



Increasing probability of record-shattering climate extremes

E. M. Fischer  , S. Sippel  and R. Knutti 

Recent climate extremes have broken long-standing records by large margins. Such extremes unprecedented in the observational period often have substantial impacts due to a tendency to adapt to the highest intensities, and no higher, experienced during a lifetime. Here, we show models project not only more intense extremes but also events that break previous records by much larger margins. These record-shattering extremes, nearly impossible in the absence of warming, are likely to occur in the coming decades. We demonstrate that their probability of occurrence depends on warming rate, rather than global warming level, and is thus pathway-dependent. In high-emission scenarios, week-long heat extremes that break records by three or more standard deviations are two to seven times more probable in 2021–2050 and three to 21 times more probable in 2051–2080, compared to the last three decades. In 2051–2080, such events are estimated to occur about every 6–37 years somewhere in the northern midlatitudes.

Society has often been surprised by the magnitude by which recent climate extremes exceeded previous observed records, such as during the extreme rainfall of Hurricane Harvey^{1–3}, the 2020 warm anomaly over Siberia⁴ or the 2003 European and 2010 Russian heatwaves^{5,6} that caused tens of thousands of heat-related fatalities^{6–8}. Week-long temperature maxima during the latter heatwaves exceeded previous records by more than two standard deviations (σ) (Extended Data Fig. 1) and thereby literally ‘shattered’ previous records by a large margin.

Such unprecedented events need to be taken into account when designing critical infrastructure, such as power plants, or heatwave preparedness strategies. In contrast to most of the scientific literature, we here take a complementary perspective and do not define the intensity of extremes as anomalies relative to pre-industrial or present-day climates but specifically focus on the margin by which previous records are exceeded, with the records being updated with every occurrence. Events that break previous local records by large margins are hereafter defined as record-shattering extremes and their intensity quantified as the standardized anomaly by which the previous record is exceeded.

We argue that the record-shattering nature of extremes has not received much attention but is very relevant for impacts, as there is a tendency to adapt at most to the highest anomalies experienced during a lifetime^{9–11} or documented in observational or historical archives. There is a lack of methods to quantify whether record-shattering events of much higher intensities than observed are possible or plausible today and in the near future. Most of the literature focuses on moderate extremes that occur several times a year, or every few years, or record-breaking events that often only marginally exceed previous events^{12–15}. Climate models suggest that record-breaking monthly and seasonal events increase particularly over low latitudes¹⁶. It is known that the signal-to-noise ratio is high for seasonal averages leading to rapid increase in record-breaking seasonal events¹⁶ but changes in record-shattering events at typical synoptic time scales of heatwaves of about 1–2 weeks are unknown.

Record-shattering events in climate models

In a gradually warming climate with a century of observations, one would not necessarily expect previous temperature records

to be broken by large margins. Nevertheless, large climate model ensembles simulate individual events in the near future that shatter previous records set over a period of 175 yr (Fig. 1a), with record margins much larger than the ones of recent heatwaves (Extended Data Fig. 1). In the illustrative case shown in Fig. 1a, simulated in a 84-member Community Earth System Model version 1.2 (CESM1.2) ensemble forced with representative concentration pathway RCP8.5 (Methods), the hottest week (Tx7d; Methods) over Central North America exceeds the previous record in the respective simulation by $>5\sigma$. In the corresponding model realization the trend in Tx7d before 2019 is seemingly small. Heat extremes observed over North America also show no clear trend in the last decades^{17,18}. It is unclear to what extent this ‘warming hole’ behaviour is forced or due to variability^{18–20} (Extended Data Fig. 1b). Since observations only correspond to one possible realization of climate, it is challenging to identify the underlying forced warming trend.

In the model context, the forced warming trend can be estimated by averaging all 84 members. The 84-member single-model initial condition large ensemble reveals that internal variability obscures the forced warming trend before the record-shattering event in the corresponding simulation (Fig. 1a). If the forced warming is taken into account, the anomaly of the event in Fig. 1a relative to the warming background climate turns out to be about 1.7°C smaller than relative to the historical period 1951–2019 in the corresponding simulation. If extremes are defined in the traditional way as absolute temperature anomalies relative to a climatology²¹, their statistically expected return periods rapidly decrease as a result of warming. What used to be a 1-in-1,000-yr event for the period 1951–2019 is about a 1-in-100-yr in a period centred at 2020 and a 1-in-40-yr event in the mid-2020s (Fig. 1b), the decade in which the event illustrated in Fig. 1a occurs. These ratios are consistent with other single-model initial condition large ensembles (CanESM2 shown in Extended Data Fig. 2) and with previous studies at seasonal scale demonstrating that record-breaking summer seasons of the recent past become the norm within a few decades^{22–25}.

However, in contrast to the continuous warming of the background climate and increase of statistically expected return levels, the sequence of record-breaking extremes that unfolds in reality usually does not lead to gradually increasing anomalies. Instead

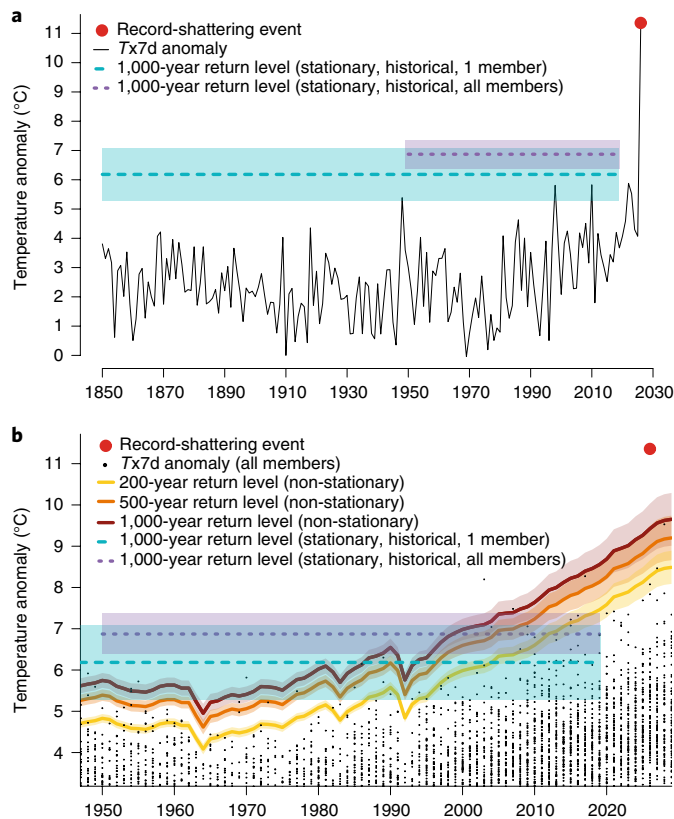


Fig. 1 | Illustrative example of a simulated record-shattering event in Central North America. **a**, Record-shattering weekly heatwave (red dot) as simulated by one selected realization (black line) of an 84-member CESM1.2 initial condition large ensemble. The black line shows simulated Tx7d calculated from Central North America average temperatures. The Tx7d anomalies are shown relative to the mean of June, July, August (JJA) mean of daily maximum temperatures 1986–2005. Best estimate of 1,000-yr return period estimated from a stationary GEV fit to the selected member (1850–2019) (dashed turquoise line) and from all 84 members (dotted violet line) (1950–2019) along with 95% CI. **b**, The Tx7d of all 84 members of the CESM1.2 initial condition large ensemble (black dots). Yellow, orange and dark red lines illustrate the non-stationary GEV estimates for 200-, 500- and 1,000-yr return periods from all members using global mean temperature as a covariate for the location and scale parameter (Methods). Stationary 1,000-yr return period (turquoise and violet lines) as in Fig. 1a.

of record-breaking events in which the previous records are broken by one- or two-tenths of a degree every few years, simulations sometimes show stagnation periods in which records are not, or only marginally, broken for several decades followed by an abrupt record-shattering extreme event such as the one illustrated in Fig. 1a.

Understanding the physical mechanisms

But is a record-shattering event such as the one illustrated in Fig. 1 plausible or not? Since the record margin and anomaly is larger than in any event observed over the region, it is challenging to rigorously demonstrate plausibility with observations. We here compare the underlying mechanisms with observed record-shattering events of smaller magnitudes. Figure 2 shows that the simulated event is associated with a large-scale weekly temperature anomaly pattern matching the area analysed and at some gridpoints reaches anomalies of around 18°C relative to the summer mean temperature for 1986–2005. This anomaly is much higher than the Tx7d anomaly during the 1995 Chicago heatwave^{18,26} (Extended Data Fig. 3b) and higher

than the 2010 Russian heatwave⁶ (Tx7d anomaly of 15°C; Extended Data Fig. 3c). The event illustrated in Fig. 1a shows a pronounced geopotential height anomaly over the Great Lakes, similar to the 3-d anomaly observed during the Chicago heatwave²⁷, yet further south and much more intense and more persistent (Extended Data Fig. 3e). The event also shows anomalously high incoming shortwave radiation and anomalously low precipitation and soil moisture during the month before the event (Fig. 2d–g). Consequently, dry soils lead to low evaporative fraction, particularly in the southern half of Central North America (Fig. 2c). Nevertheless, the model does not dry out excessively, which used to be a common bias of previous generations of land surface models²⁸. The anomalies of 500-hPa geopotential height, soil moisture and evaporative fraction in the month before and during the event are among the most extreme of all years in the 84-member ensemble even if the warming background climate is taken into account (Fig. 2d–g) but do not clearly deviate from the relationship of the bulk of all events. Thus, a combination of extreme anomalies of common drivers are coming into play during this record-shattering event (as in an ordinary event projected by the end of the century in the context of the RCP8.5 projection for the twenty-first century (Extended Data Fig. 4)) rather than different feedback mechanisms that are unique to such very rare extremes. The anomalies of the potential driving mechanisms are consistent with the long-term relationship between temperature anomalies and anomalies in contributing factors relative to a fixed historical climate (Extended Data Fig. 5). In summary, the event has drivers common to other heatwaves^{4,29–32} and cannot be ruled out as being implausible given the short observational record but more research is needed to understand whether such simulated events can act as plausible narratives for record-shattering events possible in the near future.

The event shown corresponds to the second most extreme anomaly relative to the background climate and the event that breaks the previous record by the largest margin in any of the 84 members over that region (another similar event is given in Extended Data Fig. 4b). Similar record-breaking events yet somewhat less pronounced are also simulated in different large ensembles and other regions (for example, over Central Europe (Extended Data Fig. 6)), raising the question of the overall probability of such record-shattering events and how it changes with warming.

Probability of future record-shattering extremes

In the following, we demonstrate that the changing probability of record-shattering extremes in a warming climate can be understood with events of lower magnitude and higher sample size. A comprehensive gridbox-based analysis across five initial condition large ensembles and Coupled Model Intercomparison Project Phases CMIP5 and CMIP6 multimodel ensembles (Methods) reveals that record-shattering extremes are very rare in the late twentieth and early twenty-first centuries but their expected probability increases rapidly in the coming three decades. In the CESM1.2 ensemble for any given year between 1991 and 2020 there is about a 4.5% (95% confidence interval (CI) of mean: 3.8%; 5.3%) probability for a 2 σ -record event somewhere on land in the northern midlatitudes; that is, an event that breaks the previous record by >2 σ , as for the 2003 and 2010 heatwaves. In RCP8.5, the probability increases to 22.4% (21.0; 23.7%) per year in 2021–2050 and 50.5% (48.6%; 52.5%) per year between 2051–2080 (Fig. 3a). Even for 3 σ -record and 4 σ -record events the probability reaches 17.2% (16.0%; 18.5%; Supplementary Table 1) and 5.5% (4.8%; 6.3%) in any given year, respectively, between 2051 and 2080. These probabilities imply that an event nearly as intense as illustrated in Fig. 1a is expected to occur about once every two decades somewhere in the northern midlatitudes (Fig. 3a). The exact probability of these record-shattering events differs between the five single-model initial condition large ensembles compared here (Fig. 3c–e and

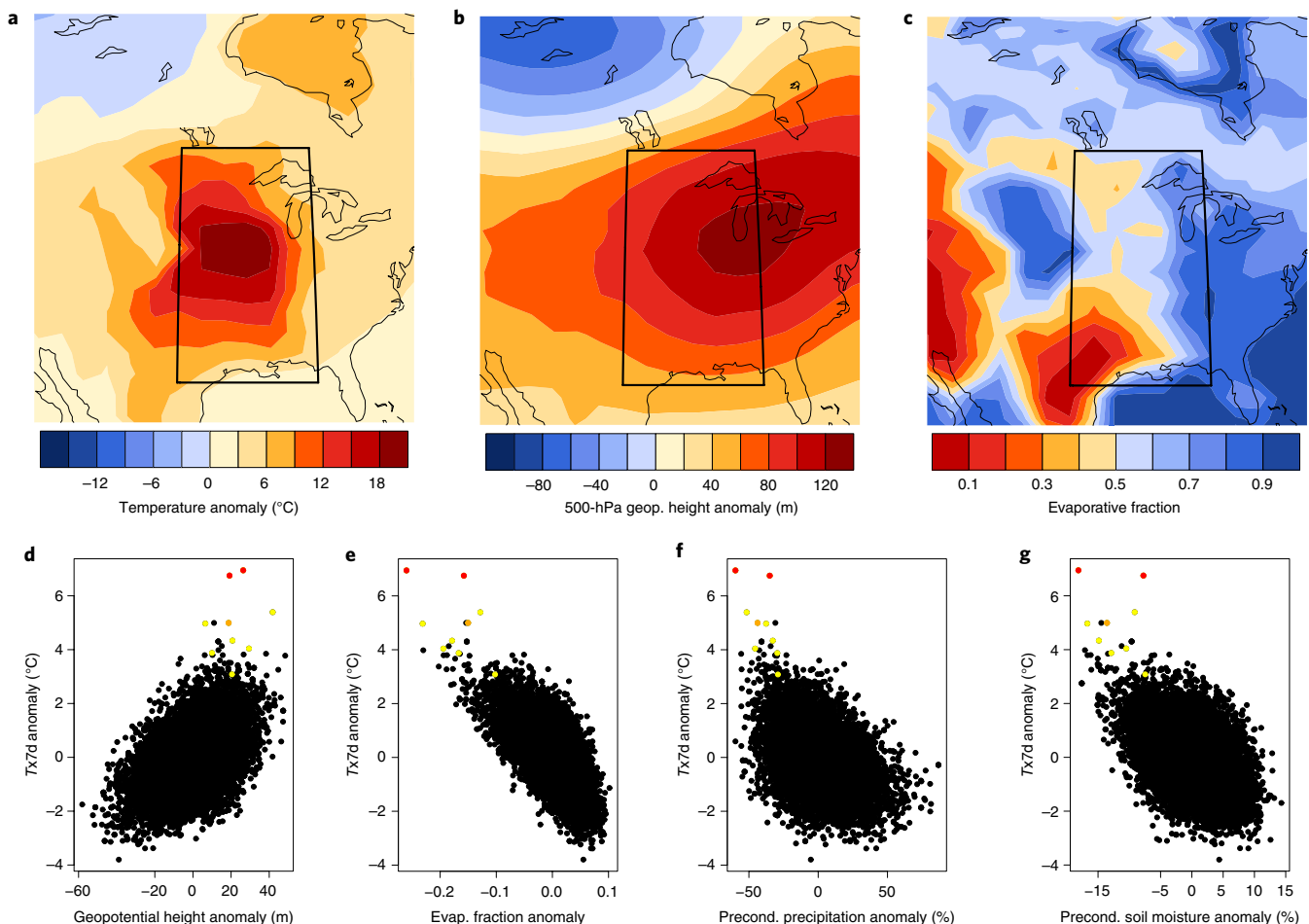


Fig. 2 | Drivers of record-shattering heatwave. **a,b**, Anomalies of 7-d mean of daily maximum temperature (**a**) and of 500-hPa geopotential height (**b**) during the event illustrated in Fig. 1a relative to the JJA average conditions in the period 1986–2005. **c**, Evaporative fraction during the event. **d–g**, Annual maximum 7-d temperature anomaly averaged across Central North America (region indicated in Fig. 2a–c) against anomalies of 7-d average geopotential height (**d**) and evaporative fraction (**e**) during the event and monthly average precipitation (**f**) and monthly soil moisture (**g**) in the month before the event. Anomalies in **d–g** are shown relative to the warming background climate; that is, relative to a moving 15-yr multimember mean. Record-shattering events are shown in yellow (2σ records), orange (3σ records) and red (4σ records).

Supplementary Table 1) and the probabilities in CESM1.2 tend to be at the high end of the model range. For instance, in 2051–2080, depending on the large ensemble, a 3σ-record event is projected about every 9–37 yr anywhere in the northern midlatitudes. All large ensembles consistently show a rapid increase in the probability with 3σ records becoming two to seven times and three to 21 times more probable than over the last three decades in 2021–2050 and 2051–2080, respectively (Supplementary Table 1 gives exact numbers for individual large ensembles). Particularly for the 3σ-record and 4σ-record events, the role of internal variability is large (see 95% CI across members in Fig. 3a,c–e) and the frequency between individual realization differs substantially, which suggests that a substantial fraction of the range across CMIP5 and CMIP6 models is due to internal variability. Thus, a robust quantification of the expected probability of record-shattering extremes requires single-model large ensembles.

CMIP5 and CMIP6 and a 50-member CanESM5 ensemble are used here to quantify the dependence on the emission pathway and reveal that the probability of record-shattering events is substantially reduced in lower emission scenarios. On average, across 16 CMIP6 models, the probability of experiencing any 3σ records is reduced by about 48% in Shared Socioeconomic Pathways SSP3–7.0 relative to SSP5–8.5 and by about a factor of 4.3 in SSP1–2.6.

The hotspot areas of the largest probability of record-shattering events vary across models but consistently include densely populated areas such as the eastern United States, central Europe, eastern Asia or parts of tropical South America and Africa (Fig. 3b and Supplementary Fig. 1). Part of the differences between regions and models are explained by a simple signal-to-noise ratio consideration particularly for 1σ-record events. Gridpoints with a higher ratio between linear warming rate during the period 2051–2080 and year-to-year variability have a larger probability of record-shattering events (Supplementary Fig. 2; correlation ranges up to 0.84 depending on the models). The tropics have previously been highlighted as a hotspot for monthly to seasonal record-breaking extremes, a result of the typically low internal variability¹⁶. Here, we specifically highlight the role of the northern midlatitudes, where the probability and intensity of week-long record-shattering extremes is remarkably high despite the high year-to-year variability, and explore the role of the corresponding Tx7d variability and potential changes in variability.

Quantifying the role of the warming rate

In the following, we show why the expected probability of record-shattering events increases so dramatically and, in contrast to the expected probability of hot extremes defined as anomalies

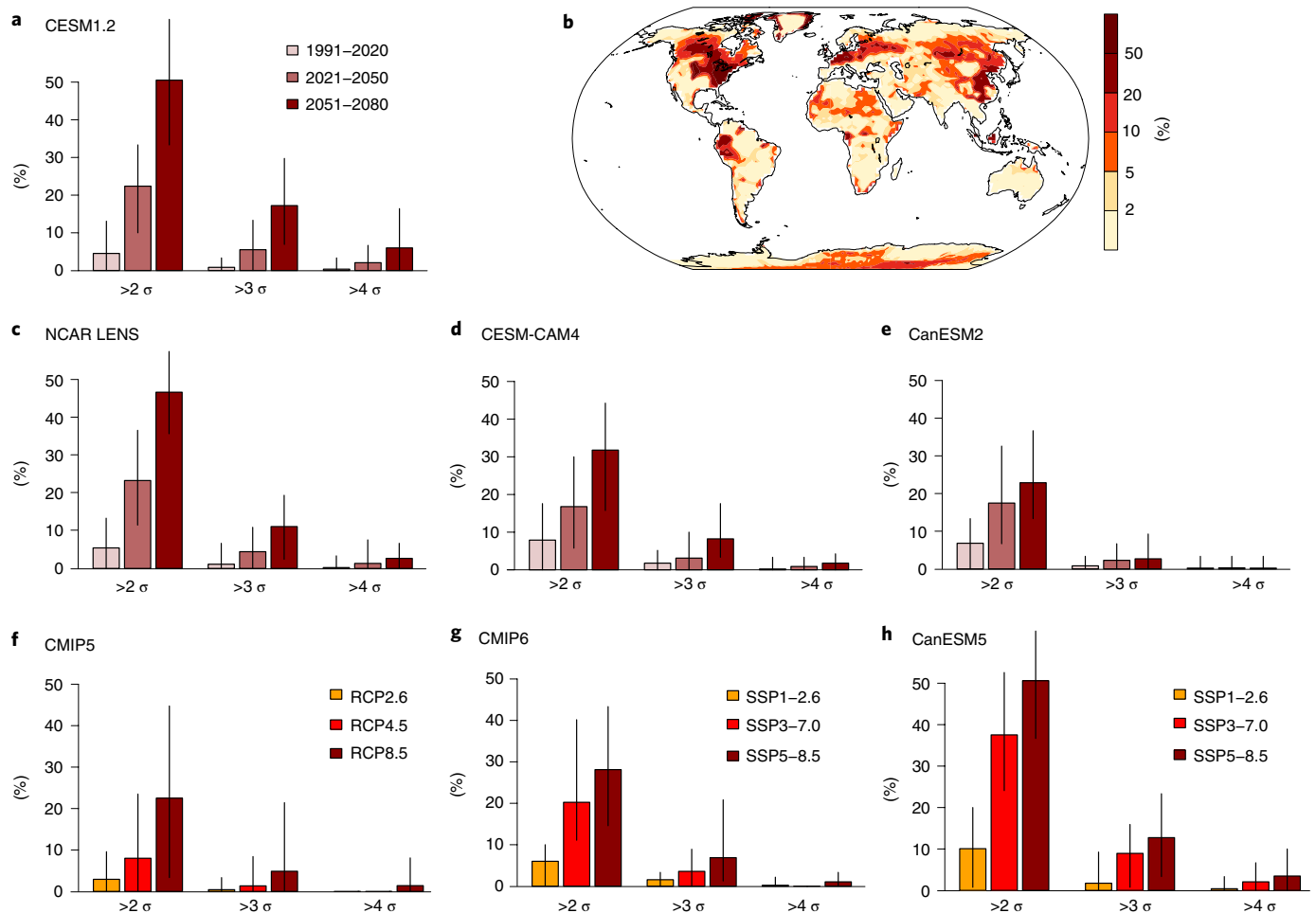


Fig. 3 | Projected occurrence of record-shattering extremes in northern midlatitudes. a, Annual probability of at least one record-shattering event per year anywhere over a land area larger than 70,000 km² in the northern midlatitudes (30–65°N) in the 84-member CESM1.2 ensemble for three different periods and event magnitudes in RCP8.5. **b**, Probability of at least one record-shattering event that breaks the previous record in the respective simulation by at least two standard deviations during the period 2051–2080. **c–e**, Same as **a** but for the NCAR LENS (**c**), the CESM-CAM4 (**d**) and the CanESM2 ensemble (**e**). **f–h**, Probability of at least one record-shattering event per year anywhere over the northern midlatitudes land regions (30–65°N) for different emission scenarios (colours) in the CMIP5 (**f**), the CMIP6 (**g**) multimodel ensemble and the CanESM5 ensemble (**h**) for the period 2051–2080. The vertical black lines show the 95% CIs across all initial condition members of the large ensembles in **a**, **c–e** and **h** and across all CMIP5 and CMIP6 models in **f** and **g**, respectively.

relative to a climatology, does not depend on the level of global warming but rather on the rate of warming. The probability for a record-breaking event in any given year n in a time series is a function of (1) the probability that no value equal to or larger than the event threshold has occurred in the complete past record and (2) the probability that the threshold is actually exceeded in this particular year³³. We here adjust the analytical solution for record-breaking events³³ to account for different margins $c \geq 0$ by which a threshold is broken and for different underlying forced responses (Methods). We obtain the probability $P(t_n)$ of an extreme exceeding the previous record by margin c to occur at a given timestep t_n for the most general case as

$$P(t_n) = \int_{-\infty}^{\infty} f(x+c, t_n) \prod_{i=1}^{n-1} F(x, t_i) dx, \quad (1)$$

where $f(x+c, t_n)$ denotes the probability density function at time t_n and $F(x, t_i)$ the cumulative distribution function at time t_i , respectively.

For a Gaussian distribution, the respective equation reads

$$P(t_n) = \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi\sigma_{t_n}^2}} e^{-\frac{(x+c-\mu_{t_n})^2}{2(\sigma_{t_n})^2}} \right] \prod_{i=1}^{n-1} \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{x-\mu_{t_i}}{\sigma_{t_i} \sqrt{2}} \right) \right) dx \quad (2)$$

where c is the margin by which the previous record is broken (in original units) and μ_t is the time-dependent average of the distribution (that is, given by the forced response), which could be linear³³ or more complex as in the examples below. The standard deviation of the distribution may be assumed constant in time (σ or time-dependent (σ_{t_i})). Equation (1) can be also used to analytically describe the probability of record-breaking or record-shattering extremes for random variables that follow extreme value distributions (Supplementary Information).

Whether the increase in record-shattering extremes can be understood with the analytical solution given in equation (2) assuming random uncorrelated variables can be tested. We compare annual probability of 1σ records for the illustrative example

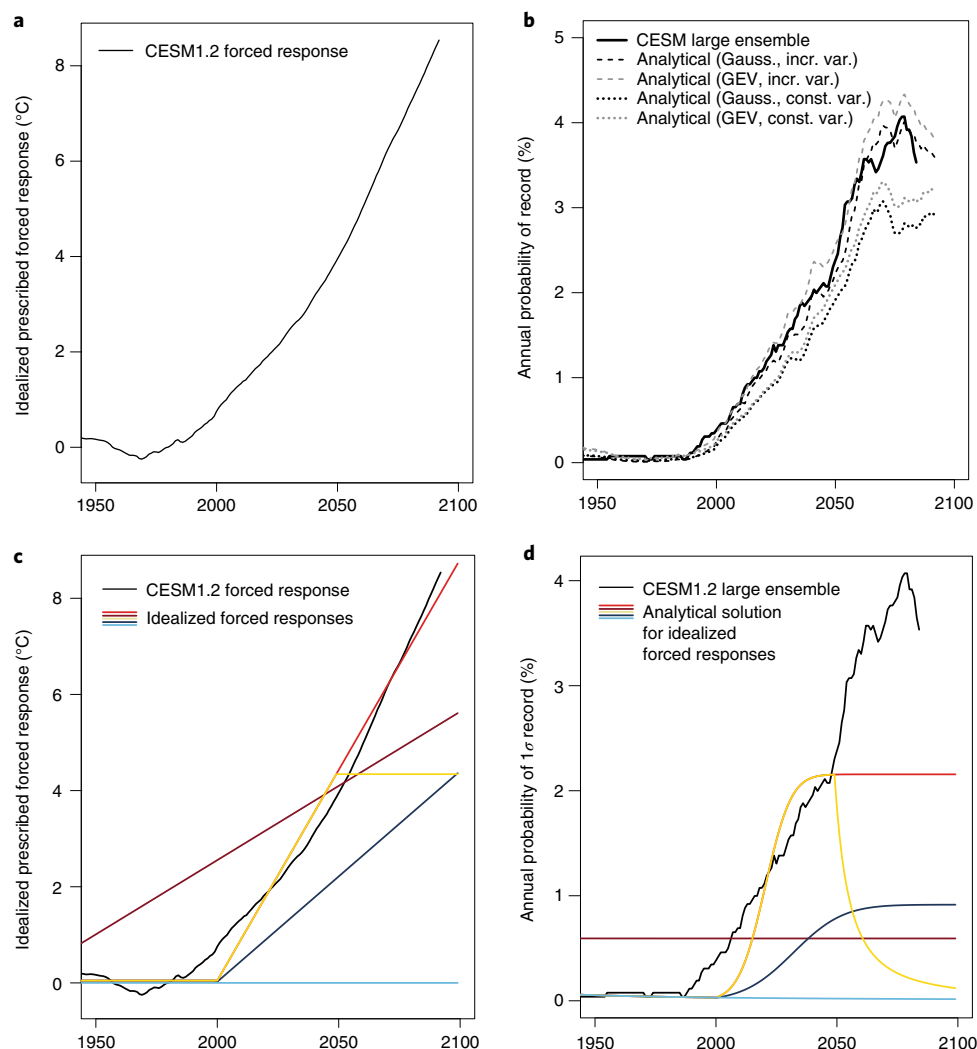


Fig. 4 | Path-dependence of occurrence of record-shattering extremes. **a**, Forced response of area-average Tx7d across Central North America (multimember mean μ_t , time series smoothed using a 15-yr moving average). **b**, Empirical annual probability of a 1σ record-shattering event calculated from (solid line) 84-member CESM1.2 large ensemble (time series smoothed with 31-yr moving average). Dashed and dotted black lines show annual probabilities based on the analytical solution (equation (2)) using the forced change in the mean (μ_t), shown in **a** and for a constant standard deviation σ (dotted line) and for a time-dependent standard deviation as simulated by the climate model (σ_t). The standard deviation σ is derived from the residuals of the 84-member CESM1.2 after the subtraction of the forced response μ_t for the period 1961–2090 (σ is also used as the threshold that defines a 1σ record-shattering event for the entire period). The time-evolving standard deviation σ_t is extracted for each year from the climate model ensemble's residuals after the subtraction of the forced response μ_t and smoothed using a 31-yr moving average. Dashed and dotted grey lines show the same but assuming a GEV distribution. The GEV distribution has been fitted with a time-dependent location parameter (proportional to the forced response shown in Fig. 4a) and a constant (grey dotted line) and time-dependent (grey dashed line) scale parameter. **c**, As in **a** but with additional idealized forced responses. **d**, Annual probability of 1σ record for analytical solutions based on the Gaussian distribution accounting for the corresponding idealized warming pathways shown in **c**. See also Supplementary Figs. 3–5 for more detailed illustrations for varying the idealized forced responses and the effect of forced changes in the standard deviation and Supplementary Fig. 7 for the effect of the shape parameter in GEV-distributed variables.

of Central North America. In CESM1.2, the probability of these area-average events over Central North America rapidly increases from very low values in the twentieth century to about a 1.5% chance per year in 2030, 2% around by 2050 and about 4% by 2080 in RCP8.5 (Fig. 4b), which corresponds to an increase by about a factor of 50. The analytical solution quantifying the probability increase in record-shattering extremes captures the one simulated by the climate model well (Fig. 4b, dotted line) but tends to underestimate the change in the second-half of the twenty-first century. When the projected variance increase is further taken into account, the probability change is very well captured (Fig. 4b, long dashed line). Note that the variance increase over Central North America is particularly pronounced in CESM1.2. Other large ensembles

analysed show a weaker or no variance increase. When assuming a non-stationary generalized extreme value (GEV) distribution with time-varying location and scale parameter for the random variables, rather than a Gaussian distribution, the simulated probability increase is also captured well though slightly overestimated (Fig. 4b and Supplementary Fig. 7).

We next test how the probability of record-shattering extremes changes with idealized forced responses (Fig. 4c and Supplementary Figs. 3 and 4). In the absence of forced warming, the probability for an event that exceeds the previous record by 1σ after a 150-yr period starting in 1850 would be extremely low (Fig. 4d, light blue line). For a linear warming trend and assuming constant variance, the event probability converges to a constant in time (Fig. 4d,

brown line) with higher warming rates implying higher probability (Supplementary Fig. 3). If a period of no or weak warming is followed by a rapid linear warming period (Fig. 4c, dark blue line) the event probability rapidly increases (Fig. 4d, dark blue line) and converges faster for higher warming rates (Supplementary Fig. 3). For a quadratic warming (Supplementary Fig. 4) similar to the accelerating warming rate in RCP8.5, the probability rapidly and continuously increases (Supplementary Fig. 4), which is consistent with the rapid increase in event probability projected by all models considered (Fig. 3). Increasing variability as simulated in CESM1.2 can further enhance the probability (Supplementary Fig. 5). Note that if temperatures stabilize after a period of increase, the event probability would not remain constant but decrease quickly (Fig. 4d, yellow line) consistent with the case of no warming. The dependence of the warming rate is consistent for 2σ -record events, with the probability being substantially lower but the relative increase is higher the more intense the record-breaking margin c . Very similar results were found when fitting an AR(1) model to the residuals as autocorrelation is very small. The results are remarkably consistent if the residual variability is estimated from ERA5 (Supplementary Fig. 6) instead of the estimate of the climate model.

Conclusions

Previous studies have demonstrated that the level of global warming primarily determines the return periods and return levels of temperature and precipitation extremes and that, in many models, these changes are independent of the emission pathways^{34–36}. In contrast, for the type of record-shattering extremes emphasized here, where the extreme is placed in the context of previous record events rather than a reference climatology, it is the warming rate (Supplementary Fig. 3), and particularly the accelerating warming rate (Supplementary Fig. 4) after periods of little to no warming, that determines the probability of record-shattering events. Increasing year-to-year variability as projected by CESM1.2 can moderately enhance the probability (Supplementary Fig. 5). Thus, primarily due to the accelerating warming rate the probability of record-shattering extremes rapidly increases in high-emission scenarios from low values in the twentieth century to high values by the second-half of the twenty-first century. Instead, if the forced warming were stabilized (for example, consistent with SSP1–2.6 or below), the frequency and intensity of heat extremes would be higher than in the historical periods but the probability of record-shattering extremes would rapidly decline. However, that is unlikely in the next few decades and in intermediate pathways such as SSP2–4.5 the rate of warming will continue to be high and in high to very high pathways (SSP3–7.0, SSP5–8.5 and RCP 8.5) even accelerate for many decades.

The focus on record-shattering extremes should be interpreted as a complementary perspective to the traditional approach of defining extremes as anomalies to a reference period. Our findings suggests that events of much higher magnitude than in the observational record are possible today and in the near future, and the probability of very rare extremes today is much higher than local observations of the last decades seem to suggest. At the same time, climate models project that record-shattering heat extremes, which break previous records by three and more standard deviations are expected about every 9–37 yr in 2051–2080 somewhere in the northern midlatitudes in high-emission scenarios. Some of these events may be considered as events that are unthought-of when only considering the past record up to the event, analogous to the concept of black swan events, referring to events that come as a complete surprise and are outside the realm of regular expectations and have been argued to be systematically underestimated in risk assessments³⁷. Taking into account that similar drivers, but with higher anomalies, come into play as in observed events, the record-shattering events illustrated here rather qualify as grey swan

events that are foreseeable using physical knowledge together with historical data³⁸.

We argue that it is vital to further investigate the plausibility of such record-shattering events, as the impacts tend to be particularly large at the first occurrence of such record-shattering events at a certain location. There is still a tendency in society to respond and adapt to the maximum event experienced during one's lifetime (as measured in the observational record or documented in historical archives) but not more^{9–11}. If events of similar magnitude reoccur several years later, impacts may be substantially smaller^{39,40} as society had time to at least partly adapt. We show that taking into account the warming rate is vital for adaptation decisions. However, in practice, this is challenging particularly in regions such as the Midwest United States where the trend in hot extremes is small, potentially because the forced warming trend is temporarily suppressed by internal variability. Such a suppressed trend may contribute to an unusual quiescent phase or disaster gap⁴¹, increase the probability of being followed by a record-shattering extreme and at the same time lead to a lack of adaptation efforts and thus in combination to a serious systematic underestimation of the risk record-shattering events in the near future.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41558-021-01092-9>.

Received: 23 January 2020; Accepted: 27 May 2021;

Published online: 26 July 2021

References

- Risser, M. D. & Wehner, M. F. Attributable human-induced changes in the likelihood and magnitude of the observed extreme precipitation during Hurricane Harvey. *Geophys. Res. Lett.* **44**, 12457–12464 (2017).
- Emanuel, K. Assessing the present and future probability of Hurricane Harvey's rainfall. *Proc. Natl Acad. Sci. USA* **114**, 12681–12684 (2017).
- Van Oldenborgh, G. J. et al. Attribution of extreme rainfall from Hurricane Harvey, August 2017. *Environ. Res. Lett.* **12**, 124009 (2017).
- Overland, J.E. & Wang, M. The 2020 Siberian heat wave. *Int. J. Climatol.* **41**, E2341–E2346 (2020).
- Miralles, D. G., Teuling, A. J., Van Heerwaarden, C. C. & Vilà-Guerau de Arellano, J. G. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.* **7**, 345–349 (2014).
- Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M. & García-Herrera, R. The hot summer of 2010: redrawing the temperature record map of Europe. *Science* **332**, 220–224 (2011).
- Robine, J. M. et al. Death toll exceeded 70,000 in Europe during the summer of 2003. *C. R. Biol.* **331**, 171–178 (2008).
- García-Herrera, R., Trigo, R. M., Luterbacher, J., Schär, C. & Fischer, E. M. A review of the European summer heat wave of 2003. *Crit. Rev. Environ. Sci. Technol.* **40**, 267–306 (2010).
- Haden, V. R., Niles, M. T., Lubell, M., Perlman, J. & Jackson, L. E. What attitudes and beliefs motivate farmers to mitigate and adapt to climate change? *PLoS ONE* **7**, e2882 (2012).
- Blennow, K., Persson, J., Tomé, M. & Hanewinkel, M. Climate change: believing and seeing implies adapting. *PLoS ONE* **7**, e50182 (2012).
- Weber, E. U. Experience-based and description-based perceptions of long-term risk: why global warming does not scare us (yet). *Climatic Change* **77**, 103–120 (2006).
- Coumou, D., Robinson, A. & Rahmstorf, S. Global increase in record-breaking monthly-mean temperatures. *Climatic Change* **118**, 771–782 (2013).
- Meehl, G. A., Tebaldi, C., Walton, G., Easterling, D. & McDaniel, L. Relative increase of record high maximum temperatures compared to record low minimum temperatures in the U.S. *Geophys. Res. Lett.* **36**, L23701 (2009).
- Elguindi, N., Rauscher, S. A. & Giorgi, F. Historical and future changes in maximum and minimum temperature records over Europe. *Climatic Change* **117**, 415–431 (2013).
- King, A. D. Attributing changing rates of temperature record breaking to anthropogenic influences. *Earth's Future* **5**, 1156–1168 (2017).

16. Power, S. B. & Delage, F. P. D. Setting and smashing extreme temperature records over the coming century. *Nat. Clim. Change* **9**, 529–534 (2019).
17. Perkins, S. E., Alexander, L. V. & Nairn, J. R. Increasing frequency, intensity and duration of observed global heatwaves and warm spells. *Geophys. Res. Lett.* **39**, L20714 (2012).
18. Kunkel, K. E., Liang, X.-Z., Zhu, J. & Lin, Y. Can CGCMs simulate the twentieth-century “warming hole” in the central United States? *J. Clim.* **19**, 4137–4153 (2006).
19. Meehl, G. A., Arblaster, J. M. & Branstator, G. Mechanisms contributing to the warming hole and the consequent US east–west differential of heat extremes. *J. Clim.* **25**, 6394–6408 (2012).
20. Yu, S. et al. Attribution of the United States ‘warming hole’: aerosol indirect effect and precipitable water vapor. *Sci. Rep.* **4**, 6929 (2014).
21. Davison, A. C. & Huser, R. Statistics of extremes. *Annu. Rev. Stat. Its Application* **2**, 203–235 (2015).
22. Sun, Y. et al. Rapid increase in the risk of extreme summer heat in Eastern China. *Nat. Clim. Change* **4**, 1082–1085 (2014).
23. Mueller, B., Zhang, X. & Zwiers, F. W. Historically hottest summers projected to be the norm for more than half of the world’s population within 20 years. *Environ. Res. Lett.* **11**, 044011 (2016).
24. Christidis, N., Jones, G. S. & Stott, P. A. Dramatically increasing chance of extremely hot summers since the 2003 European heatwave. *Nat. Clim. Change* **5**, 46–50 (2015).
25. Moore, F. C., Obradovich, N., Lehner, F. & Baylis, P. Rapidly declining remarkability of temperature anomalies may obscure public perception of climate change. *Proc. Natl Acad. Sci. USA* **116**, 4905–4910 (2019).
26. Changnon, S. A., Kunkel, K. E. & Reinke, B. C. Impacts and responses to the 1995 heat wave: a call to action. *Bull. Am. Meteorol. Soc.* **77**, 1497–1506 (1996).
27. Meehl, G. A. & Tebaldi, C. More intense, more frequent, and longer lasting heat waves in the 21st century. *Science* **305**, 994–997 (2004).
28. Seneviratne, S. I., Pal, J. S., Eltahir, E. A. B. & Schär, C. Summer dryness in a warmer climate: a process study with a regional climate model. *Clim. Dynam.* **20**, 69–85 (2002).
29. Horton, R. M., Mankin, J. S., Lesk, C., Coffel, E. & Raymond, C. A review of recent advances in research on extreme heat events. *Curr. Clim. Change Rep.* **2**, 242–259 (2016).
30. Schumacher, D. L. et al. Amplification of mega-heatwaves through heat torrents fuelled by upwind drought. *Nat. Geosci.* **12**, 712–717 (2019).
31. Fischer, E. M., Seneviratne, S. I., Lüthi, D. & Schär, C. Contribution of land–atmosphere coupling to recent European summer heat waves. *Geophys. Res. Lett.* **34**, L06707 (2007).
32. Seneviratne, S. I. et al. Investigating soil moisture–climate interactions in a changing climate: a review. *Earth Sci. Rev.* **99**, 125–161 (2010).
33. Rahmstorf, S. & Coumou, D. Increase of extreme events in a warming world. *Proc. Natl Acad. Sci. USA* **108**, 17905–17909 (2011).
34. Seneviratne, S. I., Donat, M. G., Pitman, A. J., Knutti, R. & Wilby, R. L. Allowable CO₂ emissions based on regional and impact-related climate targets. *Nature* **529**, 477–483 (2016).
35. Fischer, E. M., Sedláček, J., Hawkins, E. & Knutti, R. Models agree on forced response pattern of precipitation and temperature extremes. *Geophys. Res. Lett.* **41**, 8554–8562 (2014).
36. Pendergrass, A. G., Lehner, F., Sanderson, B. M. & Xu, Y. Does extreme precipitation intensity depend on the emissions scenario? *Geophys. Res. Lett.* **42**, 8767–8774 (2015).
37. Taleb, N. N. The black swan: the impact of the highly improbable. *Rev. Austrian Econ.* **21**, 361–364 (2007).
38. Lin, N. & Emanuel, K. Grey swan tropical cyclones. *Nat. Clim. Change* **6**, 106–111 (2016).
39. Fouillet, A., Rey, G. & Laurent, F. Excess mortality related to the August 2003 heat wave in France. *Int. Arch. Occup. Environ. Health* **80**, 16–24 (2006).
40. Green, H. K., Andrews, N., Armstrong, B., Bickler, G. & Pebody, R. Mortality during the 2013 heatwave in England—how did it compare to previous heatwaves? A retrospective observational study. *Environ. Res.* **147**, 343–349 (2016).
41. Wetter, O. et al. The largest floods in the High Rhine basin since 1268 assessed from documentary and instrumental evidence. *Hydrol. Sci. J.* **56**, 733–758 (2011).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

Definition. $Tx7d$ is defined as the annual maxima of 7-d running means of daily maximum temperature and corresponds in simple words to the average maximum temperature of the hottest week per year. For the regional analysis, daily maximum temperatures are first averaged over the SREX region (Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, SREX) of Central North America (CNA, 28.6–50°N, 85–100°W) and Central Europe (CEU, the area enclosed by the points 48°N/10°W, 61.3°N/40°E, 45°N/40°E and 45°N/10°E) and then $Tx7d$ is calculated on the basis of this regional average time series.

Record-shattering events are defined here as $Tx7d$ anomalies ($Tx7d'$) that in year n exceed the record of the corresponding simulation from the start of the simulation in 1850 up to the year before the event $\max(Tx7d'_1, Tx7d'_2, \dots, Tx7d'_{n-1})$ by a large margin c . Since not all large ensembles start in 1850, for the model intercomparison in Fig. 3 we use the concatenated period 1950–1959 from the first ten members, instead of the period 1850–1949, as a conservative approximation to typical historical temperature records in the corresponding model. A direct comparison in CESM1.2 revealed that using 1850–1949 or 1950–1959 from ten members hardly affects the record statistics in the twenty-first century and, if anything, leads to conservative estimates of the probability of record-shattering events as 1950–1959 was on average warmer than 1850–1949. The magnitude of record-shattering event is defined as the margin c by which previous records are broken and expressed in standard deviations (σ) of annual $Tx7d$ calculated across all members for each year and averaged across the period 1961–1990. For CMIP5 and CMIP6 where only one member was available the margins c are expressed σ of annual $Tx7d$ across the pre-industrial control simulation. To quantify the magnitude of the observed events shown in Extended Data Fig. 1 and the illustrative event (Fig. 1a) the σ are calculated on the basis of the corresponding regional average $Tx7d$ time series of the corresponding dataset. Thus, the estimated variability is based on a relatively small sample size particularly for the observed cases, which may lead to an underestimation relative to the true standard deviation⁴² and thus a tendency to overestimate the record margin of the observed events shown in Extended Data Fig. 1.

Model ensembles. Single-model initial condition large ensembles are essential for a robust quantification of the probability of record-shattering events. Particularly the events that break previous records by large margins are rare and thus it is important to have a large sample size to get robust estimates of the changes in their probability. We specifically performed a very large 84-member initial condition ensemble with daily output for this project. The ensemble is performed with the NCAR-CESM model version 1.2 using the Community Atmosphere Model CAM5 run at $1.9 \times 2.5^\circ$ horizontal resolution. The model is initialized from 21 different ocean initial conditions taken from timesteps of a long pre-industrial simulation that are 20 yr apart. Twenty-one ocean initial condition members are run from 1850 to 2100 using the CMIP5 historical forcing up to 2005 and RCP 8.5 thereafter. To increase the ensemble size to a total of 84 members, three atmospheric initial condition are branched from each ocean initial condition member on 1 January 1940 by randomly perturbing the atmospheric initial temperature on each atmospheric level by a roundoff error of 10^{-13} K (as in refs. ^{43,44}). After a period of 5–10 yr the simulations reach very different atmospheric states and are here found to be independent realizations after 1950.

We here compare the results with the three additional single-model initial condition large ensembles that provide daily output for the historical forcing up to 2005 and for the RCP 8.5 scenario thereafter, referred to as CanESM2 ensemble, NCAR LENS and CESM-CAM4 ensemble hereafter. The CanESM2 ensemble⁴⁵ is a 50-member ensemble from 1950 to 2100 that is similar to our setup branched off from five ocean initial condition members in 1950 using perturbations in the atmospheric initial conditions. The NCAR LENS ensemble⁴⁶ consists of 30 members covering the period 1920–2080 starting from different atmospheric initial conditions in 1920. Finally, the CESM-CAM4 ensemble⁴⁴ is a 21-member large ensemble starting from different atmospheric initial conditions in 1950 and running up to 2100. While the latter two experiments cannot be considered as independent lines of evidence because they are performed with another model version and another native model resolution but still with a model of the NCAR-CESM model family, we argue that a comparison in such a complex statistic as the probability of record-shattering extremes is still informative and can reveal whether the behaviour is controlled by the same factor. Finally, the CanESM5 ensemble⁴⁷ uses CMIP6 forcing and consists of two times 25 members for the three emission scenarios SSP1–2.6, SSP3–7.0 and SSP5–8.5. We here combine the two 25-member ensembles with slightly different parameter sets (p1 and p2) into a 50-member ensemble because their response to forcing is found to be indistinguishable⁴⁷. As Fig. 3 illustrates, there can be substantial differences in the magnitude of the changes but it can be shown that the general increase in events and the dependence on the signal-to-noise ratio is consistent across the different large ensembles performed with NCAR-CESM models as well as with CanESM2. Furthermore, Supplementary Fig. 6 illustrates that the results of CESM1.2 for Central North America are highly consistent with a case in which the variability is estimated from ERA5 (linearly detrended $Tx7d$ residuals 1979–2018) or if the forced response is estimated from CMIP6 multimodel mean for SSP 5–8.5 instead of CESM1.2.

For CMIP6 we use daily output of the historical, SSP 1–2.6, SSP 3–7.0 and SSP 5–8.5 simulations of the following 16 models: AWI-CM-1-1-MR, BCC-CSM2-MR, CanESM5, CNRM-CM6-1, CNRM-ESM2-1, EC-Earth3, EC-Earth3-Veg, FGOALS-g3, GFDL-CM4, GFDL-ESM4, IPSL-CM6A-LR, MIROC6, MPI-ESM1-2-HR, MRI-ESM2-0, NESM3 and UKESM1-0-LL. For CMIP5 we use daily output of historical, RCP 2.6, RCP 4.5 and RCP 8.5 simulations of the 24 models ACCESS1-0, ACCESS1-3, bcc-csm1-1, CanESM2, CESM1-BGC, CMCC-CM, CMCC-CMS, CNRM-CM5, CSIRO-Mk3-6-0, EC-EARTH, GFDL-ESM2G, GFDL-ESM2M, HadGEM2-CC, HadGEM2-ES, inmcm4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MIROC-ESM, MIROC-ESM-CHEM, MPI-ESM-LR, MPI-ESM-MR, MRI-CGCM3 and NorESM1-M.

Projected probability of record-shattering extremes. To quantify the projected probability of record-shattering extremes (Fig. 3), we first regrid the annual $Tx7d$ data of all single-model initial condition large ensembles and all CMIP5 and CMIP6 models to a common $1.9 \times 2.5^\circ$ grid. For each simulation and each year we quantify the area over which a certain record-shattering magnitude (2σ , 3σ or 4σ) is exceeded. We then check for each year whether the respective event thresholds are exceeded over at least 70,000 km² land area of the northern midlatitudes and quantify the probability of at least one event of this magnitude per year. The minimum land area corresponds to an area that is typically larger than individual gridpoints, even in coarse resolution models. The bars in Fig. 3 show the 95% range across members of the large ensembles and the 95% range across models for the CMIP5 and CMIP6 ensembles. The uncertainties are documented in more detail in Supplementary Tables 1 and 2 and the corresponding text.

Extreme value analysis. Return periods and return levels for $Tx7d$ in Fig. 1b are calculated by fitting a non-stationary GEV using global annual mean temperature as a covariate of the location parameter μ and the scale parameter σ to account for the warming. The role of a constant versus warming-dependent scale parameter σ has been tested but only plays a secondary role for the change in return periods. In the illustration in Fig. 1, the location and scale parameter are a function of global mean temperature whereas the shape parameter ξ is kept constant. Since we are focusing on very high return level events we sample maxima over a block size of 84 yr for the CESM1.2 ensemble. Specifically, we draw the maximum of each model year across 84 ensemble members of the CESM1.2 large ensemble (Fig. 1b) and across 50 members of the CanESM2 ensembles, respectively (Extended Data Fig. 2). Note that the serial autocorrelation is very small. We also repeat the analysis with a block size of 20 members and find very consistent results. By using multiyear block sizes we minimize the risk of using too-small real block sizes^{21,48} as the hottest week of a year tends to occur in the same few weeks of the year. For the period around 2020–2030, we compare the results with empirical estimates directly sampling from the distribution of members and find very consistent results.

Analytical solution. It has been demonstrated that for independent, identically distributed and thus stationary variables the probability of breaking a record in a given year or timestep t_n is inversely related to the number of timesteps before and equals $P(t_n) = 1/n$ where n is the number of years^{13–16,49–51}.

Rahmstorf and Coumou³³ adjusted the analytical solution quantifying the probability of record-breaking events to account for a linear warming trend. They suggested that the probability of breaking a record with value x at timestep t_n can be thought of as a combination of the probability that a value x has not been exceeded in any prior timestep t_1, t_2, \dots, t_{n-1} given by $\prod_{i=1}^{n-1} F(x, t_i)$ and the ‘probability’ that it is reached at timestep t_n given by $f(x, t_n)$, with $F(x, t_i)$ denoting the cumulative distribution function and $f(x, t_n)$ the probability density function. Integrating over all values of x then yields (equivalent to equation (1)),

$$P(t_n) = \int_{-\infty}^{\infty} f(x, t_n) \prod_{i=1}^{n-1} F(x, t_i) dx$$

They then showed that for a given trend $\mu_0 + \mu_1 t$ and assuming uncorrelated Gaussian residuals the probability can be written as

$$P(t_n) = \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu_0-\mu_1 t_n)^2}{2\sigma^2}} \right] \prod_{i=1}^{n-1} \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{x-\mu_0-\mu_1 t_i}{\sigma\sqrt{2}} \right) \right) dx$$

Here we shift the focus from ‘simple’ record-breaking events to record-shattering events in which a certain record set in a previous timestep is exceeded by a certain margin c , which here is expressed in the original physical units (but then normalized by standard deviation in the formula). Thus, as for the record-breaking equation, the probability for having a record-shattering event is given by the probability that a certain threshold x is not exceeded at any previous timestep $\prod_{i=1}^{n-1} F(x, t_i)$ and the ‘probability’ to reach a record margin c above the corresponding threshold at timestep t_n given by $f(x+c, t_n)$. Furthermore, the underlying trend does not necessarily need to be a linear warming trend $\mu_0 + \mu_1 t$ but can be thought of as any forced response μ_i and σ_i might also change over time as part of a forced climate change. Thus, we can then write the probability for a record-shattering of at least the magnitude c as

$$P(t_n) = \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi\sigma_{t_n}^2}} e^{-\frac{(x+\epsilon-\mu_{t_n})^2}{2(\sigma_{t_n})^2}} \right] \prod_{i=1}^{n-1} \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{x-\mu_{t_i}}{\sigma_{t_i}\sqrt{2}} \right) \right) \partial x$$

For no warming $P(t_n)$ of record-shattering events with $c=3\sigma$ converges towards zero at a rate that is faster than for ‘simple’ record-breaking events $c=0$ (that is, the negative linear slope in log–log space is steeper than for ‘simple’ record-breaking events that decline as $p(n)=1/n$, that is converging with $\log(p(n))=-\log(n)$ in log–log space).

In summary, the generalizations to the equation given in ref. ³³ made in this paper allows us to explore three additional aspects:

1. We can assess the probability of record-shattering events that exceed previous records by a specified margin c (in original physical units), in addition to assessing the probability of record-breaking events in general.
2. The time-dependence of the mean μ_{t_n} is not required to be linear but can follow the shape of any forced response. In addition, the scale parameter (here, σ_t) of the distribution can change as well as part of a forced change in variability.
3. The general formula given in equation (1) is not restricted to Gaussian distributed random variables. In fact, we show in Fig. 4b and Supplementary Fig. 7 that equation (1) can be used to analytically describe the probability of record-breaking and record-shattering extremes in random variables that follow extreme value distributions. We illustrate in Supplementary Fig. 7 the influence of idealized forced responses in the location parameter and the effect of different tail (shape) parameters on the probability of record-shattering and record-breaking extreme events.

Data availability

All original CMIP5 and CMIP6 data, the CanESM2 and NCAR LENS ensembles, and the ERA5 reanalysis used in this study, are publicly available as follows: CMIP5 model data, <https://esgf-node.llnl.gov/projects/cmip5/>; CMIP6 model data, <https://esgf-node.llnl.gov/projects/cmip6/>; CanESM2 ensemble reanalysis, <https://open.canada.ca/data/en/dataset/aa7b6823-fd1e-49ff-a6fb-68076a4a477c>; NCAR LENS ensemble, <https://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html>; and ECMWF5 ERA5 reanalysis, <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels>. The output from CESM1.2 and CESM-CAM4 used in this analysis is available at https://data.iac.ethz.ch/Fischer_et_al_2021_RecordExtremes/.

Code availability

All computer code to reproduce the main results and all figures are available at https://data.iac.ethz.ch/Fischer_et_al_2021_RecordExtremes/

References

42. Sippel, S. et al. Quantifying changes in climate variability and extremes: pitfalls and their overcoming. *Geophys. Res. Lett.* **42**, 9990–9998 (2015).

43. Fischer, E. M., Beyerle, U. & Knutti, R. Robust spatially aggregated projections of climate extremes. *Nat. Clim. Change* **3**, 1033–1038 (2013).
44. Deser, C., Phillips, A. S., Alexander, M. A. & Smoliak, B. V. Projecting North American climate over the next 50 years: uncertainty due to internal variability. *J. Clim.* **27**, 2271–2296 (2014).
45. Kirchmeier-Young, M. C., Zwiers, F. W. & Gillett, N. P. Attribution of extreme events in Arctic Sea ice extent. *J. Clim.* **30**, 553–571 (2017).
46. Kay, J. E. et al. The Community Earth System Model (CESM) large ensemble project: a community resource for studying climate change in the presence of internal climate variability. *Bull. Am. Meteorol. Soc.* **96**, 1333–1349 (2015).
47. Swart, N. et al. The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci. Model Dev. Discuss.* **12**, 4823–4873 (2019).
48. Ana, F. & De Haan, L. On the block maxima method in extreme value theory: PWM estimators. *Ann. Stat.* **43**, 276–298 (2015).
49. Falk, M., Chokami, A. K. & Padoan, S. A. Some results on joint record events. *Stat. Probab. Lett.* **135**, 11–19 (2018).
50. Falk, M., Khorrami Chokami, A. & Padoan, S. Records for time-dependent stationary Gaussian sequences. *J. Appl. Probab.* **57**, 78–96 (2020).
51. Ahsanullah M. & Nevzorov V.B. in *International Encyclopedia of Statistical Science* (Ed. Lovric, M.) 1195–1202 (Springer, 2011).

Acknowledgements

We acknowledge funding received from the Swiss National Science Foundation (grant no. 200020_178778) (E.M.F.). We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP5 and CMIP6. We thank the climate modelling groups for producing and making available their model output. We acknowledge Environment and Climate Change Canada’s Canadian Centre for Climate Modelling and Analysis for executing and making available the CanESM2 large ensemble simulations. We acknowledge the US CLIVAR working group on large ensembles for compiling the Multi-Model Large Ensemble Archive.

Author contributions

E.M.F. and R.K. designed the study. E.M.F. performed the model analysis and S.S. implemented the analytical solution. All authors contributed to the interpretation of the results and the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

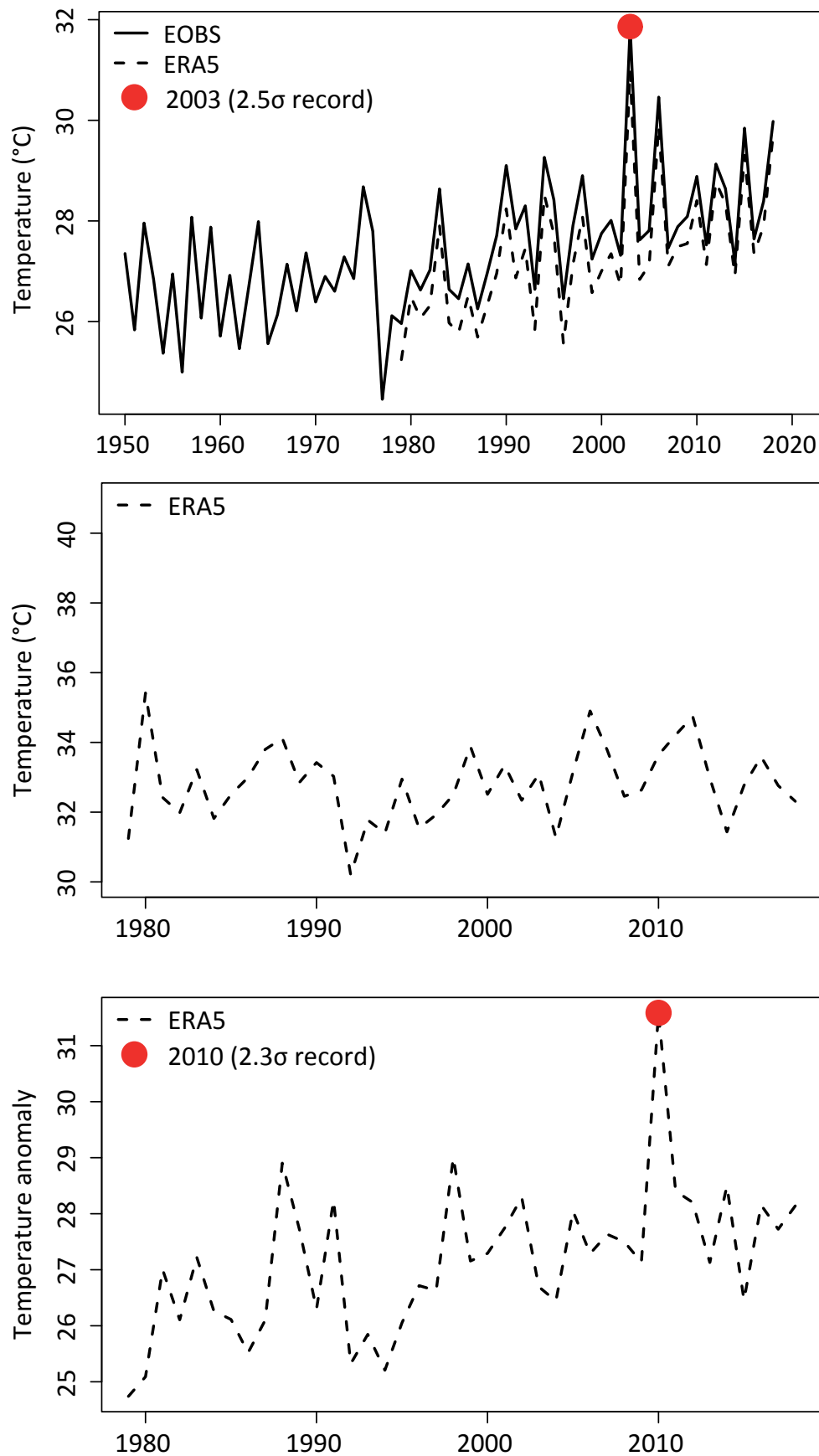
Extended data is available for this paper at <https://doi.org/10.1038/s41558-021-01092-9>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41558-021-01092-9>.

Correspondence and requests for materials should be addressed to E.M.F.

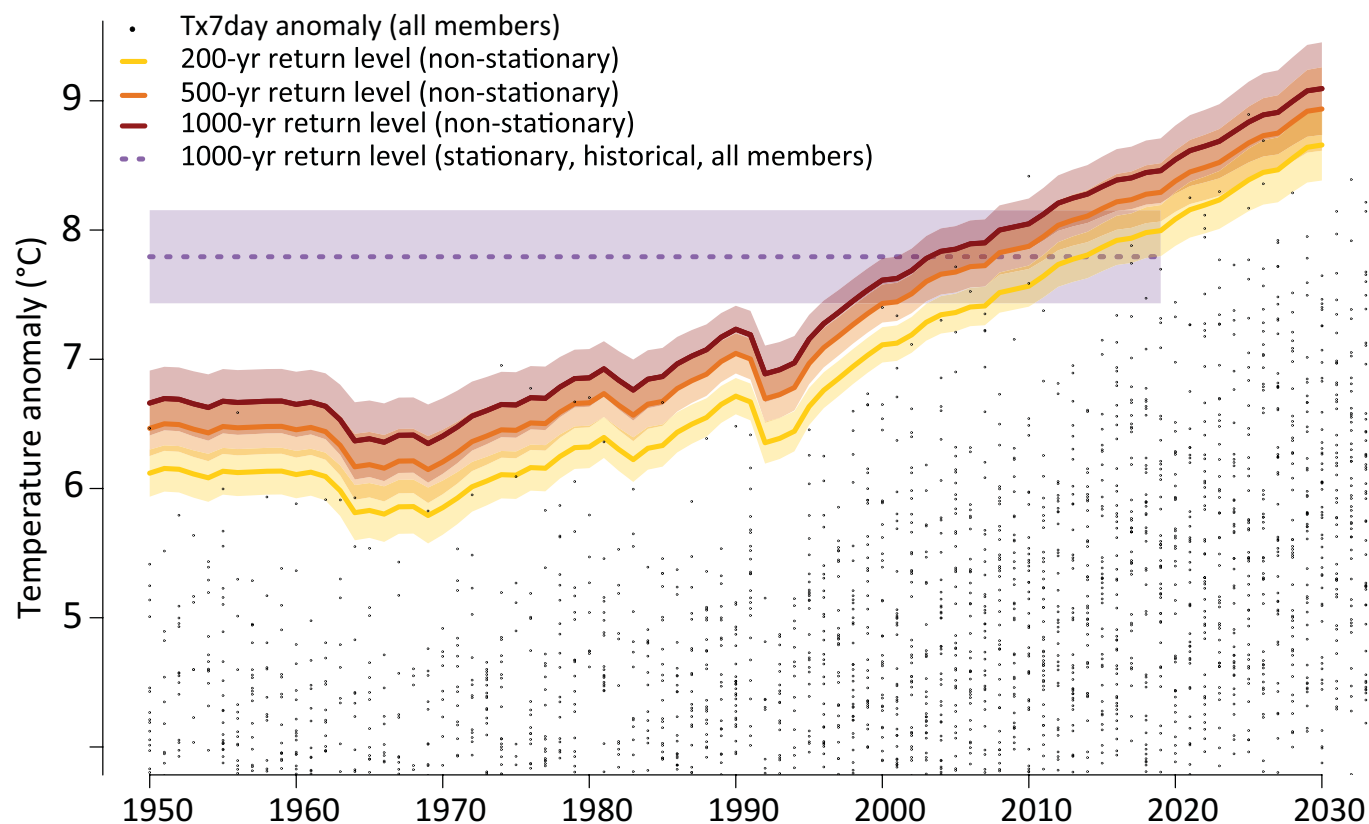
Peer review information *Nature Climate Change* thanks Raphaël Huser and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

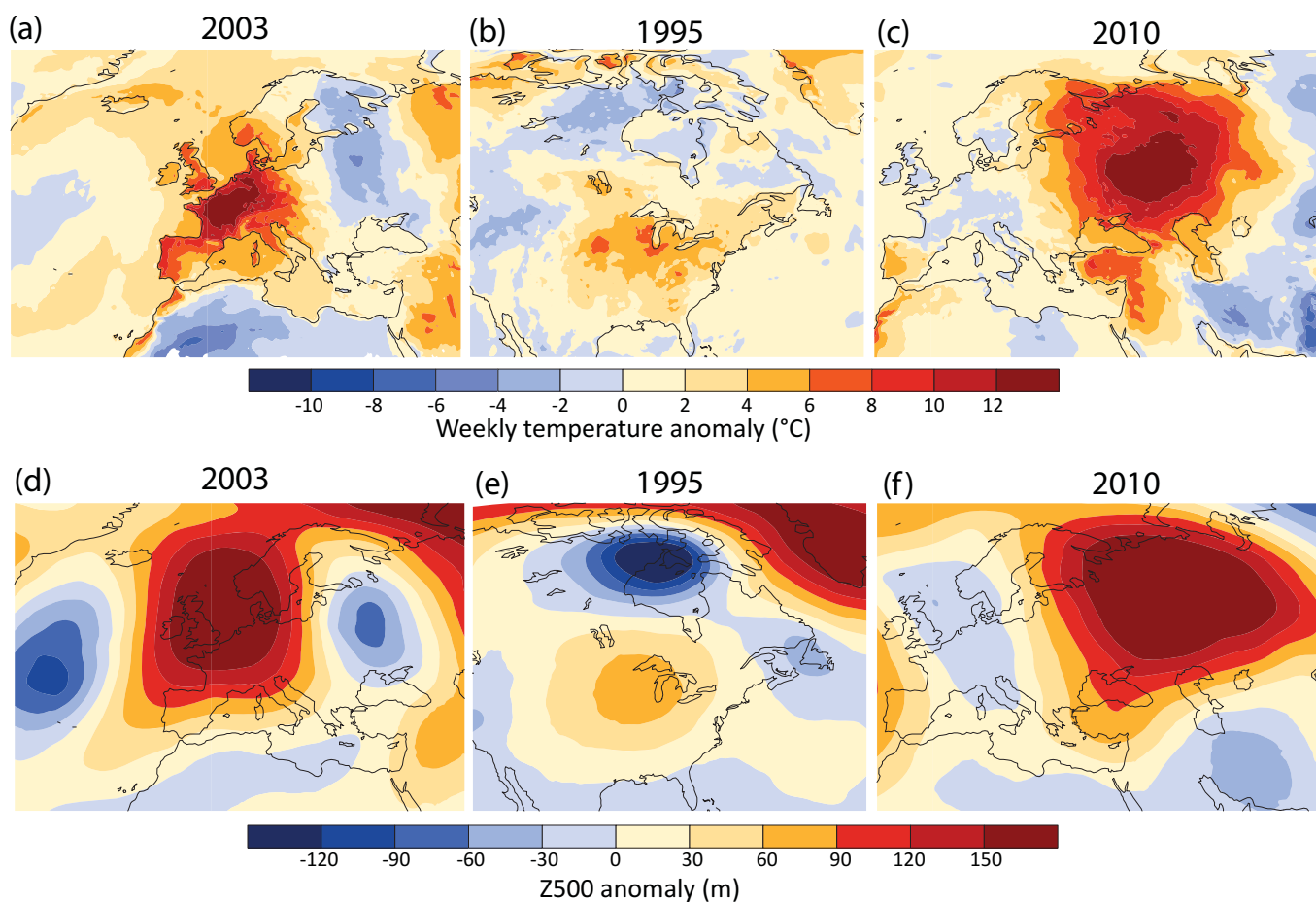


Extended Data Fig. 1 | See next page for caption.

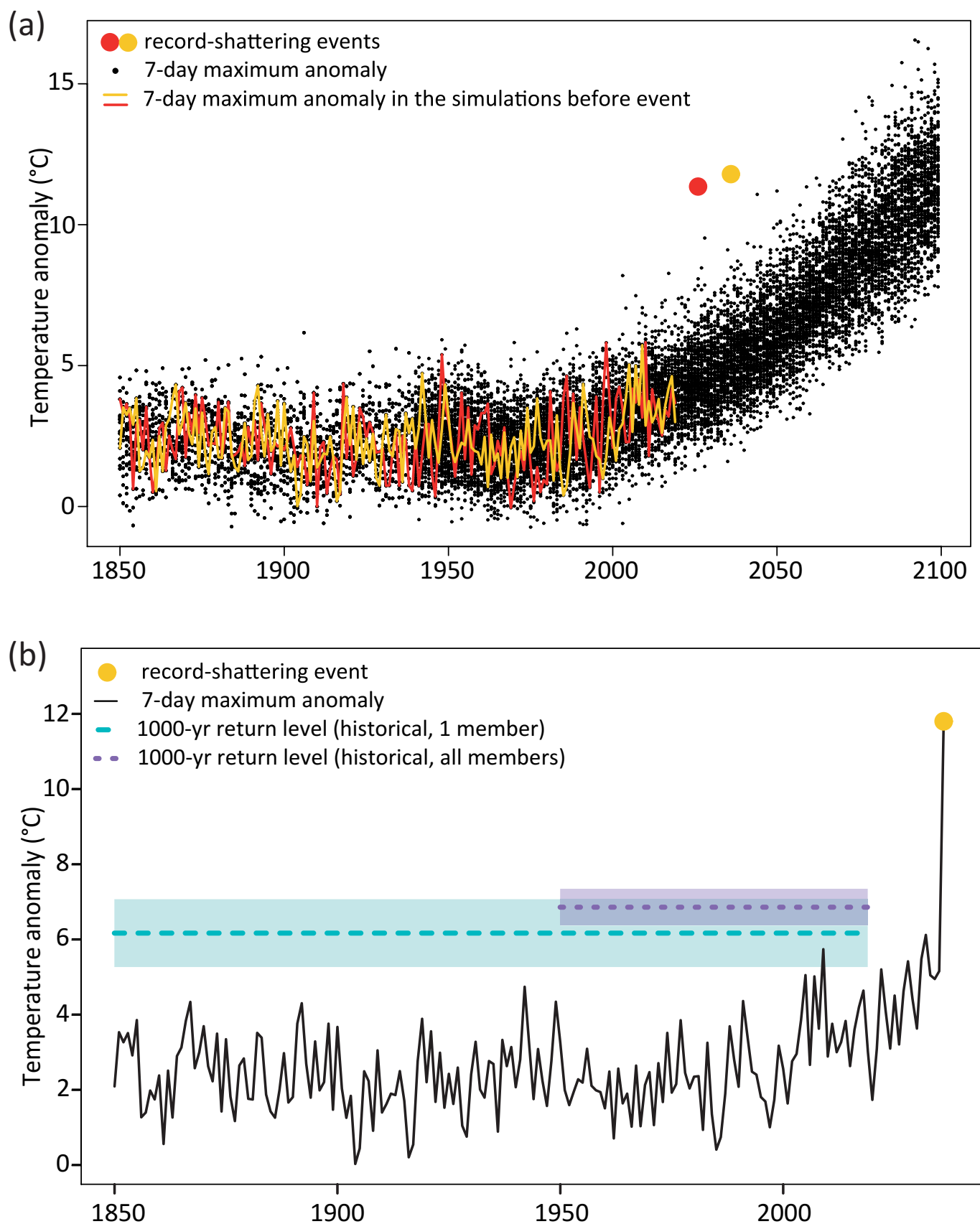
Extended Data Fig. 1 | Observed record-shattering events. Regional average Tx7day time series for (top) the European 2003 heatwave region 35–60°N and 10°W–20°E, (middle) the CNA region used in Fig. 1, (bottom) the Russian 2010 heatwave region 40–70°N and 20–70°E. The solid line shows a regional average of EOBS v19 gridded observations³⁸ and the dashed lines of ERA5 reanalysis product³⁹. The two dots illustrate two record-shattering events, and the magnitude is expressed in standard deviations of the detrended time series.



Extended Data Fig. 2 | Changing return levels in other large ensemble. Same as Fig. 1b but based on CanESM2 large ensemble. Yellow, orange and dark red lines illustrate the non-stationary GEV estimates for 200-, 500-, and 1000-yr return periods from all members using global mean temperature as a covariate for the location and scale parameter (see Methods). Best estimate of 1000-yr return period estimated from a stationary GEV fit from all 50 members (dotted violet line) (1950–2019) along with 95% CIs.

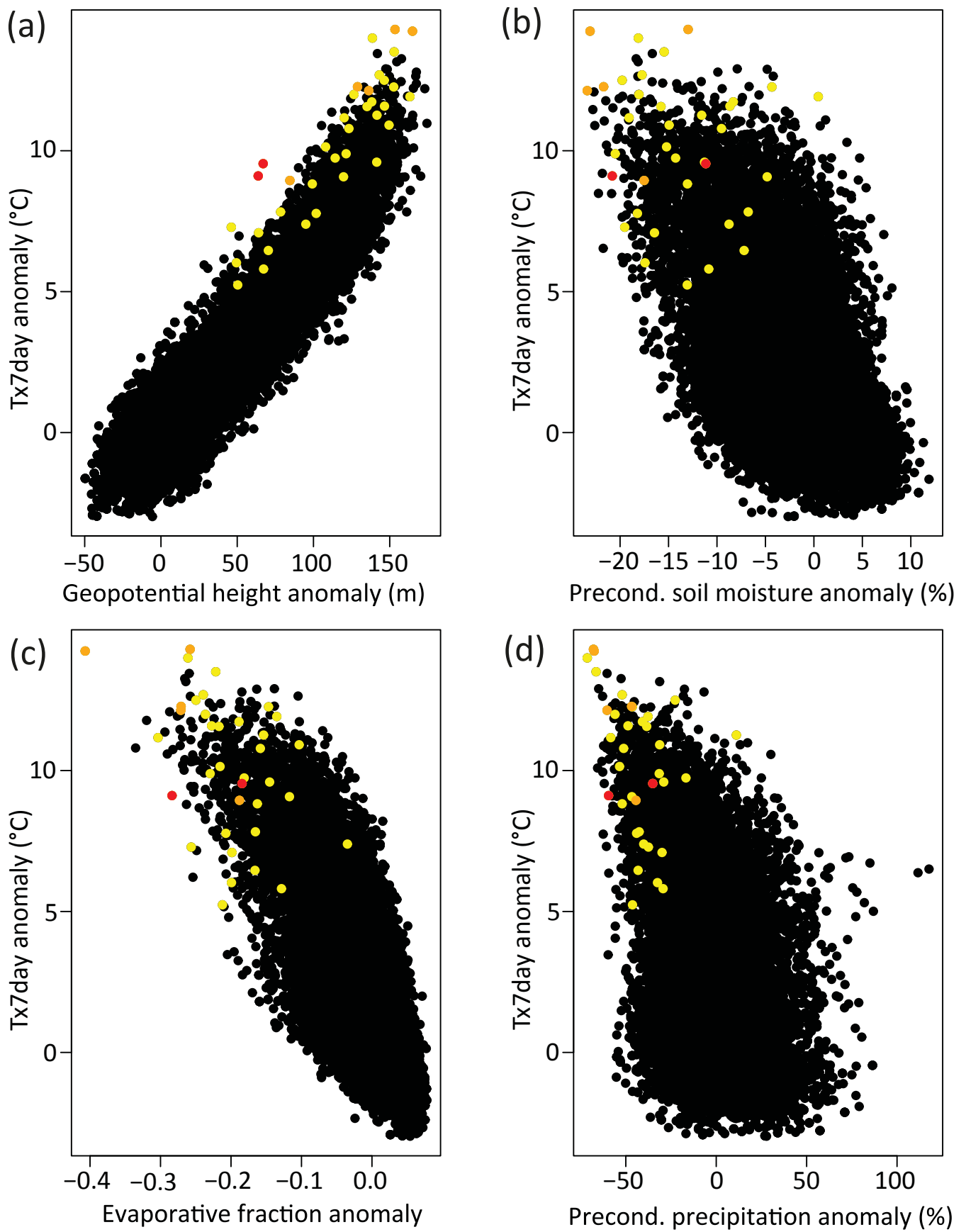


Extended Data Fig. 3 | Observed heatwaves and their drivers. (a–c) Tx7day anomaly and (d–f) associated 500 hPa geopotential during the 2003 European summer heatwave, the 1995 Chicago heatwave and the 2010 Russian heatwave. Anomalies are calculated based on the ECMWF ERA5 reanalysis and expressed as anomalies relative to the period 1986–2005.



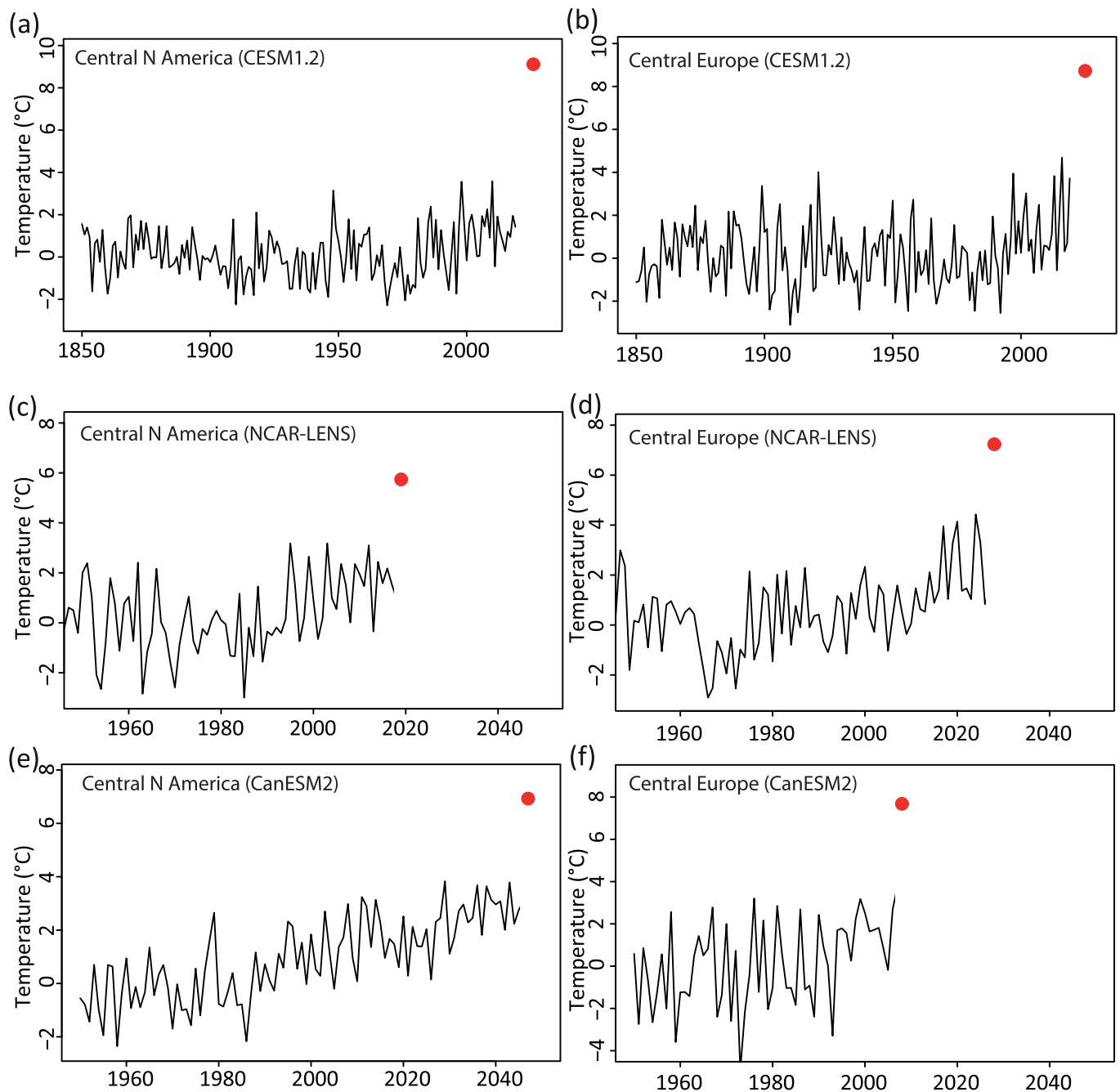
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Record-shattering event in long-term context. (top) Tx7day anomalies averaged across Central N America in all CESM1.2 members from 1850–2100. The red and yellow dots illustrate the two record-shattering events discussed in the main text. The red line and red dot highlight the event and corresponding simulations shown in Fig. 1a. The yellow line and yellow dot highlight the second most extreme record-shattering event. (bottom) The second most intense record-shattering event discussed in the main text is illustrated in the same way as the most extreme event in Fig. 1a. Best estimate of 1,000-yr return period estimated from a stationary GEV fit to the selected member (1850–2019) (dashed turquoise line) and from all 84 members (dotted violet line) (1950–2019) along with 95% CI.



Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Drivers of record-shattering extremes in long-term context. Same as Fig. 2d–g but relative to a constant reference period expressed as the multimember average across 1981–2010. The record-shattering extremes exceeding the previous record by more than 2σ , 3σ , and 4σ are marked with yellow, orange and red dots, respectively.



Extended Data Fig. 6 | Examples of record-shattering events in other large ensembles and other regions. Record-shattering event (red dot) and annual maximum 7-day temperature anomalies averaged over (a,c,e) Central N America) and (b,d,f) Central Europe in the corresponding simulation before the event (black line).