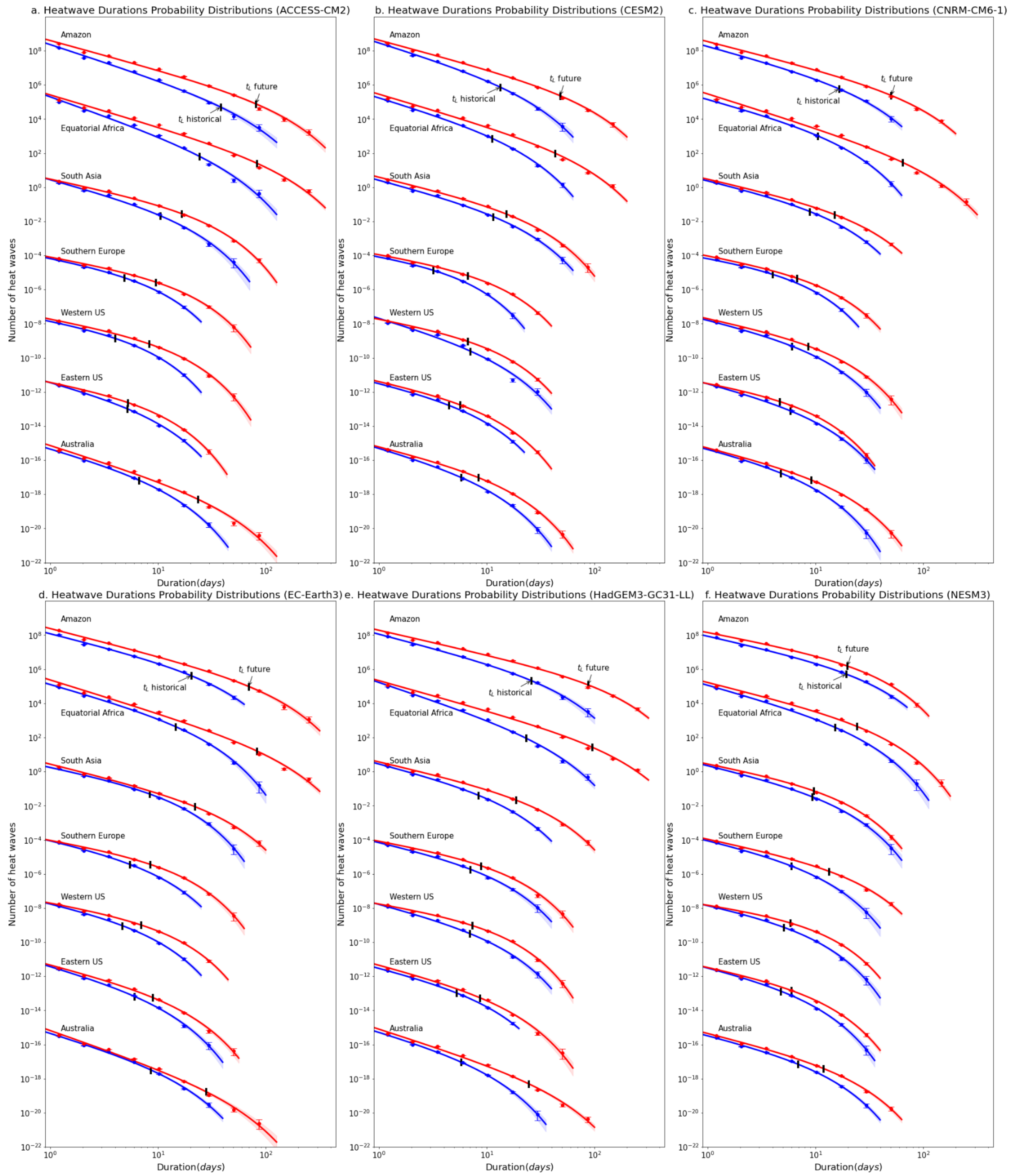
$$\frac{dT}{dt} = \frac{-(T-\Delta T)}{t_D} + \sqrt{\frac{2}{t_D}}\sigma_{fut}\xi.$$

As a leading case, consider $\sigma_{fut} = \sigma_{his}$. Under this configuration, empirically we find that the shape of the duration probability distributions is modified from the no-memory case by the addition of a power law range controlling the probability of low and moderate durations, similar to distributions seen in observations and GCM output (Fig. 1 main text), maintaining an approximately exponential probability tail for long durations (Supplementary Figure 12).

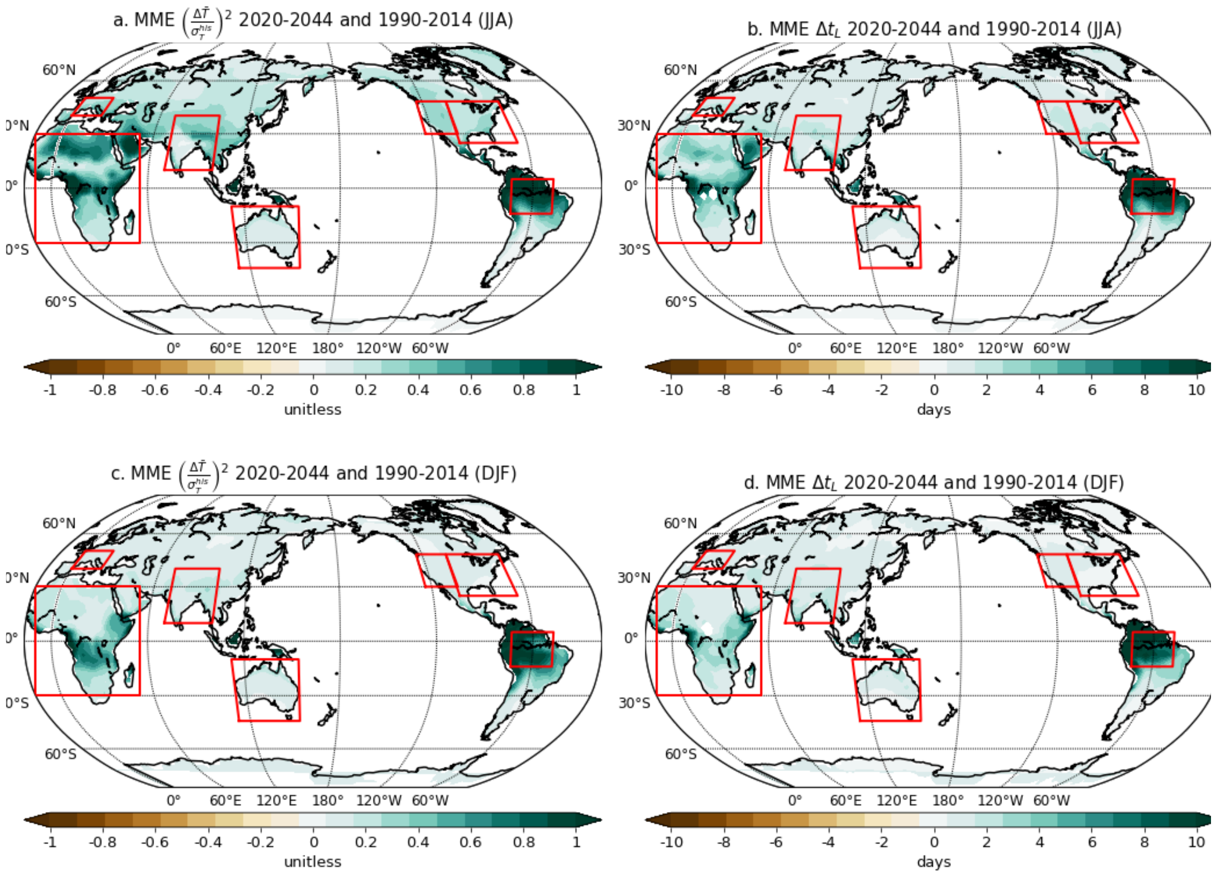
Similarly to the "no-memory" model, we find empirically that $\frac{t_q^{fut}}{t_q^{his}}$ depends on $\frac{\Delta T}{\sigma_{his}}$ also approximately quadratically (Fig. 3 in the main text; see also Supplementary Figures 10 and 11), although with a smaller rate of increase.

Impacts of changes in standard deviation and autocorrelation

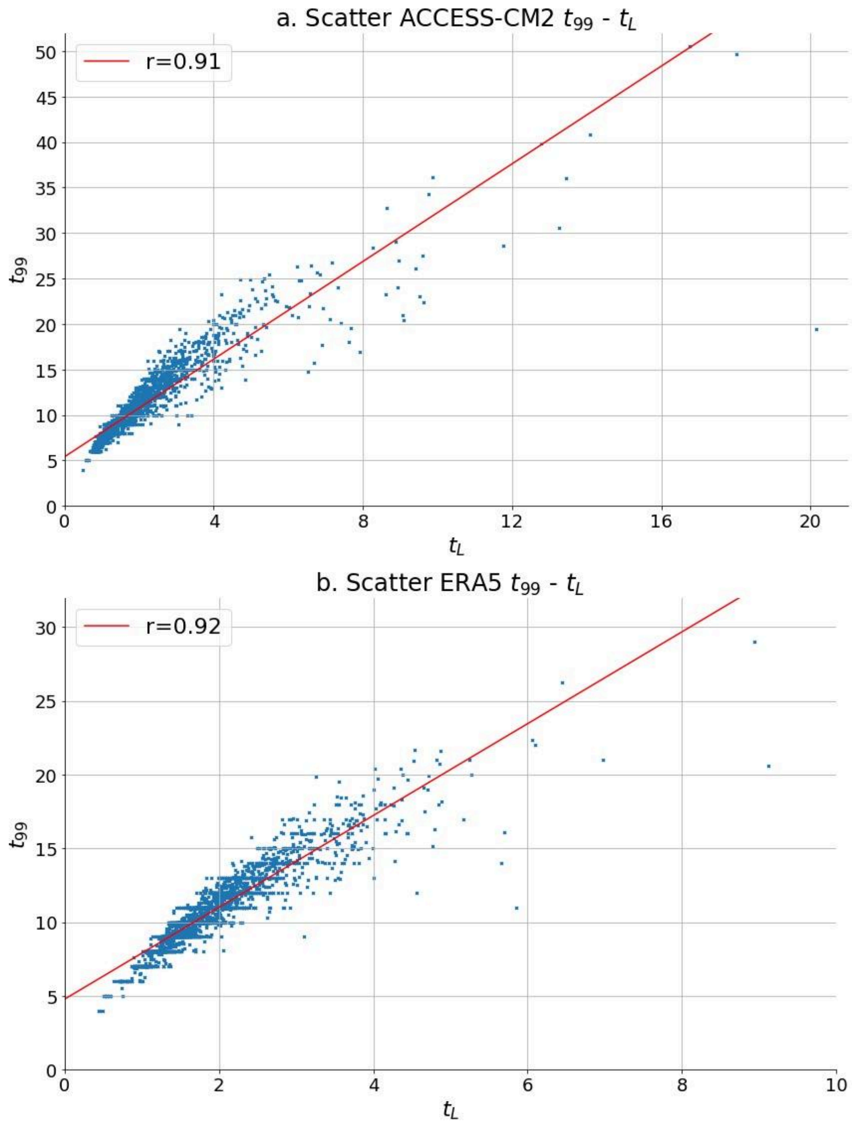
Changes in standard deviation σ and in autocorrelation timescale t_D can potentially affect duration distributions with warming. Supplementary Figure 4 shows example temperature distributions from CMIP six models normalized by standard deviation $\frac{T}{\sigma}$, where T is shown relative to the historical mean and the normalization for the future case is done alternately by σ_{his} and σ_{fut}. These illustrate the extent to which changes in σ constitute a second-order effect compared to changes in the mean when normalizing by σ_{his} instead of σ_{fut}. The region in Supplementary Figure 4 is chosen based on its relatively large percent change in temperature anomaly standard deviation, as seen in Supplementary Figure 5. Elsewhere, effects are even more modest. Supplementary Figure 6a shows changes in multi-model mean autocorrelation timescale t_D between future and historical, which are very small compared to the percent changes in duration, which explains why an approximation that neglects these changes with climate can still work well. We underline that the changes *across regions* in a given climate are not small. Supplementary figure 10 addresses this within the autocorrelated prototype, showing the relative insensitivity of duration curves as a function of $\frac{\Delta T}{\sigma}$ to different autocorrelation timescales, as would occur from region to region. This explains why we can successfully collapse curves across a broad range of regions (Figure 3 of the main text).



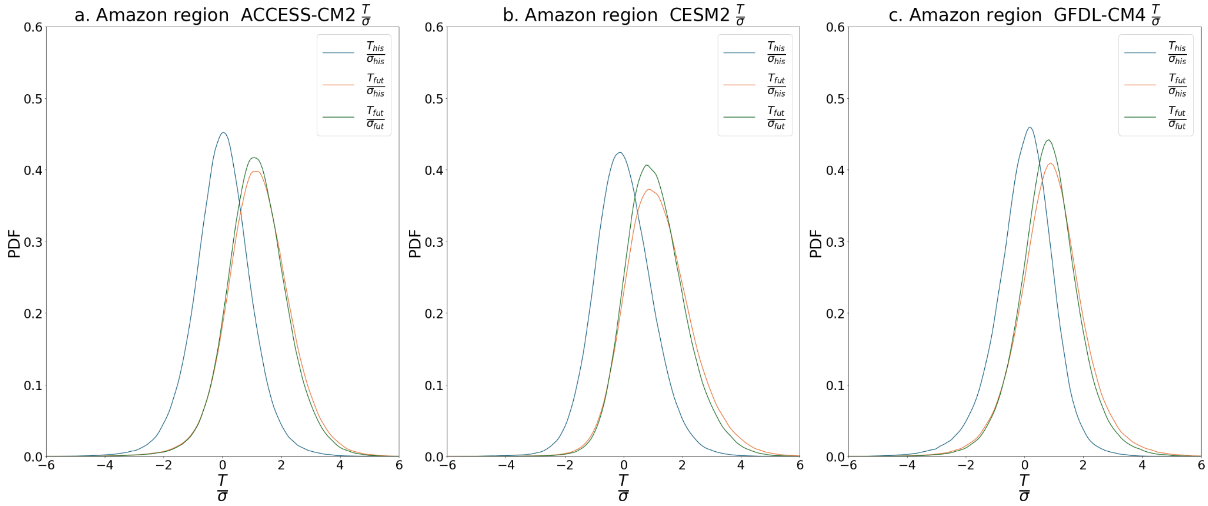
Supplementary Figure 1. Similar to Fig. 1a in the main text but for six other models. PDFs in each panel are computed from daily temperature data over 25 years for historical and future periods (n=9,125 days per period), aggregated across land grid points within each region. Error bars reflect the 5th-95th percentile range from 100 bootstrap realizations, sampled with replacement.



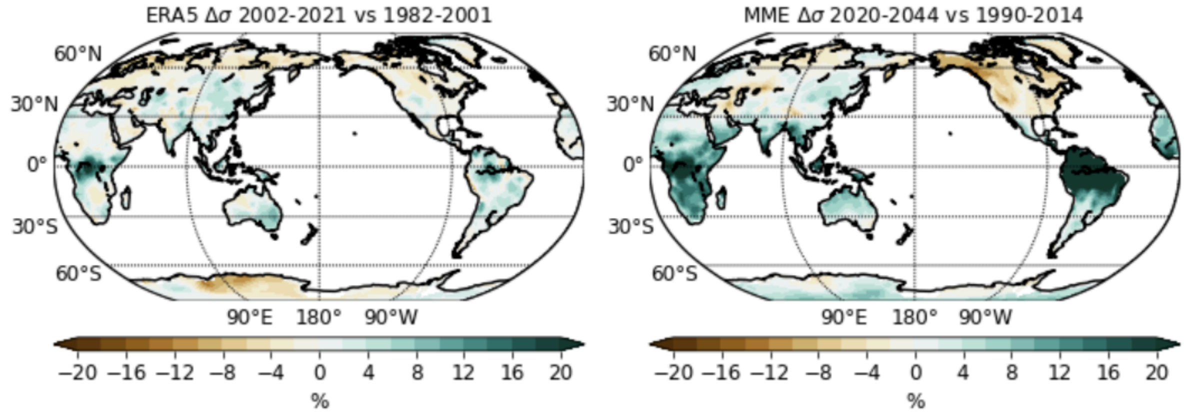
Supplementary Figure 2. Same as Figure 4 in the main text but focused on boreal summer (Jun-Jul-Aug; top) and austral summer (Dec-Jan-Feb; bottom). We classify the seasonality of the heatwaves by the month the heat wave started.



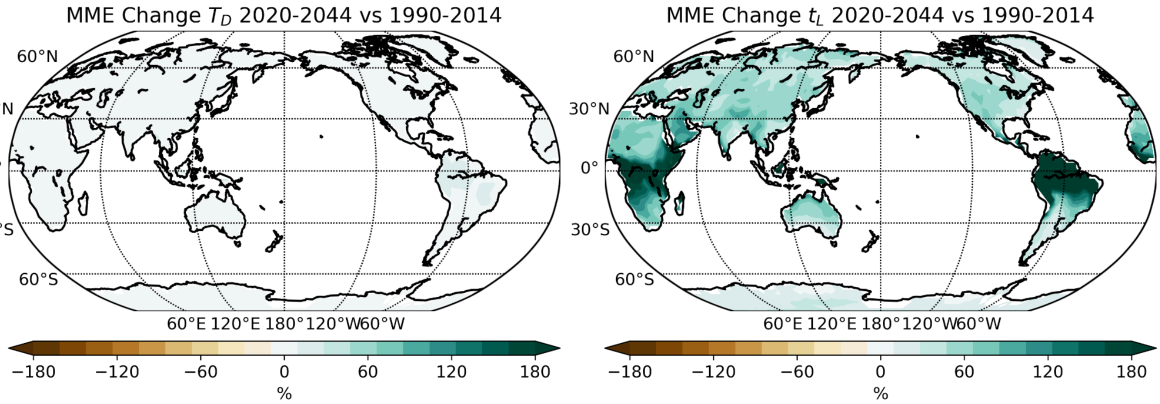
Supplementary Figure 3. Scatter plot between the 99th percentile of heat wave durations and duration scale t_L in one climate model (ACCESS-CM2; top) and ERA5 (bottom).



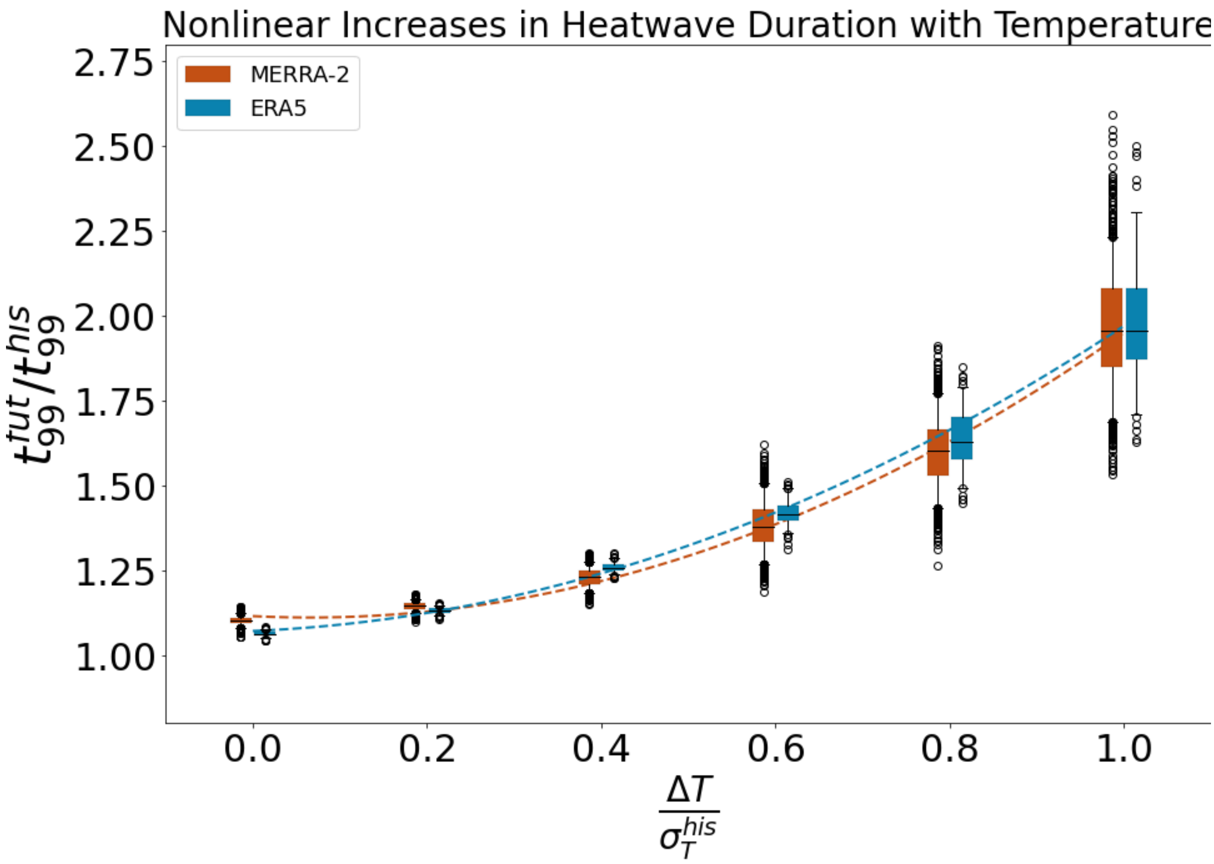
Supplementary Figure 4. Examples of changes in temperature anomaly distributions (2020-2044 vs 1990-2014) in 3 models (ACCESS-CM2, CESM2, GFDL-CM4) in the Amazon region. The historical distributions are shown standardized by the historical standard deviation (i.e., the variable is $\frac{T_{his}}{\sigma_{his}}$). To illustrate first-order and second-order effects, the future temperature anomaly distributions are shown standardized by the historical standard deviation ($\frac{T_{fut}}{\sigma_{his}}$) and future standard deviation ($\frac{T_{fut}}{\sigma_{fut}}$). The similarity between $\frac{T_{fut}}{\sigma_{his}}$ and $\frac{T_{fut}}{\sigma_{fut}}$ distributions illustrate that the leading effect is the change in the mean. A second-order effect (much smaller than the leading effect) can be seen in the different widths of $\frac{T_{fut}}{\sigma_{his}}$ and $\frac{T_{fut}}{\sigma_{fut}}$ distributions, with the $\frac{T_{fut}}{\sigma_{fut}}$ distribution being narrower, consistent with a larger temperature anomaly standard deviation in the future in the region (see Supplementary Figure 5).



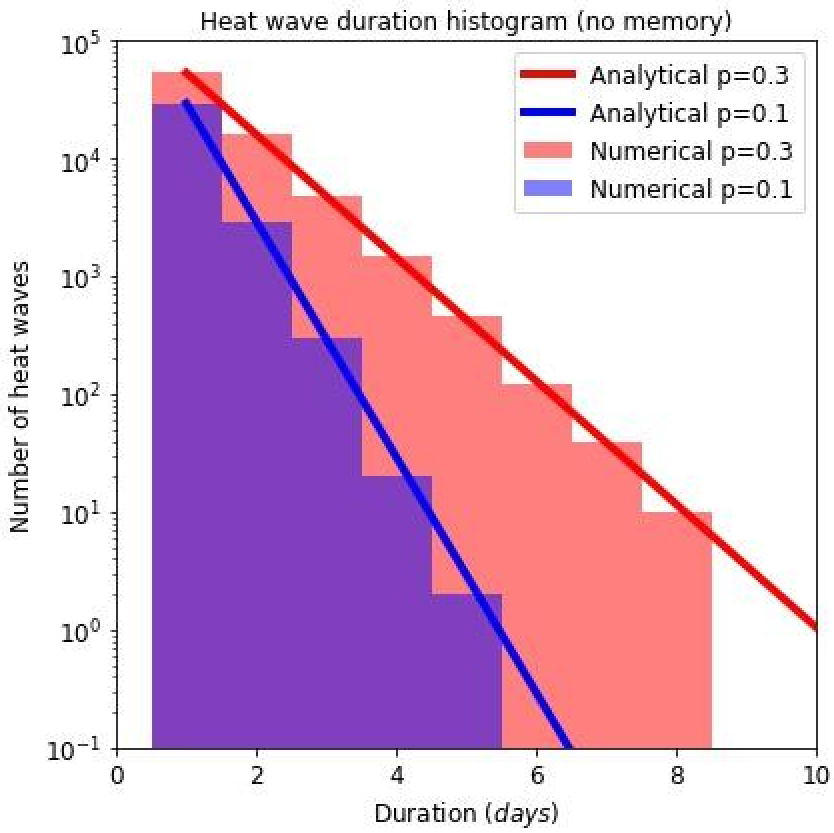
Supplementary Figure 5. Percent changes in temperature anomaly standard deviation in **a.** ERA5 (2002-2021 vs 1982-2001) and **b.** CMIP6 (multi-model mean; 2020-2044 vs 1990-2014).



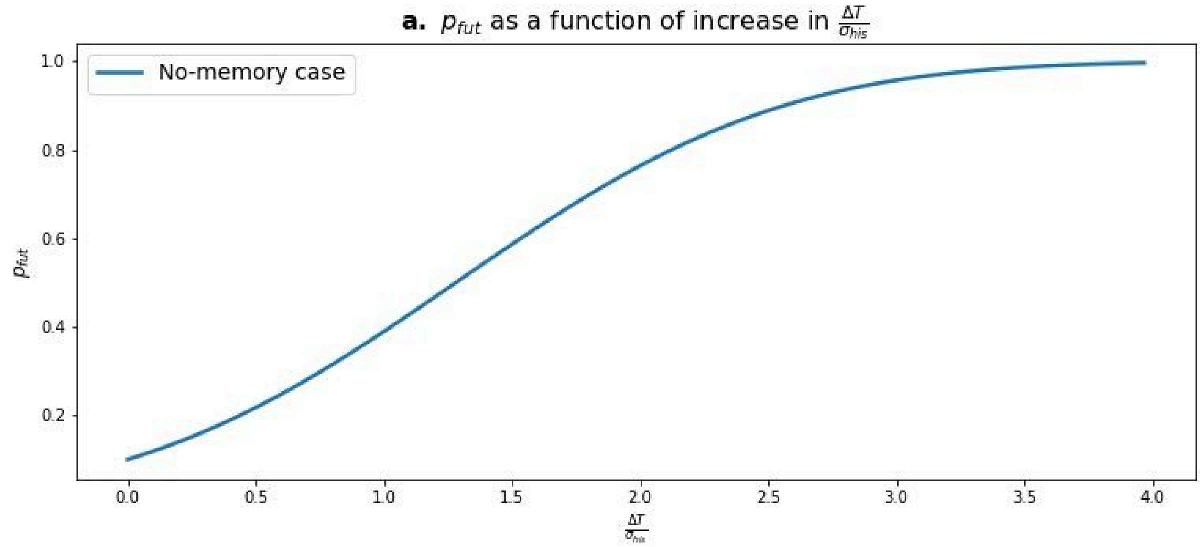
Supplementary Figure 6. Multi-model mean percent change in **a** autocorrelation timescale t_D and **b** extreme heatwave duration timescale t_L between 2020-2044 compared with 1990-2014. On the same scale, increases in t_L are much larger than changes in t_D, and are mainly controlled nonlinearly by normalized increases in temperature (Fig. 4 in the main text).



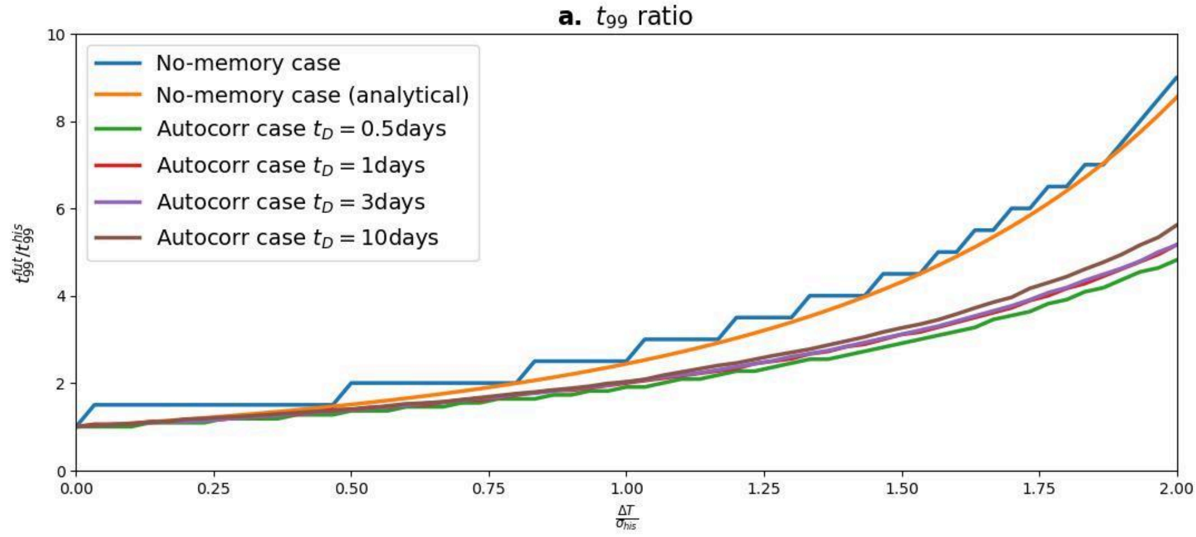
Supplementary Figure 7. Similar to Figure 3 in the main text but using ERA5 and MERRA-2 for periods 1982-2001 and 2002-2021. Statistics are derived from daily data over 20 years for each period (n=7,300 days), at each land grid point. These daily values are used to compute heat wave durations and their extreme percentiles. Box and whisker definitions follow those in Fig. 3 of the main text: the box spans the 25th (Q1) to 75th (Q3) percentiles, and whiskers extend to $Q_1 - 1.5IQR$ and $Q_3 + 1.5IQR$, where IQR is the interquartile range (Q3 – Q1).



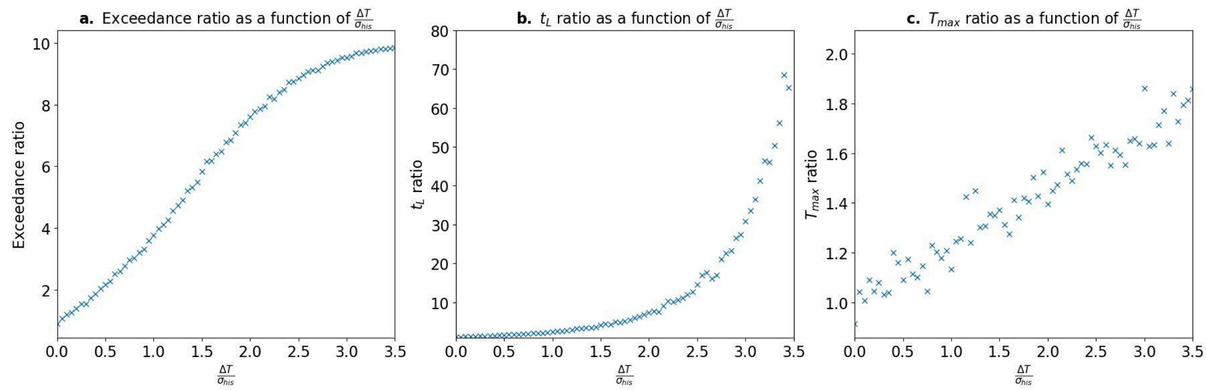
Supplementary Figure 8. Analytical (solid lines; eq. S1) and numerical (bars) probability distributions of heat wave durations generated by the "no-memory" model. Blue (red) colors represent the duration probability distribution calculated with a probability of exceedance $p = p_{his} = 0.1$ ($p = p_{fut} = 0.3$) representative of historical (global warming) conditions. The probability distributions are normalized to integrate to the number of heat waves in each case.



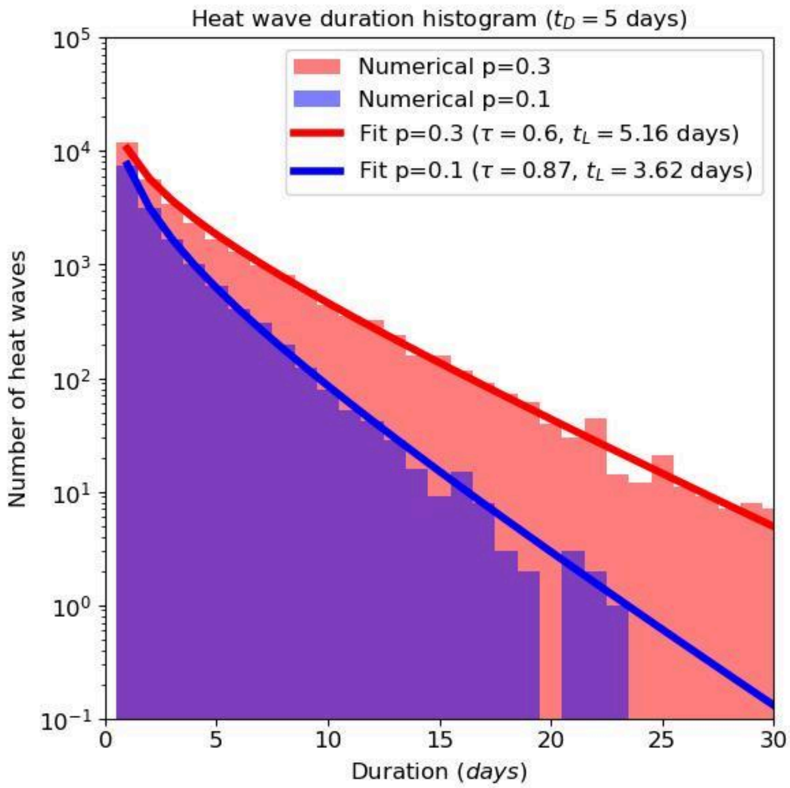
Supplementary Figure 9. Probability of exceedance (p_{fut}) as a function of $\frac{\Delta T}{\sigma_{his}}$, for the no-memory case as derived in equation S2.



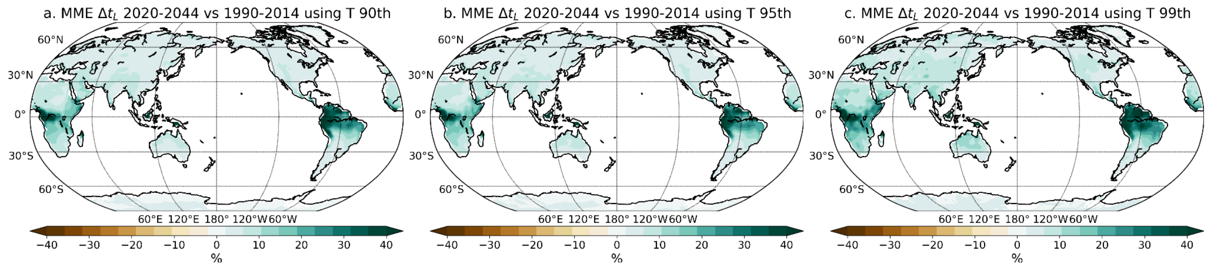
Supplementary Figure 10. Ratio of extreme heat wave durations $\frac{t_{99}^{fut}}{t_{99}^{his}}$ as a function of $\frac{\Delta T}{\sigma_{his}}$ generated by the "no-memory" model (numerical and analytical; eq. S3) and the autocorrelated model for different t_D timescales.



Supplementary Figure 11. Ratio between historical and global warming **a.** exceedances (above the 90th percentile of historical temperature anomaly), **b.** duration scale t_L, and **c.** maximum temperature as a function of the mean increase in temperature divided by the historical standard deviation ($\frac{\Delta T}{\sigma_{his}}$) in the autocorrelated stochastic prototype. The “no-memory” case showcases similar results. Each point is calculated from a 200yr run with a decorrelation timescale of 10 days. Results are not sensitive to the decorrelation timescale chosen.



Supplementary Figure 12. Similar to Supplementary Figure 3 but generated from the autocorrelated model. We fitted a histogram of the form $At^{-\tau}exp\left(\frac{-t}{t_L}\right)$, with coefficients A, τ, and t_L calculated using multivariate linear regression as in Martinez-Villalobos and Neelin 2019. Only bins with 10 or more counts are taken into account in the fit. These synthetically generated PDFs have similar shapes as the ones in observations and models (Fig. 1).



Supplementary Figure 13. Multi-model mean percent changes in t_L for three different heat wave definition thresholds: 90th, 95th and 99th percentiles of local daily temperature anomalies. The spatial patterns are highly correlated with correlation coefficients of 0.99 between panels **a** and **b**, 0.95 between panels **a** and **c**, and 0.96 between panels **b** and **c**.

Model Name	Modeling Group
ACCESS-CM2	CSIRO-ARCCSS
ACCESS-ESM1-5	CSIRO
AWI-CM-1-1-MR	AWI
BCC-CSM2-MR	BCC
CAMS-CSM1-0*	CAMS
CanESM5	CCCma
CESM2	NCAR
CESM2-WACCM	NCAR
CMCC-CM2-SR5	CMCC
CMCC-ESM2	CMCC
CNRM-CM6-1	CNRM-CERFACS
CNRM-CM6-1-HR	CNRM-CERFACS
CNRM-ESM2-1	CNRM-CERFACS
EC-Earth3	EC-Earth-Consortium
EC-Earth3-CC	EC-Earth-Consortium
EC-Earth3-Veg	EC-Earth-Consortium
EC-Earth3-Veg-LR	EC-Earth-Consortium
FGOALS-g3	CAS
GFDL-CM4	NOAA-GFDL
GFDL-ESM4*	NOAA-GFDL
HadGEM3-GC31-LL	MOHC NERC
IITM-ESM*	CCCR-IITM
INM-CM4-8	INM
INM-CM5-0	INM
IPSL-CM6A-LR	IPSL

MIROC6*	MIROC
MIROC-ES2L	MIROC
MPI-ESM1-2-HR	DKRZ DWD MPI-M
MPI-ESM1-2-LR	DKRZ DWD MPI-M
NESM3	NUIST
NorESM2-LM	NCC
NorESM2-MM	NCC
TaiESM1	AS-RCEC

Supplementary Table 1. CMIP6 models used. Asterisks denote models not used in Figure 3 of the main text.

Region	$\frac{t_{99}^{fut}}{t_{99}^{his}}$
Amazon	2.80
Equatorial Africa	4.01
South Asia	1.55
Southern Europe	1.66
Western US	1.36
Eastern US	1.34
Australia	1.96

Supplementary Table 2. Multi model ensemble mean fractional increase in duration scale t_L comparing 2020-2044 to 1990-2014.