PNAS RESEARCH ARTICLE | EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES OPEN ACCESS

Global emergence of regional heatwave hotspots outpaces climate model simulations

Kai Kornhuber[a,b,c,1], Samuel Bartusek[b,d], Richard Seager[b], Hans Joachim Schellnhuber[a,1], and Mingfang Ting[b,c]

Affiliations are included on p. 9.

Contributed by Hans Joachim Schellnhuber; received June 5, 2024; accepted September 30, 2024; reviewed by Erich Fischer and Regina R. Rodrigues

Multiple recent record-shattering weather events raise questions about the adequacy of climate models to effectively predict and prepare for unprecedented climate impacts on human life, infrastructure, and ecosystems. Here, we show that extreme heat in several regions globally is increasing significantly and faster in magnitude than what state-of-the-art climate models have predicted under present warming even after accounting for their regional summer background warming. Across all global land area, models underestimate positive trends exceeding 0.5 °C per decade in widening of the upper tail of extreme surface temperature distributions by a factor of four compared to reanalysis data and exhibit a lower fraction of significantly increasing trends overall. To a lesser degree, models also underestimate observed strong trends of contraction of the upper tails in some areas, while moderate trends are well reproduced in a global perspective. Our results highlight the need to better understand and model the drivers of extreme heat and to rapidly mitigate greenhouse gas emissions to avoid further harm from unexpected weather events.

heatwaves | extreme weather | climate change

The frequency and magnitude of extreme weather events that exceeded the margins of local and regional climatology by multiple SD in recent years have caused substantial impacts on ecological and social systems and attracted much attention in the public and scientific domain. Early examples of such "record-shattering" heat events (1) include the European and Russian heatwaves of 2003, 2010, and 2018 (2–6) and the Siberian heatwave of 2020 (7, 8). More recently, the extreme 2021 heatwave struck the North American Pacific Northwest (9–11) (Fig. 1 *A* and *B*) and the sequential European heatwaves of 2022 (Fig. 1 *C* and *D*), which contributed to a seasonal total of over 60,000 heat related deaths (12), occurred in synchrony with record-breaking heatwaves in North America and China. Record-breaking heat returned to Europe, as well as central Russia and Southeast Asia (*SI Appendix*, Fig. S1 *A* and *B*) in 2023 (13, 14). The year 2023 also featured record-breaking heat in the Amazon basin (Fig. 1 *E* and *F*), which had substantial health impacts on the affected population. Notable past examples of record-breaking heat in the Southern Hemisphere further include the record-breaking heatwave of 2016 in South Africa (Fig. 1 *G* and *H*), which exacerbated water scarcity in that area and was part of a multiyear drought among many other events. As confirmed in multiple extreme event attribution studies, the globally observed long-term increase in such extreme heat events can be attributed to anthropogenic activities, the rise of greenhouse gas concentrations in the atmosphere, and their associated warming (15, 16).

In the global average, this warming has been accurately predicted by different generations of climate models (17). However, an overestimation of mean warming in some of the most recent CMIP6 models (commonly known as the "hot-model" problem) has sparked discussions about paying closer attention in multimodel contexts to weighting models by their skill in reproducing observed trends (18). From an extreme weather perspective, the large and unexpected margins by which recent regional-scale extremes have broken earlier records have raised questions about the degree to which climate models can provide adequate estimates of relations between global mean temperature changes and regional climate risks.

The interacting processes associated with extreme heat can cause the tails of the distributions to increase faster than the mean, with effects on regional temperature distributions (19, 20). As key heatwave drivers, soil moisture deficiencies, surface air temperature, and high-pressure systems constitute a tightly linked interacting trifold that can drive heat into extreme ranges. Bartusek et al. (9) exemplified the framework of climate-change-driven nonlinear interactions based on the case of the 2021 Pacific Northwest heatwave, showing that extreme anomalies of common drivers can push a linear dependence structure among the

Significance

Heatwaves can lead to considerable impacts on societal and natural systems. Accurate simulation of their response to warming is important for adaptation to potential climate futures. Here, we quantify changes of extreme temperatures worldwide over recent decades. We find an emergence of hotspots where the hottest temperatures are warming significantly faster than more moderate temperatures. In these regions, trends are largely underestimated in climate model simulations. Globally aggregated, we find that models struggle with both ends of the trend distribution, with positive trends being underestimated most, while moderate trends are well reproduced. Our findings highlight the need to better understand and model extreme heat and to rapidly mitigate greenhouse gas emissions to avoid further harm.

Author contributions: K.K., S.B., R.S., H.J.S., and M.T. designed research; K.K. and S.B. performed research; S.B. analyzed data; and K.K., S.B., R.S., H.J.S., and M.T. wrote the paper.

Reviewers: E.F., Eidgenossische Technische Hochschule Zurich; and R.R.R., Universidade Federal de Santa Catarina.

The authors declare no competing interest.

[1] To whom correspondence may be addressed. Email: kornhuber@iiasa.ac.at or schellnhuber@iiasa.ac.at.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2411258121/-/DCSupplemental.

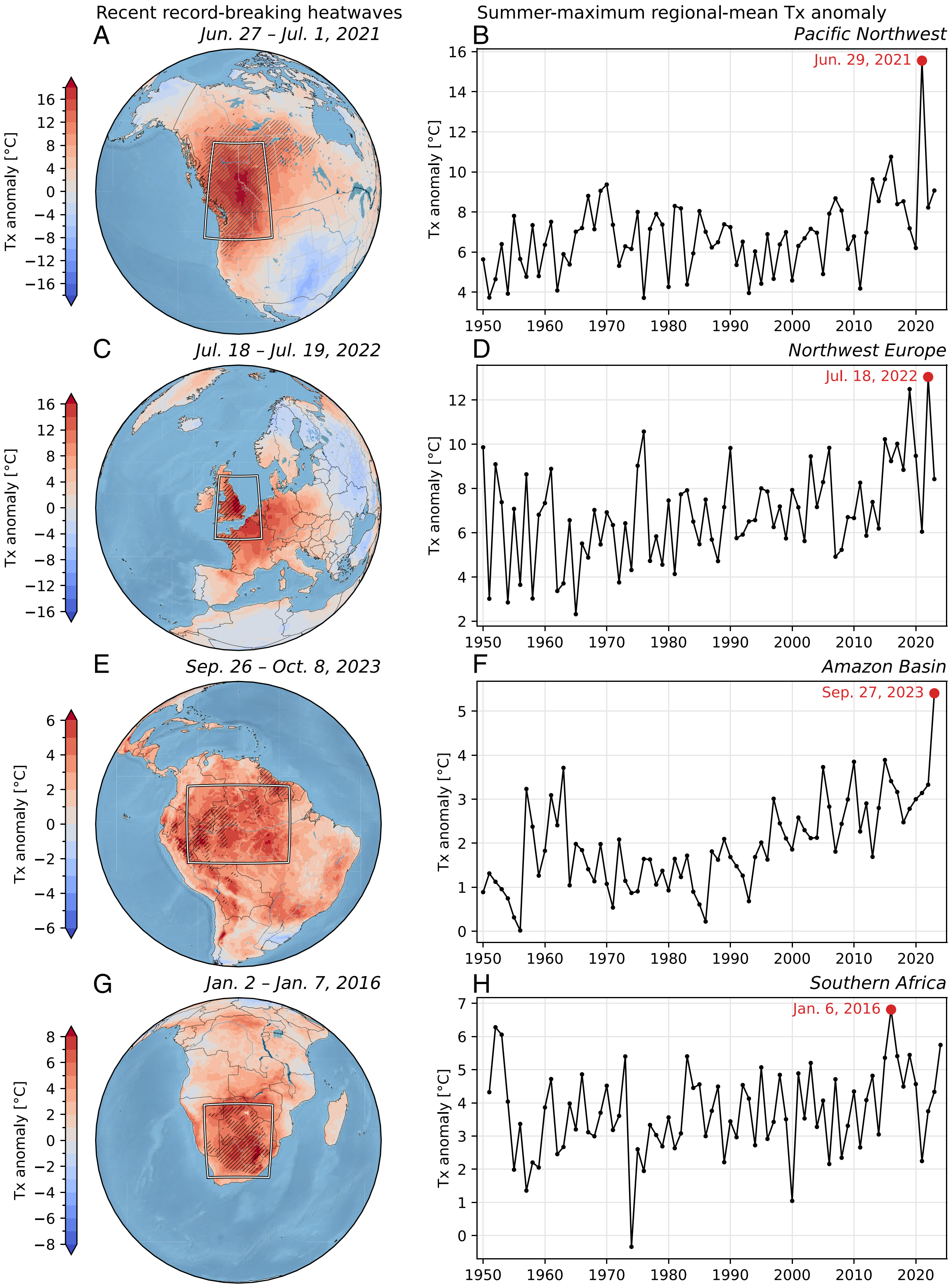

Fig. 1. Daily maximum temperature anomalies during recent record-breaking heatwaves and their temporal context. (*A*) 2-m daily maximum temperature (Tx) anomaly fields averaged over the 2021 Pacific Northwest heatwave in North America. Regions where values were record-breaking (since 1950) during the indicated time period are hatched. (*B*) Time series for the years 1950 to 2023 of the hottest annual average Tx anomaly over the region indicated by the box in *A* (relative to 1981 to 2010 June–August). The record-breaking values of regional-mean Tx and their dates are highlighted (red dot) in each time series. (*C* and *D*) same as *A* and *B* but for the western European heatwave in July 2022, (*E* and *F*) for the heatwave in the Amazon Basin in November 2023 (warm season September–November) and (*G* and *H*) the heatwave in Southern Africa (warm season December–February) in January 2016.

three variables into a nonlinear regime. These relationships are complex and can still pose challenges to models on a regional-scale. Throughout multiple generations of Coupled Model Intercomparison Projects, models have predicted that daily surface temperature variability during the warm season should increase over a large majority of land areas in response to global warming (21, 22), in contrast to strongly decreasing variability outside the warm season, especially at high latitudes (6, 23–25).

However, comparing observed trends in the upper tail of temperature distributions to modeled trends over the historical period remains relatively unexplored (20, 21, 23, 26, 27).

Here, we investigate trends in extreme temperatures by focusing on the warming of the extreme tails compared to trends in more moderate percentiles of the yearly temperature distribution for each grid-point. The manuscript is structured as follows: We first assess trends in the upper tail of local yearly distributions in daily maximum temperatures compared to the warming of typical warm season days over the past seven decades to identify regional heatwave hotspots, using a range of available reanalysis products (Fig. 2). We then compare the observed trends with those based on a large number of state-of-the-art model experiments from the HighResMIP project (28) consisting of a range of model resolutions and setups, fully coupled and forced with observed sea surface temperatures (SST) (Fig. 3) and then discuss the local and global discrepancies (Fig. 4). We conclude with a discussion of physical mechanisms that might not be accurately captured in models and suggestions on how to move forward.

Results

Trends in the most extreme daily maximum near-surface temperatures (Tx) values per year are greater than what is expected from a simple shift in the mean (i.e., a uniform shift across all percentiles) in several highly populated regions globally. Fig. 2*A* displays observed regional changes in the most extreme Tx values (99th percentile) per year compared to changes in less extreme (87.5th percentile, as the median of the 75th and 100th percentile, the limits of the upper quartile) Tx values, based on ERA5 reanalysis data from 1958 to 2022 (29). For simplicity, we will refer to this difference in trends as "tail-widening" trends in the following. The most intense signal is observed in Western Europe (Fig. 2*B*, box b in Fig. 2*A*), as has been reported elsewhere (30–32). This is a robust signal found in a series of reanalysis (JRA-55) and an observation-based dataset for Europe (E-OBS). Our analysis reveals a set of additional regions globally in which the most extreme temperatures within a year are rising significantly faster than the 87.5th percentile (nonstippled regions in Fig. 2 *A*–*I*), as selected by objective conditions (provided in the methods section. Among those are central China, southern South America, the Arabian Peninsula, Eastern Australia, Japan, and Korea, and high-latitude regions of Canada and Greenland. In the highest-latitude regions, JRA-55 trends are weaker, though still toward the extreme end of the modeled trend distributions. Here, reanalysis products disagree on the strength of the trends, possibly due to more challenging data aggregation over these regions and the fact that JRA-55 Tx is calculated from six-hourly instead of hourly data. Tail-widening is, however, by no means ubiquitous and, in fact, tail-narrowing occurs over vast regions of North America and Siberia and other regions of the world. Fig. 2 *J* and *K* highlight two of these regions. Further *SI Appendix*, Fig. S2 and the methods section detail the objective selection criteria which motivated the eight regions shown in Fig. 2. These conditions require robust trends across time periods and reanalysis products. Analogous assessments over additional regions that do not meet all the criteria are provided in *SI Appendix*, Figs. S3 and S4.

The observed occurrence of hotspots is largely, if not completely, missed over the same regions by state-of-the-art modeling frameworks, indicated with boxes in Fig. 3*A*. In these regions, the observed trends are in the extreme end or even outside of the modeled spread, even when using global mean temperature as covariate instead of time (*SI Appendix*, Figs. S6, S7*B*, and S9). Because these observed trends in extreme temperatures are only compared with modeled trends after removing the average summer warming, model-observation discrepancy resulting from seasonal-scale warming differences is minimized. Thus, comparing the tail-widening in reanalysis data and models is a more specific comparison that seeks to minimize the trend influence of changes in lower percentiles, while investigating trends in Tx alone might include a large role from trends in the 50th percentile.

Regions such as Northwest Europe (Figs. 1 *C* and *D* and 2*B*) among other areas have repeatedly witnessed record-breaking extreme heat events in recent years and the upper tail of the nearby temperature distribution has steadily been widening (Fig. 2). These large and in part densely populated regions are among those for which we find that tail-widening in observations exceed the 95th percentile of all the model spread (Fig. 3 *B*–*I*). Other notable areas where grid-point discrepancies persist but regional averages do not meet all the criteria outlined in *SI Appendix*, Fig. S2 include densely populated areas of the Southern United States and important biomes such as the Amazon and central Africa (*SI Appendix*, Figs. S1–S4).

In the case of Southern South America (Fig. 3*D*), the Arabian Peninsula (Fig. 3*E*), and the high Arctic (Fig. 3*H*) observed trends in ERA5 are stronger than in any model realization. While reanalysis datasets (ERA5, E-OBS, JRA-55) mostly agree on the sign of trends, it is notable that their magnitude can differ strongly in some regions. JRA-55 shows smaller trends throughout the regions outlined in Fig. 3. In some cases, such as over the Arctic (Fig. 3*H*) or over Northwest Canada (Fig. 3*I*), these discrepancies can be quite large, to a degree that JRA-55 is more within the model ensemble spread to a degree that the CI can overlap. As ERA5 can be considered the more modern and accurate reanalysis dataset due to a higher spatial resolution in vertical and horizontal directions and increased temporal resolution (hourly vs. six-hourly) (29, 33), these discrepancies could result in difficulties simulating the processes in extreme heat in JRA-55. Interestingly, models do underestimate negative trends in tail-widening (i.e., tail-narrowing) over some areas as well. Northern Africa (Fig. 2*J*) and Siberia (Fig. 2*K*) show the largest areas of significant negative trends in observations, a signal that is partially missed in model simulations (Fig. 3 *J* and *K*), in particular over Siberia (Fig. 3*K*).

While the differences between observed and modeled trends in the Northwestern Europe hotspot have been investigated (30–32, 34), here we expand such an analysis to the remainder of the global land area. Additionally, in contrast to previous studies, our analysis also includes SST-forced models which, unlike the coupled models, contain the observed history of SSTs (see *SI Appendix*, Table S1 for a list of all models investigated). We find that model biases are largely independent of the type of model set up, locally (Fig. 3 *B*–*I*) and on a global scale: The land area fraction over which trends are misaligned is only slightly reduced in high-resolution (25 to 50 km) (*SI Appendix*, Figs. S8 *A*–*C* and S9 *A*–*C*) and SST-forced models (*SI Appendix*, Figs. S8 *D*–*F* and S9 *D*–*F*). In experiments where the atmosphere is forced with observed SSTs, a slight improvement is found for median values, which slightly move toward higher magnitude, in particular for regions such as eastern Australia and Japan/Korea (*SI Appendix*, Fig. S5 *F* and *G*), where median trends switch from negative trends in coupled models to positive trends in SST-forced models. This suggests a minor role for SSTs potentially through SST-forced teleconnections in these regions that could have contributed to recent trends. The reanalysis-based trends are at the outer margins or beyond the 95th percentile of the model spread, which also holds for larger SST-forced ensembles (up to a total of 109 model runs) which do not offer a substantial improvement (*SI Appendix*, Fig. S5).

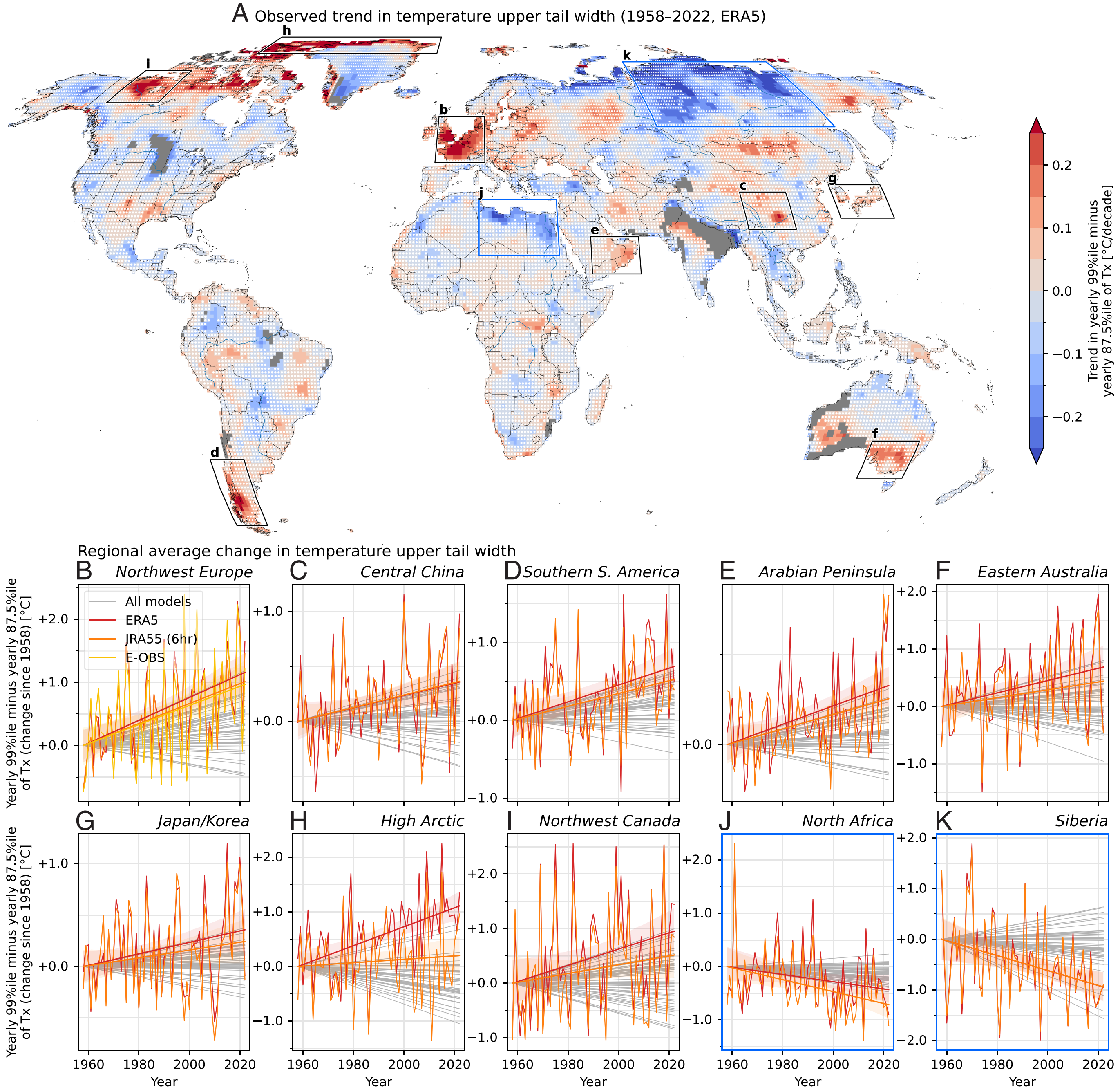

Fig. 2. A global emergence of regional heatwave hotspots. (*A*) Regional trends in extreme heat tail-widening, estimated by calculating long-term trends in the differences of the yearly 99th percentile of daily maximum temperature (Tx) and the yearly 87.5th percentile of Tx (the median of the upper quartile bounded by the 75th and 100th percentiles) at each grid-point over 1958 to 2022. Areas where trends in the annual 87.5th percentile of Tx are negative are shown in gray. Grid-points where the bootstrapped 95% CI includes zero (*Data & Methods*) are masked with white dots. A warming of the most extreme events exceeding the underlying upper quartile median warming (i.e., a widening of the upper tail of the temperature distribution) is observed in various regions globally. (*B–I*) Timeseries and linear trends of regionally aggregated changes for areas where trends have been increasing as highlighted by black boxes in a and using ERA5 (red), JRA-55 (orange), and E-OBS (yellow, for Europe only) (see labels in *A*). Gray lines show the trends retrieved from a suite of climate models, which largely fail to reproduce observed trends (see Fig. 3 for further details). Bootstrapped 95% CI of the linear trends are shaded. (*J* and *K*) shows the same analysis for regions in which the hottest Tx values have increased at a slower rate compared to the more typical warm season hot days (blue boxes in *A*).

Finally, we find that discrepancies of strongly positive trends in the upper tails of surface temperature distributions are notable also when aggregated globally or assessed over specific latitudinal ranges (Fig. 4 and *SI Appendix*, Figs. S10 and S11). While climate models exhibit a higher fraction of land area overall with positive tail width trends than in the observations (55% in models, 48% in ERA5), they simulate a much smaller area of statistically significant positive trends than seen in observations (*SI Appendix*, Fig. S10 *A*, *B*, and *D*). At $P < 0.05$, 16.3% of land-weighted positive trends are significant in ERA5, vs. 10.5% in models, and this discrepancy intensifies at even higher significance levels (7.9% vs. 3.5% at $P < 0.01$) (*SI Appendix*, Fig. S10 *C* and *E*). Probability density histograms show that for moderate trends, models largely reproduce those found in reanalysis (Fig. 4*A*), while still underestimating the extreme trends, as indicated by the higher skewness of the observed distribution (Fig. 4*A*, note the log-scale at the y-axis). Interestingly, the distribution of low-resolution models matches the reanalysis better (skewness of 0.24 and 0.4, respectively) compared to high-resolution models

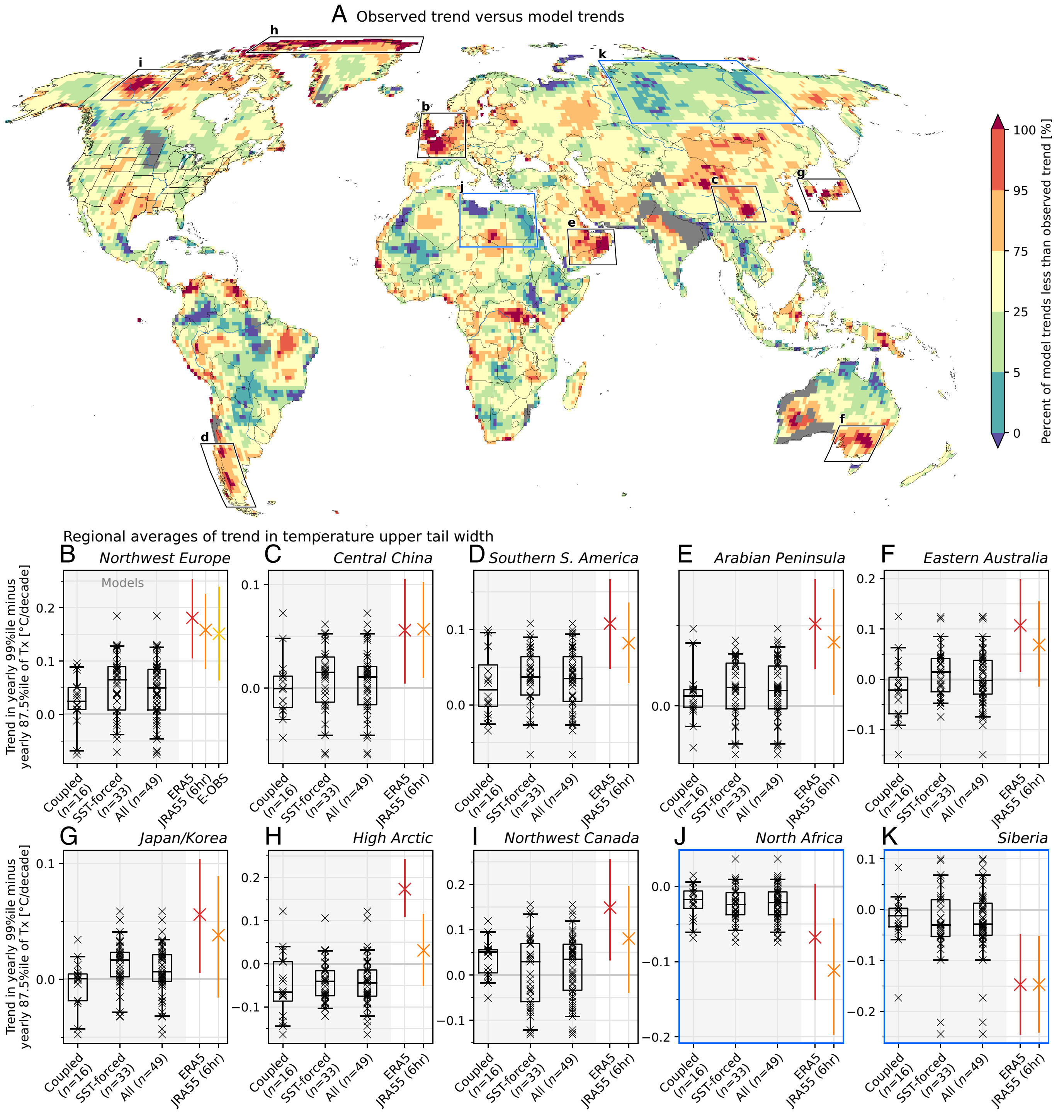

Fig. 3. Regional trends of extreme temperatures are underestimated in climate model experiments in multiple regions globally. (*A*) Comparison of observed trends in tail-widening (yearly 99th percentile minus 87.5th percentile) with 49 simulations from coupled and SST-forced climate models (*SI Appendix*, Table S1). Observed trends are outside of the modeled range in several regions globally (dark red). Areas where the annual 87.5th percentile of Tx shows a negative trend in observations are shown in gray. (*B–I*) Distributions of modeled trends in the 99th percentile relative to the 87.5th percentile each year in areas as in Fig. 2 *B–I* for coupled and SST-forced models compared to ERA5 (red), JRA-55 (orange), and E-OBS (in *B* only) observations-based gridded climate data, displayed as box-and whisker-plots. Boxes span the 25th and 75th percentiles of the model spread while the median is shown as a horizontal black line. The whiskers denote the 5th and 95th percentiles, while the single model values are provided as scatters. In the same fashion, panels *J* and *K* show exemplary regions in which models exhibit an overestimation of historic trends in tail width. Bootstrapped 95% CI bounds are shown for observations (*Data & Methods*).

(skewness of 0.15), while all models combined exhibit a skewness of 0.2 (*SI Appendix*, Fig. S11*A*) underestimating the tail. Low-resolution models also exhibit the strongest single grid-point trends, which may be attributable to isolated grid-points in Arctic regions, in the vicinity of coastlines and potential artifacts of challenging cryosphere modeling (see exemplary trend map based on one ensemble member in *SI Appendix*, Fig. S12). The underestimation of positive trends is also expressed by the cumulative probability ratio between ERA5 and models (Fig. 4*B*).

We find that for positive trends, models underestimate the land area exhibiting an occurrence of positive tail-widening trends exceeding 0.5 °C per decade by a factor of four. Considering trends of both signs rather than just positive trends, models underestimate the occurrence of trends exceeding 0.5 °C per decade by a

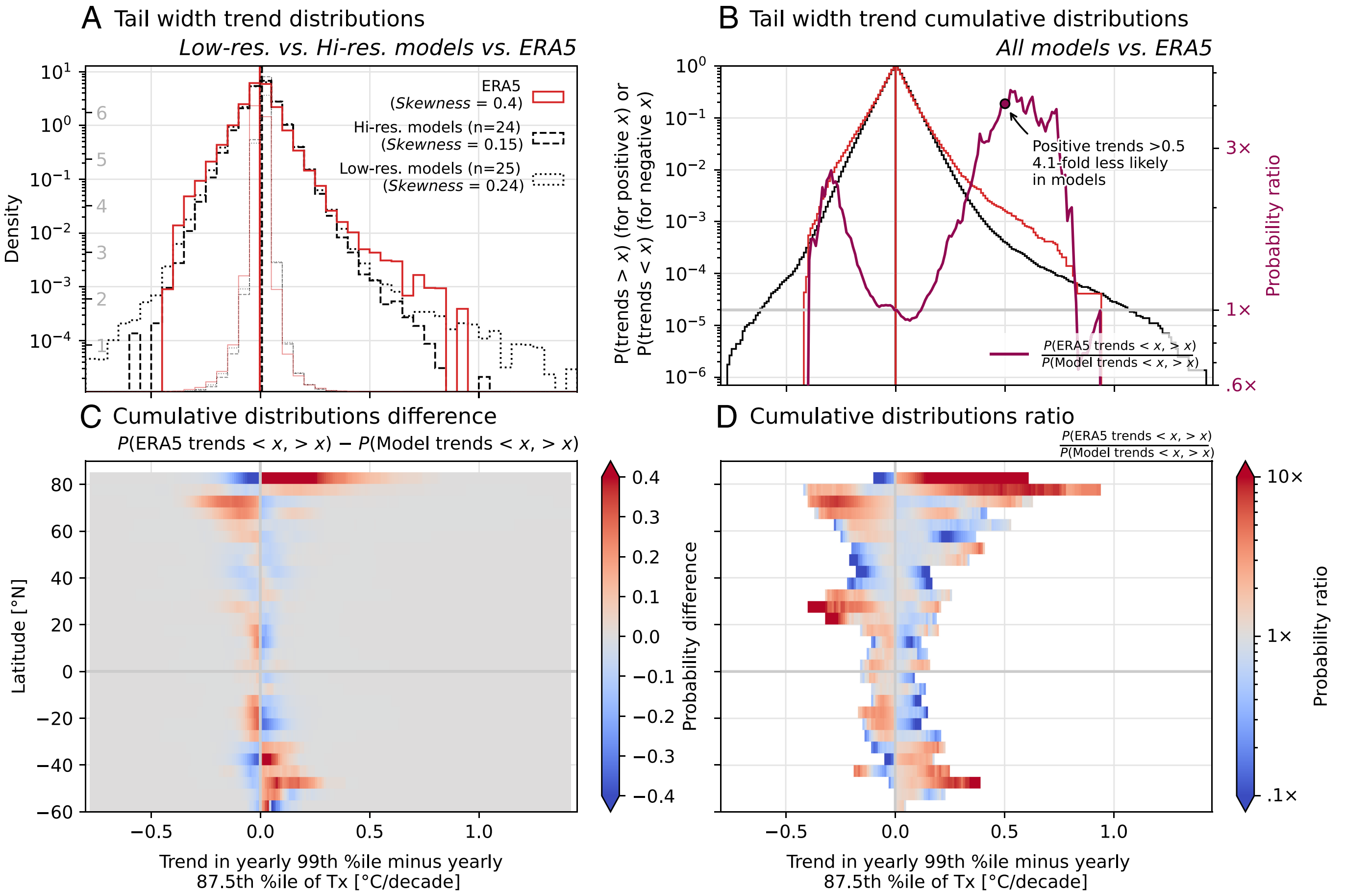

Fig. 4. Global underestimation of trends in extreme temperature tail width in climate models compared to reanalysis data. (*A*) Area weighted probability density histogram of trends over land in the difference between the yearly 99th percentile and 87.5th percentile Tx in ERA5 (red), high-resolution climate models (black, dashed) and low-resolution climate models (black, dotted). Number of models included and the skewness of the distribution are noted in the legend. Histograms are shown according to a log scale (saturated colors) and a linear scale (translucent colors). Medians of the distributions are shown as vertical lines. (*B*) Cumulative probability distributions of negative trends (left of zero) and positive trends (right of zero) separately, and their ratios between models and ERA5. Models underestimate positive trends exceeding 0.5 °C/decade by a factor of four. (*C*) Differences and (*D*) ratios between cumulative distributions of positive and negative trends in reanalysis and models, by latitude. The largest discrepancies are identified in Northern Hemisphere high latitudes and the mid-latitudes of the Northern and Southern Hemispheres.

factor of 3.5 (*SI Appendix*, Fig. S11*B*). Note that despite that discrepancy in strong trends (here chosen as above 0.5 °C per decade), they simulate positive trends slightly more often overall. Biases are found to be strongest in the Northern high latitudes, in both cumulative probability differences (Fig. 4*C*) and ratios (Fig. 4*D*), while Northern and Southern Hemisphere mid-latitudes, where multiple hotspots are located, are also emphasized. Red areas on the positive side at around 50 to 60°N and 50°S are likely the reflection of the observed hotspots of Northwest Europe and Southern South America. For lower mid-latitudes (around 40°N), an overestimation of moderate (~0.2 °C per decade) positive and negative trends is identified, which are mostly nonsignificant (Fig. 1*A*). For the case of high northern latitudes, the underestimated Arctic mean warming in climate models reported previously (35) might also affect the tails of the distribution.

Summary and Discussion

While coupled climate models have been useful tools in modeling and projecting the past global mean temperature response to anthropogenic activities over the historical period (17) (*SI Appendix*, Fig. S6), we find that observed long-term trends in the tail behavior of extreme heat events are indeed outside of what historical model ensembles suggest in several regions globally (Fig. 3 and *SI Appendix*, Figs. S5, S8, and S9). High-impact extreme weather events are mostly without exception the outcome of several compounding factors acting together, with regionally varying importance of each component involved.

Dry soils and associated land–atmosphere feedbacks are major heatwave drivers (9, 19, 36). It has been found that an amplified warming trend of hot days vs. mean warming in the tropics can largely be explained by a "dry gets hotter" mechanism (37), while precipitation trends were found to govern the occurrence of hot-dry extremes globally (38). Huntingford et al. (26) found regionally varying causes for trends in the yearly 90th percentile of daily temperatures: While in the Northern Hemisphere, extratropics dry soils were emphasized, trends in tropical Africa were linked to increased available energy. Similar results were reported by ref. 27 who found a good agreement of model simulations and reanalysis data over the Northern Hemisphere in a more moderate definition of tail-widening (trends average Tx vs hottest Tx). They also find that positive and negative trends (as observed here over Siberia or Northern Africa) can be related to local changes in the hydroclimate. This is consistent with earlier findings (39) that attributed amplified warming in Mediterranean-type regions. Simpson et al. (40) found that trends in humidity, which are strongly dependent on the accurate depiction of rainfall patterns (38), evaporation (which is partially controlled by vegetation), and hydrological characteristics of the land surface, including vegetation are still not accurately reproduced, which could in part explain the discrepancies reported here.

Persistent high-pressure systems, which materialize as local blocking patterns (41) or zonally elongated stationary Rossby waves (3, 42, 43) are important contributors to weather extremes especially in the mid-latitudes (44). Atmospheric circulation is considered a major source of uncertainty which also affects precipitation trends (Shepherd, 2014) and, hence, surface water and energy balances and temperatures. Specifically, Europe has been identified as a global heatwave hotspot (31), where the hottest days of the year are warming twice as fast as mean summer days (30), a trend that is driven by atmosphere dynamical patterns (31, 34) and is largely missed by climate models (32). Circulation trends could also be contributing to the decreasing tail width over Siberia and Northern Africa where geopotential height trends show smaller or negative trends compared to the remaining mid-latitudes (45) and have suppressed warming over the central United States (masked in gray in Figs. 1 and 2 as 87th percentiles show a negative trend) (46).

Although the newest generation of climate models shows some improvement in the representation of the frequency and magnitude of atmospheric blocking (41, 47, 48), these measures are still underestimated in CMIP6 models (49). Recent research has shown that while models do accurately reproduce the location and strength of upper-level wave patterns, they also substantially underestimate the surface response to quasistationary wave patterns of the type involved in several of recent extreme weather events, e.g., the European heatwave of 2003 (50, 51).

The slight improvement we find in SST-forced models, in particular for Eastern Australia and Japan and Korea (Fig. 3 and *SI Appendix*, Fig. S5), provides further evidence for a potential role for specific SST patterns, possibly by forcing certain atmospheric dynamical circulation patterns and/or rainfall patterns and associated land–atmosphere feedbacks, which have played an important role in recent high-impact heatwaves (3, 5, 52, 53). Heatwaves in North America are often linked to persistent ridges in the jet stream, which have been related to SST patterns in the Pacific (54, 55).

Persistent and extreme heat has particularly increased over western and southern North America (31, 56). Tropical Pacific SSTs exert a powerful control on climate and weather variability worldwide, primarily via El Niño-Southern Oscillation (ENSO) cycles. El Niño has been suggested as a potential contributor to some extreme heat and precipitation in the Northern mid-latitudes in summer 2023 (13). Further, it is known that a La Niña-like SST trend in the tropical Pacific has contributed to the two-decades-long megadrought in southwest North America (57). ENSO events also have an important role in favoring specific heat extremes such as the 2010 heatwave in Russia which was associated with a La Niña-like SST pattern (52). State-of-the-art climate models predict that rising greenhouse gas emissions (GHG) should reduce the west-to-east warm-to-cool SST gradient across the equatorial Pacific while, in observations, the gradient has strengthened over recent decades along with rising GHG concentrations (58, 59). Regional biases in heatwave intensification may therefore be partially linked to diverging SST signals in models and observations and how they teleconnect to precipitation and temperature worldwide. However, since bias reduction in SST-forced experiments is small, this cannot be the sole explanation (Fig. 3 and *SI Appendix*, Fig. S5).

The representation of aerosols and their interaction with clouds remain a major challenge for climate models (60), but these factors can play an important role in regional heatwave trends (61). Aerosol reduction has been identified as a contributing driver of European heatwave trends, largely missed by regional models (62). China has substantially reduced aerosol and ozone precursor emissions in recent years, which have contributed to increased local temperature trends in some locations (63).

Conclusion

Actionable climate assessment for effective climate adaptation and mitigation requires skillful and reliable projections of extreme weather risks under different emission scenarios on a regional to local level. This holds particularly true for the representation of recently observed extremes of large magnitude that might be rare under current climatic conditions but will become more likely under continued GHG emissions (1, 56, 64). Skillful projections of trends in such "extreme-extremes" (unprecedented or record-shattering extremes) must build on a thorough physical understanding of why they are emerging and the nonlinear behavior responsible so that model simulations can be benchmarked and potential biases can be accounted for.

In large and densely populated areas such as western Europe and China and other areas that feature important biomes for the world climate such as the Amazon, and polar regions around Greenland and Canada, some of which have been discussed in the context of climate tipping points (65, 66), the multimodel mean of climate simulations of the past decades does not show the enhanced warming of the temperature distributions' upper tails observed in these regions (Fig. 1 and *SI Appendix*, Fig. S5). Note that for the Amazon, the strongest trends have emerged over the past 23 y and are found for ERA5 only (*SI Appendix*, Fig. S4).

Often, the multimodel mean is used and prioritized in many assessments of climate risks, while upper percentiles are treated as implausible scenarios and are at times rejected as outliers. For instance, the 1.5 °C warming target established by the Paris Agreement was set largely based on avoiding "dangerous climate change," in part associated with critical tipping elements and/or thresholds in the Earth system (65, 67). However, if impacts of global warming, such as amplified extreme heat, proceed faster than expected based on the multimodel mean projections used to support such a warming target, its utility may deserve reconsideration. We find that in numerous regions (Figs. 2 and 3), trends in the tail-widening of extreme heat distribution over the past 65 y exceed the 95th percentile of the model spread and, in some cases, even exceed the spread entirely. Trends shown in ERA5 reanalysis are outside of the modeled range for southern South America, the Arabian Peninsula, and Arctic Canada (Fig. 3 *D*, *E*, and *H*), irrespective of any model configuration investigated here, while the observed uncertainty intervals determined by bootstrapping overlap with the model spread. These findings hold for model simulations at higher resolution, or forced with historical SSTs, as well as with greatly expanded ensemble sizes (*SI Appendix*, Figs. S5, S8, and S9).

Newer modeling initiatives such as super-high-resolution frameworks suggested, e.g., in the Earth Virtualization Engine (EVE) (68) promise convection permitting resolution and may offer possibilities in improving the depiction of important mechanisms. However, no substantial improvement for the higher resolved subset of the investigated models was found. Super-high-resolution, convection-resolving models may better represent processes that link SSTs with Rossby waves and associated extremes (45), regional blocking, and realistic surface response of heat events to such atmospheric patterns (50, 51). However, limitations due to data storage and computing costs might be significant constraints for the study of extreme events with high-resolution modeling frameworks, as the long time series lengths and large ensemble sizes needed for adequate statistics and trend attribution may be too resource intensive and not readily available. Newer generation models have also shown an improved skill in modeling blocking events which is more pronounced in high-resolution models (47, 69). Given the importance of nonlinear feedbacks involving hydroclimatic processes, a proper representation of the seasonal relationships of the flow of energy and water in the soil–vegetation–atmosphere continuum needs to be assured (7). Reasonable forecasts of past extreme heatwaves suggest that models can in principle produce such extreme-extremes when directly forced with the correct boundary conditions (11, 70). Ensemble boosting techniques can be used to create large ensembles of extraordinary extremes at reduced computational cost (71, 72). In an evolutionary manner, these algorithms preserve those that follow an extreme trajectory while filtering out others. This allows a sampling around a specific event characteristic. A large ensemble of highly anomalous events, which would be featured only at an extremely low rate in large ensembles (20), allows for an in-depth and statistically robust analysis of the governing physics of extreme-extremes in models.

However, disentangling the relative importance of externally forced and internal variability in the observed trends may be key to attributing the sources of model–observation discrepancies. Coordinated single forcing large ensemble experiments such as the new Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP) (73) might help in improving our understanding in the relative role of various external or internal drivers in extreme event trends.

Further, machine learning (ML) approaches have shown promising results for providing more reliable bias adjustment of climate model output. These are based on methods from image processing and are better in retaining the relationships between variables compared to more traditional quantile-mapping approaches. This is particularly important when analyzing risks and impacts from compound extremes. ML techniques could also assist in detecting nonlinear and regime-changing behavior in the ocean–atmosphere–land–vegetation system and provide causality where common drivers experience strong coupling and feedbacks (9, 74). Beyond using ML for analysis, recent advances in ML-driven weather forecasts exemplify its potential in climate modeling (75, 76). In addition, ML might offer accurate and less computationally costly solutions for resolving important subgrid processes (77, 78), compared to purely numerical frameworks. ML approaches, however, must be combined with others that can physically explain and understand the causal flows identified by ML. New assimilation techniques that integrate observational datasets and exploit advanced interpolation frameworks have been proven to improve the depiction of extremes compared to reanalysis datasets (79) and provide climate information at a higher resolution.

While our findings provide many avenues for interesting and relevant new research, the authors stress that the best way to reduce both uncertainty in and exposure to climate impacts is a rapid transition of relevant societal sectors away from fossil fuels to stabilize global temperature rise.

Data & Methods

Data. The analysis is based on daily maximum temperature (Tx) at 2-m height. All model-derived results in the main analysis use data from the HighResMIP project (28), which provides a good balance and coherent setup of coupled and SST-forced experiments. Within HighResMIP, configurations labeled "SST-forced" in *SI Appendix*, Table S1 refer to a concatenation of the "highresSST-present" experiment from 1958 to 2014 with the "highresSST-future" experiment from 2015 to 2022 (with matching member IDs only). Both are atmosphere-only, with highresSST-present forced by historic SST/sea-ice fields and highresSST-future forced by SSP585 SST/sea-ice fields.

Configurations labeled "Coupled" in the table above refer to a concatenation of the "hist-1950" experiment from 1950 to 2014 with the "highres-future" experiment from 2015 to 2022 (with matching member IDs only). Both are coupled, with hist-1950 subject to historical forcing and highres-future subject to SSP585 forcing. All model data were pre-processed with *xmip* to standardize metadata and data structures.

ERA5 reanalysis data (29) from years 1958 to 2023 were used in Fig. 1 to display 2-m Tx and ERA5 u and v components of wind at the 300 hPa pressure level and were downloaded from the Copernicus Data Store (https://cds.climate.copernicus.eu/#!/home). For Fig. 2, the data were limited to 1958 to 2022 as a global analysis requires the availability of the entire annual dataset.

Six-hourly temperature data from the Japanese 55-year Reanalysis (JRA-55) (80) ranging from years 1958 to 2022 were used in Figs. 2 and 3 and were retrieved from https://jra.kishou.go.jp/JRA-55/index_en.html. Note that the difference in temporal resolution, might influence the accuracy of Tx values based on JRA-55 (six-hourly) compared to ERA5 (hourly).

Temperature data from the gridded observation-based E-OBS (81) were used for calculating trends over Europe as shown in Figs. 2*B* and 3*B* and were downloaded from: https://www.ecad.eu/download/ensembles/download.php.

Due to a better representation of various fields possibly linked to a higher horizontal and vertical spatial and also temporal resolution which makes the quantification of Tx values more accurate ERA5 is considered as the best estimate of historical trends in this analysis.

Methods. Tx is defined as the hottest daily temperature based on one-hourly data for ERA5 and six-hourly data for JRA-55. In the main analysis, the yearly 99th and 87.5th percentiles (as the median of the upper quartile, spanning the 75th to 100th percentile) of Tx were calculated for each grid-point from model runs (on their native grids) and ERA5, for all years 1958 to 2022. The 99th percentile represents the median day of the hottest 2% days of the year, while the 87.5th percentile approximates the average summer day, as it represents the median of the hottest quarter of days of the year. This percentile approach is therefore generalizable to all areas of the globe, not dependent on specific calendar definitions of seasons. Computed percentiles were conservatively regridded to a common 1-degree grid using *xesmf*. We compute linear trends in the yearly difference of these two percentiles, similar to a quantile regression approach (82, 83). In Figs. 2 and 3, significance and uncertainty ranges for observed trends are calculated via a bootstrap approach. In Fig. 2*A*, for each grid-point, the 65 yearly datapoints (the difference between yearly 99th and 87.5th percentile Tx) from 1958 to 2022 were resampled 1,000 times (with replacement, and without shuffling in time), calculating a linear trend in time for each iteration, and grid-points were covered with a white dot if the 2.5th to 97.5th percentile range of these trend values crossed zero. In Fig. 3 *B–K*, for each region and observational dataset, a yearly mean of the difference between the yearly 99th and 87.5th percentile Tx at each grid-point was calculated, the resulting 65 yearly datapoints were resampled 10,000 times as described above, and the 2.5th to 97.5th percentile range in the iterations' trend values is shown as vertical lines.

Regions shown in Figs. 2 and 3 were selected based on a set of conditions outlined in *SI Appendix*, Fig. S2, requiring that the trend of the regional average is positive and significant ($P < 0.05$) in ERA5 over the period 1958 to 2022, positive in JRA55 over the period 1958 to 2022 and positive over the period 1980 to 2022 in all available datasets [including MERRA-2 (84)]. Those regions are marked black in *SI Appendix*, Fig. S1, and regional average trends are shown in *SI Appendix*, Figs. S3 and S4. In *SI Appendix*, Fig. S10 (and where it is referenced in the main text), *P*-values for both ERA5 and model data are calculated parametrically, via a two-tailed Wald Test with a t-distribution of the test statistic, rather than applying the full iterative bootstrapping approach to all model ensemble members (in contrast to the significance testing in Figs. 2 and 3).

Data, Materials, and Software Availability. All figures were produced using Python v.3.6 (https://www.python.org/downloads/release/python-360/) (85). All code needed to reproduce the main figures are available at https://doi.org/10.5281/zenodo.13993309 (86). All model data were accessed through Earth System Grid Federation (ESGF) (https://wcrp-cmip.org/nodes/aims3-llnl-gov/)

(87). All model runs in main analysis are from CMIP6's HighResMIP (https://gmd.copernicus.org/articles/9/4185/2016/gmd-9-4185-2016.html) (88). All model runs listed in *SI Appendix*, Table S1 are available from ESGF (hosted at the Lawrence Livermore National Laboratory data node) at the time of accessing, given matching member IDs across historical and future experiments, and with uninterrupted data across the entire 1958 to 2023 period, were used. ERA5 data was downloaded from the Copernicus Climate Data Store.

ACKNOWLEDGMENTS. We want to thank the European Centre for Medium-Range Weather Forecasts and the Copernicus Climate Change Service for providing the ERA5 reanalysis and World Climate Research Programme, which, through its Working Group on Coupled Modeling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the ESGF for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. K.K. was supported by the German Federal Ministry of Education and Research within the ClimXtreme project subprojects PERSEVERE (01LP2322D). R.S. and M.T. were supported by NOAA award NA20OAR4310379 and R.S. additionally by NSF award AGS-2127684. S.B. was supported by the NASA under the Future Investigators in NASA Earth and Space Science and Technology program, grant 21-EARTH21-0134.

Author affiliations: [a]International Institute for Applied Systems Analysis, Laxenburg 2361, Austria; [b]Lamont-Doherty Earth Observatory, Columbia University, New York, NY 10964; [c]Columbia Climate School, Columbia University, New York, NY 10025; and [d]Department of Earth and Environmental Sciences, Columbia University, New York, NY 10027

1. E. M. Fischer, S. Sippel, R. Knutti, Increasing probability of record-shattering climate extremes. *Nat. Clim. Change* **11**, 689–695 (2021).
2. D. Barriopedro, E. M. Fischer, J. Luterbacher, R. M. Trigo, R. García-Herrera, The hot summer of 2010: Redrawing the temperature record map of Europe. *Science* **332**, 220–224 (2011).
3. K. Kornhuber *et al.*, Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environ. Res. Lett.* **14**, 054002 (2019).
4. D. Mitchell, K. Kornhuber, C. Huntingford, P. Uhe, The day the 2003 European heatwave record was broken. *Lancet Planet. Health* **3**, e290–e292 (2019).
5. E. Rousi *et al.*, The extremely hot and dry 2018 summer in central and northern Europe from a multi-faceted weather and climate perspective. *Nat. Hazards Earth Syst. Sci.* **23**, 1699–1718 (2023).
6. C. Schär *et al.*, The role of increasing temperature variability in European summer heatwaves. *Nature* **427**, 332–336 (2004).
7. L. Gloege *et al.*, Land-atmosphere cascade fueled the 2020 Siberian heatwave. *AGU Adv.* **3**, e2021AV000619 (2022).
8. J. E. Overland, M. Wang, The 2020 Siberian heat wave. *Int. J. Climatol.* **41**, E2341–E2346 (2021).
9. S. Bartusek, K. Kornhuber, M. Ting, 2021 North American heatwave amplified by climate change-driven nonlinear interactions. *Nat. Clim. Change* **12**, 1143–1150 (2022).
10. V. Thompson *et al.*, The 2021 western North America heat wave among the most extreme events ever recorded globally. *Sci. Adv.* **8**, eabm6860 (2022).
11. R. H. White *et al.*, The unprecedented Pacific Northwest heatwave of June 2021. *Nat. Commun.* **14**, 727 (2023).
12. J. Ballester *et al.*, Heat-related mortality in Europe during the summer of 2022. *Nat. Med.* **29**, 1857–1866 (2023).
13. S. Perkins-Kirkpatrick *et al.*, Extreme terrestrial heat in 2023. *Nat. Rev. Earth Environ.* **5**, 244–246 (2024).
14. M. Zachariah *et al.*, Extreme heat in North America, Europe and China in July 2023 made much more likely by climate change. http://spiral.imperial.ac.uk/handle/10044/1/105549. Accessed 11 July 2023.
15. Intergovernmental Panel on Climate Change, *Climate Change 2021–The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2023).
16. A. Robinson, J. Lehmann, D. Barriopedro, S. Rahmstorf, D. Coumou, Increasing heat and rainfall extremes now far outside the historical climate. *NPJ Clim. Atmos. Sci.* **4**, 45 (2021).
17. Z. Hausfather, H. F. Drake, T. Abbott, G. A. Schmidt, Evaluating the performance of past climate model projections. *Geophys. Res. Lett.* **47**, e2019GL085378 (2020).
18. Z. Hausfather, K. Marvel, G. A. Schmidt, J. W. Nielsen-Gammon, M. Zelinka, Climate simulations: Recognize the 'hot model' problem. *Nature* **605**, 26–29 (2022).
19. D. Barriopedro, R. García-Herrera, C. Ordóñez, D. G. Miralles, S. Salcedo-Sanz, Heat waves: Physical understanding and scientific challenges. *Rev. Geophys.* **61**, e2022RG000780 (2023).
20. K. A. McKinnon, I. R. Simpson, How unexpected was the 2021 Pacific Northwest heatwave? *Geophys Res. Lett.* **49**, e2022GL100380 (2022).
21. S. Bathiany, V. Dakos, M. Scheffer, T. M. Lenton, Climate models predict increasing temperature variability in poor countries. *Sci. Adv.* **4**, eaar5809 (2018).
22. D. Coumou, G. Di Capua, S. Vavrus, L. Wang, S. Wang, The influence of Arctic amplification on mid-latitude summer circulation. *Nat. Commun.* **9**, 2959 (2018).
23. M. G. Donat, A. J. Pitman, S. I. Seneviratne, Regional warming of hot extremes accelerated by surface energy fluxes. *Geophys. Res. Lett.* **44**, 7011–7019 (2017).
24. E. M. Fischer, C. Schär, Consistent geographical patterns of changes in high-impact European heatwaves. *Nat. Geosci.* **3**, 398–403 (2010).
25. J. A. Screen, Arctic amplification decreases temperature variance in northern mid- to high-latitudes. *Nat. Clim. Change* **4**, 577–582 (2014).
26. C. Huntingford *et al.*, Acceleration of daily land temperature extremes and correlations with surface energy fluxes. *NPJ Clim. Atmos. Sci.* **7**, 84 (2024).
27. A. K. Srivastava, M. Wehner, C. Bonfils, P. A. Ullrich, M. Risser, Local hydroclimate drives differential warming rates between regular summer days and extreme hot days in the Northern Hemisphere. *Weather Clim. Extrem.* **45**, 100709 (2024).
28. R. J. Haarsma *et al.*, High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev.* **9**, 4185–4208 (2016).
29. H. Hersbach *et al.*, The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
30. M. Patterson, North-West Europe hottest days are warming twice as fast as mean summer days. *Geophys. Res. Lett.* **50**, e2023GL102757 (2023).
31. E. Rousi, K. Kornhuber, G. Beobide-Arsuaga, F. Luo, D. Coumou, Accelerated Western European heatwave trends linked to more-persistent double jets over Eurasia. *Nat. Commun.* **13**, 3851 (2022).
32. R. Vautard *et al.*, Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends. *Nat. Commun.* **14**, 6803 (2023).
33. Y. Meng-shu, Y. Xiu-ping, Z. Tao, X. Xiao-feng, W. Shi-gong, Intercomparison of CRA-Interim Precipitation Products with ERA5 and JRA-55. *J. Trop. Meteorol.* **26**, 136–147 (2020).
34. D. Faranda, G. Messori, A. Jezequel, M. Vrac, P. Yiou, Atmospheric circulation compounds anthropogenic warming and impacts of climate extremes in Europe. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2214525120 (2023).
35. M. Rantanen *et al.*, The Arctic has warmed nearly four times faster than the globe since 1979. *Commun. Earth Environ.* **3**, 168 (2022).
36. D. G. Miralles, A. J. Teuling, C. C. van Heerwaarden, J. Vilà-Guerau de Arellano, Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.* **7**, 345–349 (2014).
37. M. P. Byrne, Amplified warming of extreme temperatures over tropical land. *Nat. Geosci.* **14**, 837–841 (2021).
38. E. Bevacqua, G. Zappa, F. Lehner, J. Zscheischler, Precipitation trends determine future occurrences of compound hot–dry events. *Nat. Clim. Change* **12**, 350–355 (2022).
39. D. Urdiales-Flores *et al.*, Drivers of accelerated warming in Mediterranean climate-type regions. *NPJ Clim. Atmos. Sci.* **6**, 97 (2023).
40. I. R. Simpson *et al.*, Observed humidity trends in dry regions contradict climate models. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2302480120 (2024).
41. L.-A. Kautz *et al.*, Atmospheric blocking and weather extremes over the Euro-Atlantic sector–A review. *Weather Clim. Dyn.* **3**, 305–336 (2022).
42. V. Petoukhov, S. Rahmstorf, S. Petri, H. J. Schellnhuber, Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5336–5341 (2013).
43. R. H. White, K. Kornhuber, O. Martius, V. Wirth, From atmospheric waves to heatwaves: A waveguide perspective for understanding and predicting concurrent, persistent, and extreme extratropical weather. *Bull. Am. Meteorol. Soc.* **103**, E923–E935 (2022).
44. J. A. Screen, I. Simmonds, Amplified mid-latitude planetary waves favour particular regional weather extremes. *Nat. Clim. Change* **4**, 704–709 (2014).
45. H. Teng, R. Leung, G. Branstator, J. Lu, Q. Ding, Warming pattern over the Northern Hemisphere midlatitudes in Boreal summer 1979–2020. *J. Clim.* **35**, 3479–3494 (2022).
46. J. Singh, S. Sippel, E. M. Fischer, Circulation dampened heat extremes intensification over the Midwest USA and amplified over Western Europe. *Commun. Earth Environ.* **4**, 432 (2023).
47. R. Schiemann *et al.*, Northern Hemisphere blocking simulation in current climate models: Evaluating progress from the Climate Model Intercomparison Project Phase 5 to 6 and sensitivity to resolution. *Weather Clim. Dyn.* **1**, 277–292 (2020).
48. T. Woollings *et al.*, Blocking and its response to climate change. *Curr. Clim. Change Rep.* **4**, 287–300 (2018).
49. P. Davini, F. D'Andrea, From CMIP3 to CMIP6: Northern Hemisphere atmospheric blocking simulation in present and future climate. *J. Clim.* **33**, 10021–10038 (2020).
50. K. Kornhuber *et al.*, Risks of synchronized low yields are underestimated in climate and crop model projections. *Nat. Commun.* **14**, 3528 (2023).
51. F. Luo *et al.*, Summertime Rossby waves in climate models: Substantial biases in surface imprint associated with small biases in upper-level circulation. *Weather Clim. Dyn.* **3**, 905–935 (2022).
52. G. Di Capua *et al.*, Drivers behind the summer 2010 wave train leading to Russian heatwave and Pakistan flooding. *NPJ Clim. Atmos. Sci.* **4**, 55 (2021).
53. A. Duchez *et al.*, Drivers of exceptionally cold North Atlantic Ocean temperatures and their link to the 2015 European heat wave. *Environ. Res. Lett.* **11**, 074004 (2016).
54. K. A. McKinnon, A. Rhines, M. P. Tingley, P. Huybers, Long-lead predictions of eastern United States hot days from Pacific sea surface temperatures. *Nat. Geosci.* **9**, 389–394 (2016).
55. D. L. Swain, D. E. Horton, D. Singh, N. S. Diffenbaugh, Trends in atmospheric patterns conducive to seasonal precipitation and temperature extremes in California. *Sci. Adv.* **2**, e1501344 (2016).
56. S. E. Perkins-Kirkpatrick, S. C. Lewis, Increasing trends in regional heatwaves. *Nat. Commun.* **11**, 3357 (2020).
57. R. Seager *et al.*, Ocean-forcing of cool season precipitation drives ongoing and future decadal drought in southwestern North America. *NPJ Clim. Atmos. Sci.* **6**, 141 (2023).
58. R. Seager *et al.*, Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nat. Clim. Change* **9**, 517–522 (2019).
59. R. Seager, N. Henderson, M. Cane, Persistent discrepancies between observed and modeled trends in the tropical Pacific Ocean. *J. Clim.* **35**, 4571–4584 (2022).
60. L. A. Lee, C. L. Reddington, K. S. Carslaw, On the relationship between aerosol model uncertainty and radiative forcing uncertainty. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5820–5827 (2016).
61. P. Wang *et al.*, Aerosols overtake greenhouse gases causing a warmer climate and more weather extremes toward carbon neutrality. *Nat. Commun.* **14**, 7257 (2023).
62. D. L. Schumacher *et al.*, Exacerbated summer European warming not captured by climate models neglecting long-term aerosol changes. *Commun. Earth Environ.* **5**, 182 (2024).
63. J. Gao *et al.*, Fast climate responses to emission reductions in aerosol and ozone precursors in China during 2013–2017. *Atmos. Chem. Phys.* **22**, 7131–7142 (2022).
64. V. Thompson *et al.*, The most at-risk regions in the world for high-impact heatwaves. *Nat. Commun.* **14**, 2152 (2023).

65. D. I. Armstrong McKay *et al.*, Exceeding 1.5°C global warming could trigger multiple climate tipping points. *Science* **377**, eabn7950 (2022).
66. T. M. Lenton *et al.*, Tipping elements in the Earth's climate system. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1786–1793 (2008).
67. H. J. Schellnhuber, S. Rahmstorf, R. Winkelmann, Why the right climate target was agreed in Paris. *Nat. Clim. Change* **6**, 649–653 (2016).
68. B. Stevens *et al.*, Earth Virtualization Engines (EVE). *Earth Syst. Sci. Data* 16, 2113–2122 (2024). https://essd.copernicus.org/preprints/essd-2023-376/.
69. B. J. Harvey, P. Cook, L. C. Shaffrey, R. Schiemann, The response of the Northern Hemisphere storm tracks and jet streams to climate change in the CMIP3, CMIP5, and CMIP6 climate models. *J. Geophys. Res. Atmos.* **125**, e2020JD032701 (2020).
70. D. Holley, S. H. Lee, Forecasting extreme heat in the UK during July 2022. *Weather* **77**, 320–321 (2022).
71. E. M. Fischer *et al.*, Storylines for unprecedented heatwaves based on ensemble boosting. *Nat. Commun.* **14**, 4643 (2023).
72. F. Ragone, J. Wouters, F. Bouchet, Computation of extreme heat waves in climate models using a large deviation algorithm. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 24–29 (2018).
73. D. M. Smith *et al.*, Attribution of multi-annual to decadal changes in the climate system: The Large Ensemble Single Forcing Model Intercomparison Project (LESFMIP). *Front. Clim.* **4**, 955414 (2022).
74. C. Lesk *et al.*, Stronger temperature–moisture couplings exacerbate the impact of climate warming on global crop yields. *Nat. Food* **2**, 683–691 (2021).
75. K. Bi *et al.*, Accurate medium-range global weather forecasting with 3D neural networks. *Nature* **619**, 533–538 (2023).
76. R. Lam *et al.*, Learning skillful medium-range global weather forecasting. *Science* **382**, 1416–1421 (2023).
77. T. Schneider *et al.*, Harnessing AI and computing to advance climate modelling and prediction. *Nat. Clim. Change* **13**, 887–889 (2023).
78. J. Yuval, P. A. O'Gorman, Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions. *Nat. Commun.* **11**, 3295 (2020).
79. C. Funk *et al.*, A high-resolution 1983–2016 Tmax climate data record based on infrared temperatures and stations by the climate hazard center. *J. Clim.* **32**, 5639–5658 (2019).
80. S. Kobayashi *et al.*, The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteorol. Soc. Jpn. Ser.* **93**, 5–48 (2015).
81. R. C. Cornes, G. van der Schrier, E. J. M. van den Besselaar, P. D. Jones, An ensemble version of the E-OBS Temperature and precipitation data sets. *J. Geophys. Res. Atmos.* **123**, 9391–9409 (2018).
82. M. A. Haugen, M. L. Stein, E. J. Moyer, R. L. Sriver, Estimating changes in temperature distributions in a large ensemble of climate simulations using quantile regression. *J. Clim.* **31**, 8573–8588 (2018).
83. K. A. McKinnon, A. Rhines, M. P. Tingley, P. Huybers, The changing shape of Northern Hemisphere summer temperature distributions. *J. Geophys. Res. Atmos.* **121**, 8849–8868 (2016).
84. R. Gelaro *et al.*, The modern-era retrospective analysis for research and applications, Version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
85. Python Software Foundation, *Python Language Reference, version 3.6*. http://www.python.org. Accessed 1 June 2023.
86. S. Bartusek, Extreme Heatwave Scripts. Zenodo. https://doi.org/10.5281/zenodo.13993309. Deposited 25 October 2024.
87. Laurence Livermore National Lab, *Earth* System Grid Federation. ESGF. https://wcrp-cmip.org/nodes/aims3-llnl-gov/. Deposited 1 March 2024.
88. R. J. Haarsma *et al.*, High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geosci. Model Dev*. 9, 4185–4208 (2016).