Supplementary Information (SI) for

Accelerating increases in heat waves duration under global warming

Cristian Martinez-Villalobos[1,2], Danning Fu[3], Paul Loikith[4], and J. David Neelin[5]

[1] Faculty of Engineering and Science, Universidad Adolfo Ibañez, Santiago, Chile
[2] Data Observatory Foundation, ANID Technology Center, No. DO210001, Santiago, Chile
[3] Department of Biomedical Data Science, Stanford University, Stanford, California, USA
[4] Department of Geography, Portland State University, Portland, Oregon, USA
[5] Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, California, USA

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

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

Importantly, 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 Supplementary Figure 3) 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. 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

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

where erf is the error function. The dependence of p_{fut} on $\frac{\Delta T}{\sigma_{his}}$ is illustrated in Supplementary Figure 4.

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 $\frac{\Delta T}{\sigma_{his}}$. 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. 5).

Autocorrelated model

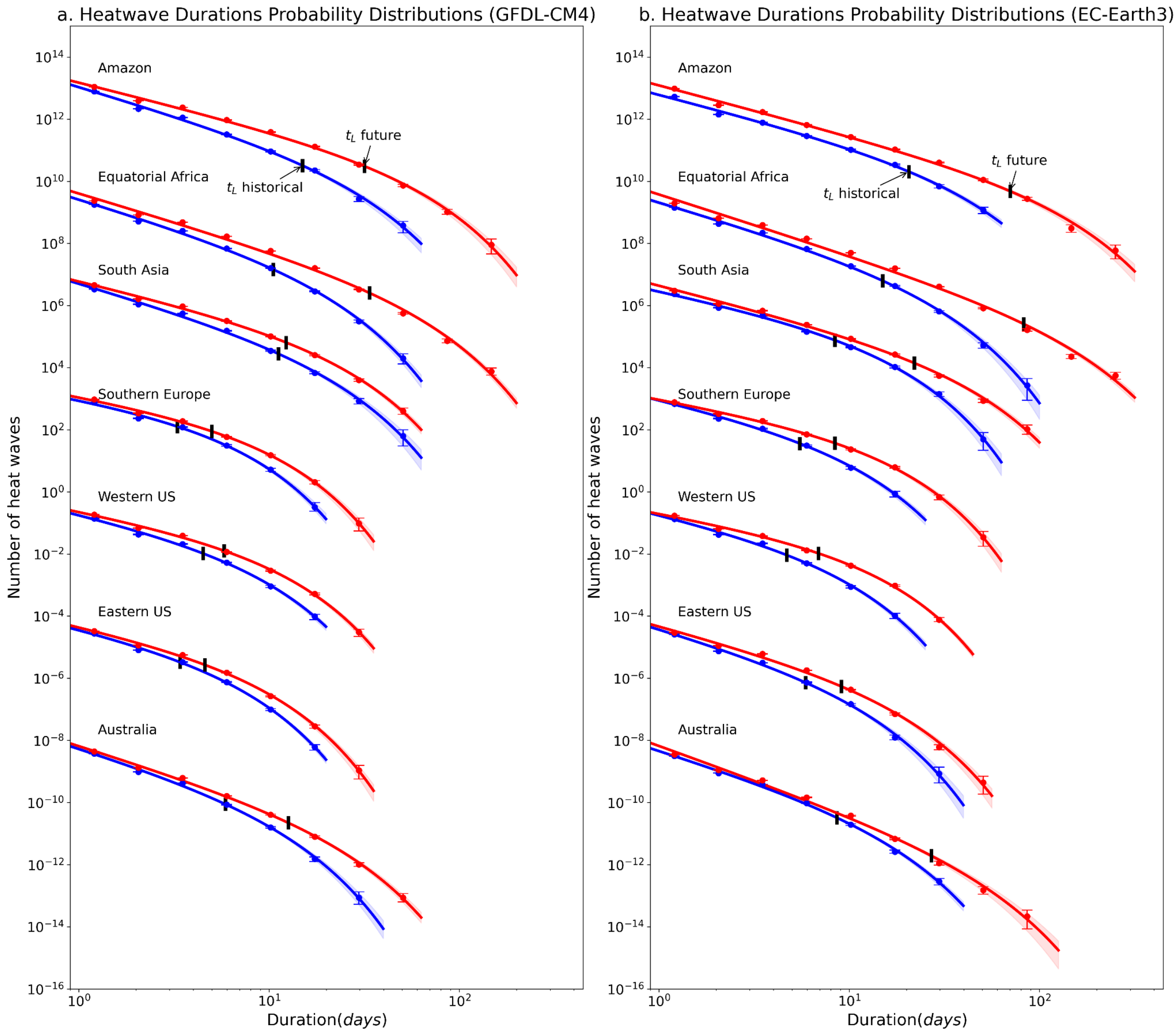
This model adds a layer of complexity by considering a decorrelation timescale τ_c for the temperature anomaly generated. In the historical case, the evolution of a temperature anomaly $T = T_{his}$ is governed by

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

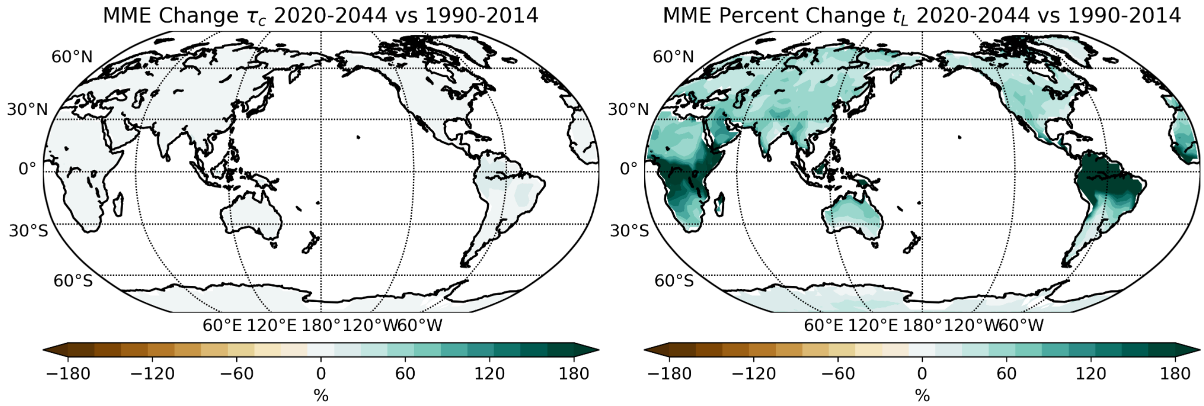
where ξ is a Gaussian white noise process. This type of model is generically referred to as an Ornstein-Uhlenbeck (OU) process. Similarly as in the no-memory model, T is distributed following a Gaussian distribution with mean 0 and standard deviation σ_{his}, but now the time series are day-to-day autocorrelated with a timescale τ_c (τ_c 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

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

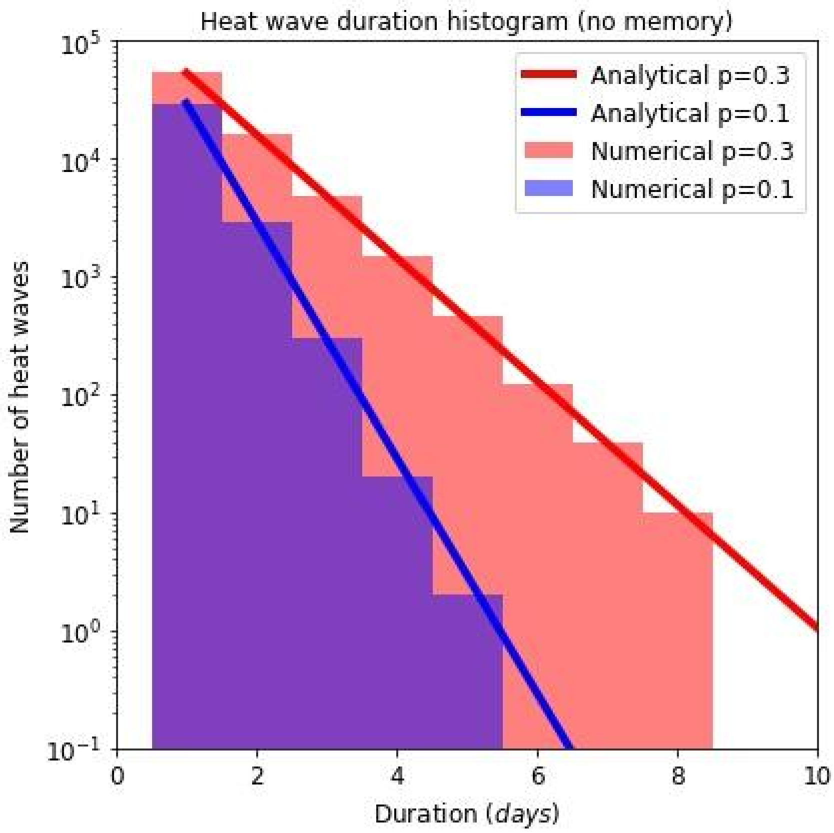
Under this configuration, empirically we find that the shape of the duration probability distributions is modified from the no-memory case by the addition of a power law range controlling the probability of low and moderate durations, similar to distributions seen in observations and GCM output (Fig. 1 main text), maintaining an approximately exponential probability tail for long durations (Supplementary Fig. 6). 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 Fig. 5), although with a smaller rate of increase.



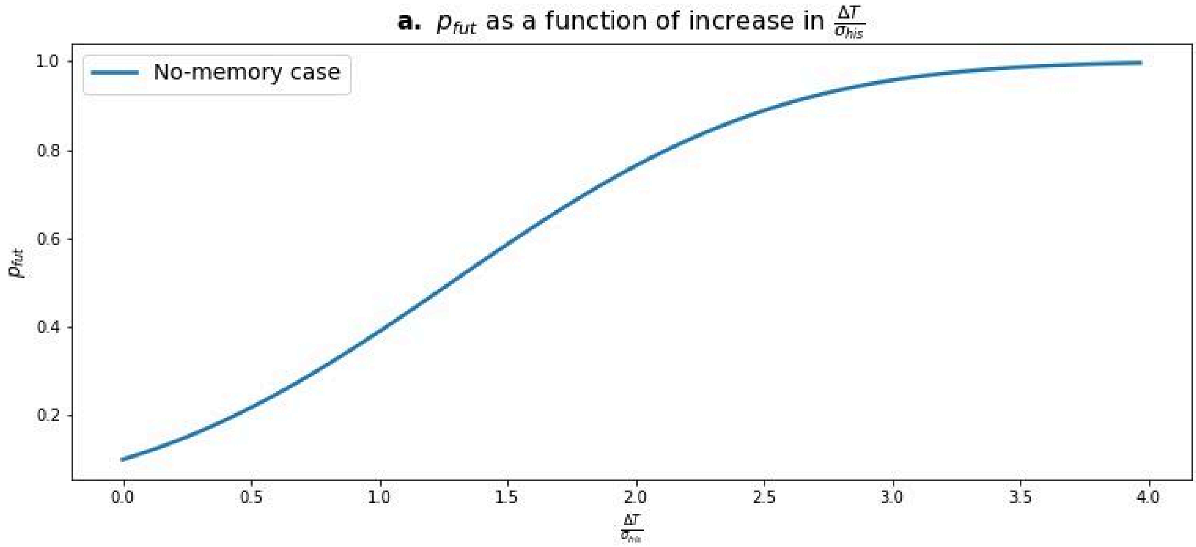
Supplementary Figure 1. Similar to Fig. 1a in the main text but for GFDL-CM4 and EC-Earth3 models.



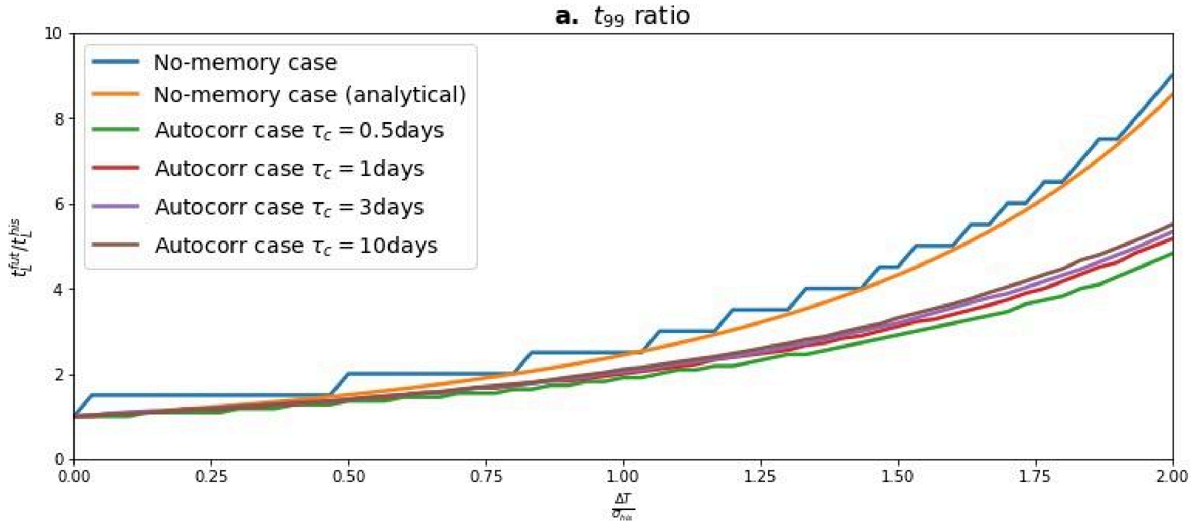
Supplementary Figure 2. Multi-model mean percent change in **a** autocorrelation timescale τ_c and **b** extreme heatwave duration timescale t_L between 2020-2044 compared with 1990-2014. On the same scale, increases in t_L are much larger than changes in τ_c, and are mainly controlled nonlinearly by normalized increases in temperature (Fig. 4 in the main text).



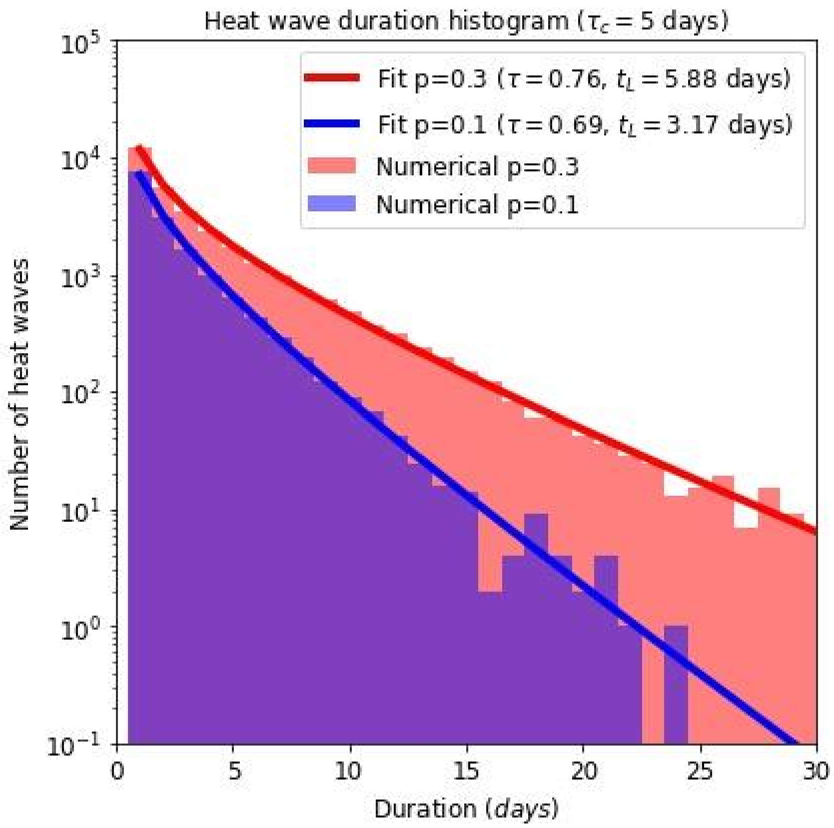
Supplementary Figure 3. Analytical (solid lines; eq. S1) and numerical (bars) probability distributions of heat wave durations generated by the "no-memory" model. Blue (red) colors represent the duration probability distribution calculated with a probability of exceedance $p = p_{his} = 0.1$ ($p = p_{fut} = 0.3$) representative of historical (global warming) conditions. The probability distributions are normalized to integrate to the number of heat waves in each case.



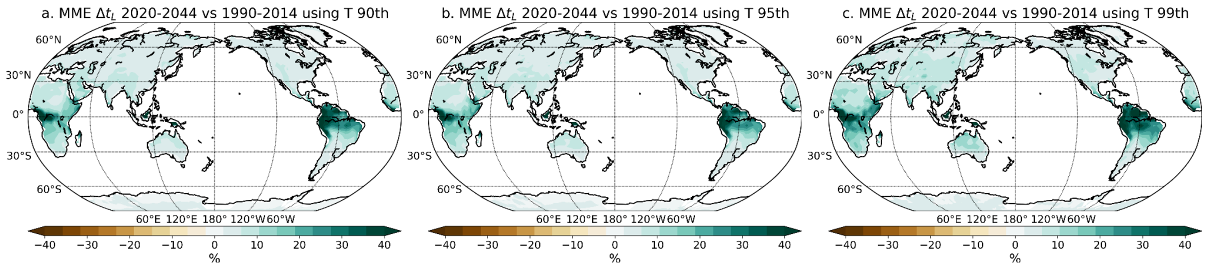
Supplementary Figure 4. Probability of exceedance (p_{fut}) as a function of $\frac{\Delta T}{\sigma_{his}}$, as derived in equation S2.



Supplementary Figure 5. Ratio $\frac{t_{99}^{fut}}{t_{99}^{his}}$ 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 τ_c timescales.



Supplementary Figure 6. 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 7. 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.