POTSDAM-INSTITUT FÜR KLIMAFOLGENFORSCHUNG

Constraining the Pattern and Magnitude of Projected Extreme Precipitation Change in a Multimodel Ensemble

MAXIMILIAN KOTZ,[a,b] STEFAN LANGE,[a] LEONIE WENZ,[a,c] AND ANDERS LEVERMANN[a,b,d]

[a] *Potsdam Institute for Climate Impact Research, Potsdam, Germany*
[b] *Institute of Physics, University of Potsdam, Potsdam, Germany*
[c] *Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany*
[d] *Lamont-Doherty Earth Observatory, Columbia University, New York, New York*

(Manuscript received 19 August 2023, in final form 2 October 2023, accepted 2 October 2023)

ABSTRACT: Projections of precipitation extremes over land are crucial for socioeconomic risk assessments, yet model discrepancies limit their application. Here we use a pattern-filtering technique to identify low-frequency changes in individual members of a multimodel ensemble to assess discrepancies across models in the projected pattern and magnitude of change. Specifically, we apply low-frequency component analysis (LFCA) to the intensity and frequency of daily precipitation extremes over land in 21 CMIP-6 models. LFCA brings modest but statistically significant improvements in the agreement between models in the spatial pattern of projected change, particularly in scenarios with weak greenhouse forcing. Moreover, we show that LFCA facilitates a robust identification of the rates at which increasing precipitation extremes scale with global temperature change within individual ensemble members. While these rates approximately match expectations from the Clausius-Clapeyron relation on average across models, individual models exhibit considerable and significant differences. Monte Carlo simulations indicate that these differences contribute to uncertainty in the magnitude of projected change at least as much as differences in the climate sensitivity. Last, we compare these scaling rates with those identified from observational products, demonstrating that virtually all climate models significantly underestimate the rates at which increases in precipitation extremes have scaled with global temperatures historically. Constraining projections with observations therefore amplifies the projected intensification of precipitation extremes as well as reducing the relative error of their distribution.

KEYWORDS: Atmosphere; Precipitation; Climate change; Empirical orthogonal functions; Pattern detection; General circulation models

1. Introduction

An intensification of the hydrological cycle (Ziegler et al. 2003) is likely to play a major role in the socioeconomic impacts of climate change. Key determinants of social welfare, such as groundwater availability (Thomas and Famiglietti 2019), agricultural productivity (Liang et al. 2017), and social stability (Hsiang et al. 2013; von Uexkull et al. 2016), are closely linked to changes in precipitation. In particular, the changing frequency and intensity of heavy precipitation extremes constitutes an important channel, given their links to flooding, which can cause considerable loss of life (Schumacher 2017) and financial losses (Davenport et al. 2021; Frame et al. 2020), as well as impacts on overall macroeconomic productivity (Kotz et al. 2022). Understanding how these characteristics of precipitation will change under anthropogenic influence is therefore crucial for informing risk assessments and climate policy. Climate models such as those in the Coupled Model Intercomparison Project (CMIP6; Eyring et al. 2016) play a crucial role in this understanding, providing projections of precipitation extremes under different levels of greenhouse forcing at the regional and temporal detail necessary for assessment of these impacts (Warszawski et al. 2014). These projections can subsequently inform detailed assessments of the consequences of both mitigation (Lange et al. 2020; Thiery et al. 2021) and adaptation (Willner et al. 2018; Boulange et al. 2021).

Despite considerable progress, barriers to an effective use of these projections remain due to uncertainty in projected precipitation change (Knutti and Sedláček 2013; Shepherd 2014). The simplest theory regarding precipitation extremes indicates that the scaling of available moisture with atmospheric temperatures based on the Clausius-Clapeyron relation (CC-relation) (Allen and Ingram 2002; Santer et al. 2007; Fischer and Knutti 2016) should lead to increases in precipitation extremes with global warming at rates of 6%–7% K^{-1}. Indeed, intensification of precipitation extremes have been detected in almost all global land areas (Min et al. 2011; Zhang et al. 2013; Fischer and Knutti 2016; Chen and Sun 2017; Kirchmeier-Young and Zhang 2020; Madakumbura et al. 2021), and when aggregated spatially are largely consistent with expectations from the CC-relation (Fischer and Knutti 2016). However, a number of other factors determine changes in precipitation extremes including the vertical structure of upward atmospheric velocities and their change (O'Gorman and Schneider 2009), as well as shifts in atmospheric circulation (Pfahl et al. 2017). These factors contribute significantly to regional differences in the intensification of precipitation extremes (O'Gorman and Schneider 2009) and are also the dominant source of uncertainty across climate model projections (Pfahl et al. 2017).

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-23-0492.s1.

Corresponding author: Maximilian Kotz, maxkotz@pik-potsdam.de

DOI: 10.1175/JCLI-D-23-0492.1

In addition to these uncertainties in the response of precipitation extremes to human forcing, further uncertainty in future projections exists due to the role of chaotic internal variability (Shepherd 2014; Fischer et al. 2014). This is particularly relevant when internal variability is large in comparison with the strength of the anthropogenically forced signal, such as when considering changes over short time scales, in scenarios of weak greenhouse forcing or in variables with a weaker response to anthropogenic influence (Blanusa et al. 2023). The latter is certainly true for changes in precipitation in comparison with temperature (King et al. 2015), emphasizing the possibility that internal variability may mask or bias assessments of how precipitation extremes change in response to human influence. An effective method to quantify and separate the contributions of human forcing and internal variability to projected precipitation change is the use of large ensembles of climate models initialized from different starting conditions (Aalbers et al. 2018; Wood and Ludwig 2020). However, such single-model initial-condition large ensembles (SMILEs) (Kay et al. 2015) do not enable the investigation of discrepancies that may exist between different climate models in the simulated response of precipitation extremes to human forcing.

Here we aim to assess uncertainties in projected changes in daily precipitation extremes that arise from the different responses to anthropogenic forcing embodied in the different members of the multimodel ensemble CMIP6. Our interest in the differences between climate models constrains us to identify changes in each ensemble member individually. To limit the influence of chaotic internal variability, we therefore use a pattern-filtering methodology that has been designed to separate the forced response of the climate system from internal variability within individual ensemble members. This method, known as low-frequency component analysis (LFCA), extends an empirical orthogonal function (EOF) decomposition to search for linear recombinations of the orthogonal basis functions with a maximal ratio of low-frequency to total variance (Wills et al. 2018). Given the lower frequency at which anthropogenically forced signals evolve in comparison with chaotic internal variability, this approach can help to discriminate between them while drawing on the structure of the spatiotemporal covariance matrix to do so. When tested in the context of large ensembles, this method shows particular promise in separating the anthropogenic component of projected temperature change from individual ensemble members with greater accuracy than large ensembles with up to 20 members (Wills et al. 2020). Although not yet tested against large ensembles in the context of precipitation, this method nevertheless presents an opportunity to limit the role of internal variability when assessing discrepancies in the projected changes of extreme precipitation in individual members of the multimodel CMIP6 ensemble.

We assess changes in both the intensity and frequency of daily precipitation extremes, focusing on changes over land because of their relevance for impacts (Davenport et al. 2021; Kotz et al. 2022) and for better comparison with observations as is made in the final section. In the following section (section 2) we provide an overview of the data and methods. Section 3a presents the low-frequency changes in extreme precipitation detected by LFCA in the individual members of the CMIP6 ensemble. In section 3b, we assess discrepancies across models in the spatial patterns of projected change and the extent to which LFCA limits this uncertainty. Section 3c addresses how differences in the magnitude of projected precipitation change are constrained by scaling relations with global temperatures between climate models, but also within individual models. In section 3d, we assess the contribution of differences in temperature–precipitation scaling rates to uncertainty in the magnitude of projected change, as well as how these rates compare to those observed historically.

2. Data and methods

a. Climate data

We use daily surface precipitation rates and daily 2-m temperature from 21 climate models participating in CMIP6 (Eyring et al. 2016). We choose models that provide output under both the historical (1850–2014) and the future (2015–2100) greenhouse forcing scenarios specified by SSP126 and SSP585. Since our approach is primarily interested in differences between climate models, and since many models provide only a single ensemble member, we use only the first ensemble member available for each model. A full list of models is displayed in Fig. 5 and in Figs. S1–S4 in the online supplemental material. Additionally, we use daily 2-m temperature and daily precipitation totals from ERA5 (Hersbach et al. 2020), as well as daily precipitation totals from the Global Precipitation Climatology Centre (Schneider et al. 2008).

b. Extreme precipitation metrics

We assess changes in both the intensity and frequency of daily precipitation extremes, following work indicating that both are relevant for impacts (Kotz et al. 2022). We assess the intensity of daily extremes by taking the annual maximum, RX1, of daily precipitation $P_{x,d}$:

$$\mathrm{RX1}_{x,y} = \max(P_{x,d} : d = 1, \ldots, N_y). \tag{1}$$

The annual frequency of daily extremes is assessed by counting the number of days in each year that exceed the 99$^{\text{th}}$ percentile of the historical distribution:

$$R > 99\%_{x,y} = \sum_{d=1}^{N_y} H(P_{x,d} > 99\%), \tag{2}$$

where N_y is the number of days in a given year, H is the Heaviside step function, and x and d respectively denote grid cell and day. The historical period used to estimate percentiles is either the preindustrial period (1850–1950) or the historical period over which observational data are available (1985–2015), depending on the use. These annual measures are estimated on the native CMIP grids before being linearly interpolated to a common 1° by 1° grid for further analysis.

c. Low-frequency component analysis

We use LFCA to estimate low-frequency changes in extreme precipitation. LFCA uses an EOF decomposition and subsequent linear-recombination under a frequency-dependent variance maximization constraint to optimally filter low-frequency changes

from higher-frequency modes. As such, it has been designed (Wills et al. 2018) and demonstrated (Wills et al. 2020) to separate forced climate changes from chaotic internal variability. We use LFCA over other methods, such as signal-to-noise maximizing pattern-filtering, because of our interest in identifying changes in individual members of the multimodel CMIP6 ensemble, in which context LFCA provides a superior performance (Wills et al. 2020). Here we provide a conceptual summary of LFCA and of its application to identifying the climatic response to anthropogenic forcing; please see Wills et al. (2018) for a more detailed introduction to and description of the method.

Linear recombinations of the leading EOFs are found that maximize the ratio of low-frequency to total variance they can explain. We retain a sufficient number of EOFs to account for a minimum of 70% of the original spatiotemporal variance and define low-frequency variance as that resulting from applying a 20-yr low-pass Butterworth filter with reflecting boundary conditions to gridcell-level departures from linear trends. We use a lower cutoff frequency than Wills et al. (2018, 2020) because of the lower signal-to-noise ratio of the climate change signal in precipitation than in temperature (Deser et al. 2014) but recover consistent results under alternative filtering specifications. The resulting linear recombinations are independent and ordered in terms of increasing frequency. They constitute both a "low-frequency component" (LFC) and "low-frequency pattern" (LFP); the LFC is a time series that describes the temporal evolution of the specific spatial pattern encompassed by the LFP. The lowest-frequency component(s) of LFCA have been found to accurately characterize the forced response of the climate system to anthropogenic forcing in the context of temperature (Wills et al. 2020) and temperature variability (Kotz et al. 2021). However, given that tests of such methods in large ensembles have not yet been conducted for precipitation extremes, we here interpret the lowest-frequency component identified by LFCA as the lowest-frequency change projected by each ensemble member rather than the forced change.

We apply LFCA to annual time series of the intensity (Rx1) and frequency ($R > 99\%$) of daily precipitation extremes over land areas and to annual mean temperature (over the globe) from 1950 to 2100 under the anthropogenic forcing of the historical period and both the SSP126 and SSP585 future scenarios, having first linearly interpolated each index to a common 1° by 1° grid. Low-frequency changes between two given time periods (usually between two decades) are then calculated as the product of the lowest-frequency LFP and the difference between temporal (usually decadal) averages of the corresponding LFC. In some cases, we also estimate the forced changes in each precipitation metric using temporal averages of the data without the use of LFCA to first identify low-frequency changes. In these cases, we use 10-, 20-, or 30-yr periods to estimate the change.

d. *Assessing similarity in the patterns of change*

We use two metrics to assess similarities in the patterns of change detected by LFCA. The first is simply the number of ensemble members that agree on the projected sign of change in a given region. This metric is aggregated in Fig. 2 over global land areas to indicate the fraction of land area on which a certain percentage of models are in agreement. The second metric is a centered pattern correlation as defined in Santer et al. (1995). This metric reflects differences in the relative magnitude of change as well as the sign of change. Pattern correlations are estimated between each unique pair of climate models (210 pairs) and distributions of these pattern correlations are shown and used to estimate the significance of improvements in intermodel agreement.

e. *Scaling between precipitation extremes and global temperature change*

Spatially averaged changes in precipitation extremes are compared with changes in global temperature change. Changes in global temperature are calculated by taking area-weighted averages of changes in global 2-m surface air temperature (SAT) estimated from the lowest-frequency component identified with LFCA. Changes in precipitation extremes are expressed as local percentages before taking an area-weighted average, to make a comparison with the theoretical expectations of the CC relation. We explore the sensitivity of our results to arid regions with low baseline values in latter parts of the paper. Changes are estimated either over the whole simulation period (from 1950–60 to 2090–2100, as in Fig. 3), or between pairs of nonoverlapping decades separated by 30 years within each individual model (Fig. 4 onward). In the first case, percentages are expressed in relation to the preindustrial baseline (1850–1950), land-area-averages are taken, and scaling relations are then estimated in a logarithmic form, $\log[(\delta P/P) + 1]$ to reflect the exponential scaling anticipated from the CC relation. In the second case, percentages are estimated with respect to the baseline periods from which each change is estimated. Here, since the baseline period for estimating percentage changes is updated with each differencing period, only a linear scaling is estimated, after taking land-area averages of local percentage changes.

Temperature–precipitation scaling relations are estimated using least squares linear regressions. In each case, the intercept of these regressions is set to zero to reflect the physical constraint that without changes in global SAT there should be no change in precipitation extremes. Methodological uncertainty in the estimates of these scaling rates is obtained via a bootstrapping approach. In Fig. 3, the climate models are resampled 1000 times with replacement, whereas in Fig. 4 the different time periods are resampled equivalently. As such, these estimates reflect the methodological uncertainty arising from the availability of climate models (in Fig. 3) and the temporal evolution of low-frequency changes in projected precipitation change (in Fig. 4). When estimating the significance of differences between model scaling rates, we estimate a distribution of differences between scaling rates from the estimates obtained from the 1000 resamples of the time series and assess whether the 95% confidence intervals of this distribution of differences encompass zero.

f. *Monte Carlo simulations to propagate uncertainties and assess relative contributions*

We use a Monte Carlo simulation procedure to assess how different sources contribute to uncertainty in the overall magnitude

of projected extreme precipitation change. Specifically, we assess two contributions, from 1) differences in the projected change in global mean temperatures for a given forcing scenario (loosely equivalent to the transient climate sensitivity), and 2) differences in the rates at which precipitation extremes scale with global temperature change. We sample 10 000 times from the distribution of these two parameters and combine them to provide estimates of the distribution of projected change under uncertainty from both sources. This approach is justified by the fact that temperature–precipitation scaling rates appear independent of global temperature change across models. The distribution of projected global temperature change is obtained from the 21 different climate models of the ensemble, while the distribution of temperature-precipitation scaling rates comes from both the different climate models and the bootstrapped assessments of uncertainty in their individual scaling rates (as described above). Last, we estimate a simple observational constraint by replacing the distribution of temperature–precipitation scaling rates with the estimates obtained from the two observational data products.

3. Results

a. *Low-frequency changes in projected precipitation extremes*

We identify low-frequency changes in the intensity and frequency of daily precipitation extremes under historical and future (SSP585) anthropogenic forcing using LFCA. Results shown in Fig. 1 are for a selection of 10 of the 21 CMIP-6 models, taking every other model when ordered in terms of SAT change (see Figs. S1–S4 in the online supplemental material for results from all members of the multimodel ensemble under both SSP585 and SSP126 future forcing). For each precipitation metric and for each model, the lowest-frequency component (LFC-1) exhibits a near-monotonic trend that closely follows the increasing concentrations of greenhouse gases in the historical and SSP585 scenario. However, both the intensity and pattern of low-frequency change (from 1950–60 to 2090–2100) show clear differences between models. While most land areas exhibit increases in both the intensity and frequency of extremes, stronger increases are typically found across tropical regions. Decreases are projected in certain models, in particular in regions where net precipitation is projected to decline such as the Mediterranean Sea basin, parts of Australia, South Africa, and central South America (IPCC 2021). These regions also correspond to those identified by Pfahl et al. (2017) as having larger contributions from changes in atmospheric circulation.

In comparing the projected low-frequency changes in extreme precipitation with the magnitude of historical variability (standard deviation of annual time series from 1850 to 1900), we find that changes by the end of the century are not always significant with respect to the magnitude of internal variability (see stippled regions in Fig. 1). Changes are most often of a significant magnitude across the tropics and high latitudes, showing the least significance across midlatitudes and in regions where precipitation extremes are projected to decrease. Moreover, models with larger SAT change typically show significant changes over larger portions of the global land area (see Figs. A1a–d in the appendix). This reflects the fact that SAT change is a strong determinant of the magnitude of changes in precipitation extremes (explored further in section 3c), and therefore of the signal-to-noise ratio of forced change to internal variability. This can be further seen in the fact that models with larger SAT change contain a larger portion of the overall spatiotemporal variance within the lowest-frequency component detected by LFCA (appendix Figs. A1e–h). These results are robust to methodological choices regarding the LFCA, such as the specific percentage of spatiotemporal variance that is retained following the EOF decomposition as well as the time scale of the filter used to define low-frequency variance.

b. *Discrepancies in the spatial pattern of projected low-frequency change*

Having identified low-frequency changes within each member of the multimodel ensemble, we assess discrepancies in the spatial pattern of projected change between models (Fig. 2). In the ensemble mean, projected low-frequency changes reflect a near-global intensification of the intensity and frequency of daily precipitation extremes (Figs. 2a,b), with some weak decreases in regions where reductions in net precipitation are particularly strong (IPCC 2021) and where changes in atmospheric circulation contribute more considerably (Pfahl et al. 2017). Agreement between models is fairly high, with 90% of models agreeing on the sign of change on approximately 80% and 70% of the global land area for changes in the intensity and frequency of extremes respectively in the high forcing scenario (Figs. 2d,f). Discrepancies are concentrated across the Mediterranean basin, Australia, South Africa, and central South America (Figs. 2a,b). These regions are those where decreases in precipitation extremes are projected, where projected changes are less significant with respect to internal variability, and where atmospheric circulation has been identified as a dominant source of uncertainty (Pfahl et al. 2017).

Agreement across climate models on the spatial pattern of change is smaller under the scenario with weaker greenhouse forcing (SSP126). This is clear both in terms of the fraction of land area on which models project the same sign of change (Figs. 2c–f) and also in terms of the distribution of pairwise centered pattern correlations between climate models (Figs. 2g–j). This likely reflects the smaller signal-to-noise ratio of anthropogenically forced precipitation change relative to chaotic internal variability in the weaker forcing scenario. This can be seen in the smaller magnitude of projected precipitation change and the smaller percentage of spatiotemporal variance explained by the lowest-frequency component in the weaker forcing scenario (Figs. A1a–h in the appendix). When internal variability is relatively larger, it will mask the anthropogenically forced signal identified by LFCA to a greater extent. This will lead to greater disagreement between models resulting from both discrepancies in the forced response, and the larger differences due to internal variability.

Furthermore, the extent of agreement between models is higher when using LFCA to identify low-frequency changes in precipitation extremes rather than temporal averages over 10-, 20-, and 30-yr periods. These benefits manifest both in terms of the fraction of land-area in agreement (Figs. 2c–f) and the

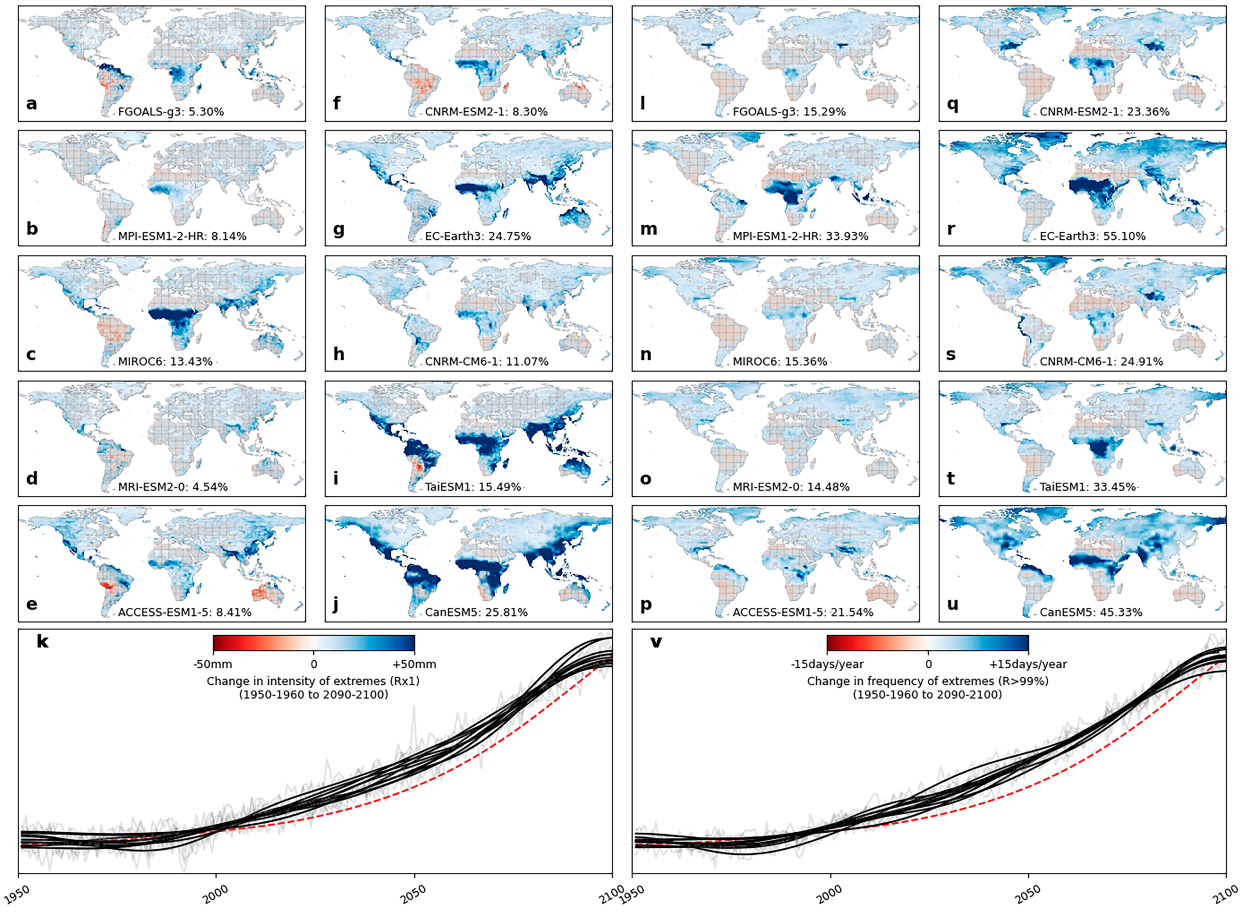

FIG. 1. Low-frequency changes in extreme daily precipitation under historical (1950–2014) and future (SSP585; 2015–2100) anthropogenic forcing, detected in individual CMIP6 climate models with low-frequency component analysis. (a)–(j) The spatial pattern of low-frequency change in the intensity of daily precipitation extremes (annual maximum; Rx1) from 1950–60 to 2090–2100 (the product of the lowest-frequency pattern and the difference between decadal averages of its corresponding component) for 10 of the 21 CMIP6 ensemble members. Regions are cross hatched where the projected change is smaller than the historical variability, estimated as the standard deviation of annual values over the period 1850–1900. (k) The temporal evolution of the corresponding lowest-frequency components (LFC-1), shown in gray with a 20-yr Butterworth filtered time series in black. Time series for each model are overlain because of their similarity. The concentrations of greenhouse gases in the historical period and SSP585 are rescaled and shown in red for comparison. (l)–(v) Corresponding results for changes in the frequency of daily precipitation extremes (the number of days with rainfall greater than the 99th percentile over the historical period 1850–1950). The model name is indicated at the bottom of each panel, along with the percentage of total spatiotemporal variance accounted for by the lowest-frequency component. We here show models chosen evenly from across the distribution of modeled global mean surface air temperature (SAT) change (ordered by increasing SAT change). Results for all ensemble members for each precipitation metric under SSP585 and SSP126 forcing scenarios are shown in Figs. S1–S4 in the online supplemental material.

centered pattern correlations (Figs. 2g–j). Improvements are strongest in the weaker forcing scenario and when compared with taking averages over shorter periods. Under SSP126 forcing, improvements in centered pattern correlations are found in 99% and 93% of model pairs when compared even with using 30-yr averages, and differences between the distributions of correlations obtained using the two methods are significant at the 0.1% level. Improvements in the stronger forcing scenario for the intensity of precipitation extremes are still significant at the 5% level, with 88% of model pairs showing improved agreement (Table 1). Assuming that multimodel agreement is an indicator of the extent to which the anthropogenic signal has been separated from chaotic internal variability, as outlined above, these improvements suggest that LFCA may improve the detection of the response of precipitation extremes to anthropogenic forcing. This conclusion would require further verification in the context of SMILEs, which is beyond the scope of this paper. Nevertheless, these results demonstrate that LFCA significantly improves agreement in the pattern of projected change between members of a multimodel ensemble, particularly in the context of weak anthropogenic forcing.

c. *The scaling of precipitation extremes with global temperatures*

In addition to uncertainty in the pattern of projected low-frequency change, the magnitude of extreme precipitation

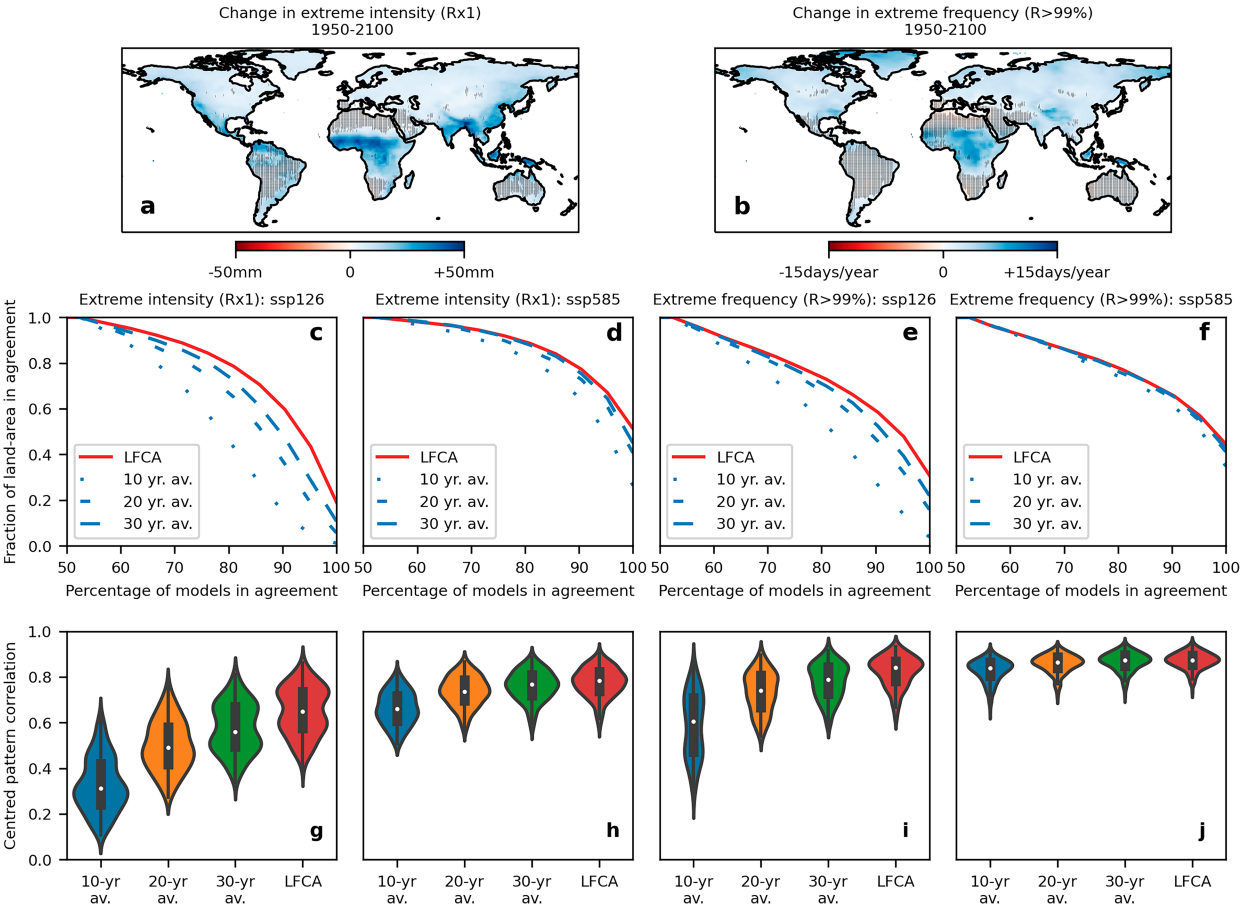

FIG. 2. An assessment of discrepancies in the spatial pattern of low-frequency change in extreme precipitation: The CMIP-6 ensemble-mean low-frequency change in the (a) intensity and (b) frequency of daily precipitation extremes. Changes are estimated from the lowest-frequency component detected with LFCA (see Fig. 1) over the period 1950–2100 under historical and SSP585 future anthropogenic forcing. Regions are stippled where less than 90% of models agree on the sign of projected change. The fraction of land area on which a certain percentage of CMIP-6 models agree on the sign of low-frequency change for the (c),(d) intensity and (e),(f) frequency of daily precipitation extremes under the low-greenhouse [SSP126; in (c) and (e)] and high-greenhouse [SSP585; in (d) and (f)] forcing scenarios. Results are shown when using either LFCA (red) or averages over 10, 20, and 30 years (blue) to detect low-frequency changes. (g)–(j) Distributions of centered pattern correlations between the low-frequency change projected by unique pairs of models for the two extreme precipitation metrics under the two forcing scenarios, when using the different methods to detect low-frequency changes. Violin plots from the Seaborn Python package indicate the interquartile range and upper and lower limits, as well as a kernel-density distribution.

changes also exhibits large differences across models (Fig. 1). Changes in global SAT explain a large proportion of these differences when averaged over global land areas (r^2 of 0.73 and 0.74 for the intensity and frequency of extremes, respectively; Fig. 3). Moreover, changes in the intensity of extremes scale with global SAT at a rate of 6.5% K^{-1}, closely following the expectations of the CC relation within the uncertainty of the scaling relation we estimate from the climate

TABLE 1. An assessment of the significance of improvements in multimodel agreement on the spatial pattern of low-frequency change in extreme precipitation metrics, showing the percentage of unique model pairs with improved centered pattern correlations in low-frequency precipitation change when detected using LFCA in comparison with averages over 10-, 20-, and 30-yr periods. The significance of the difference between the distributions of pattern correlations between models under the two different methods is assessed using the nonparametric Mann–Whitney test, with significance at the 5%, 1%, and 0.1% levels indicated by one, two, and three asterisks, respectively.

Metric: SSP	LFCA > 10-yr avg	LFCA > 20-yr avg	LFCA > 30-yr avg
Rx1: SSP126	100^{***}	100^{***}	99^{***}
Rx1: SSP585	100^{***}	100^{***}	88^{*}
$R > 99\%$: SSP126	100^{***}	95^{***}	93^{***}
$R > 99\%$: SSP585	97^{***}	82^{**}	58

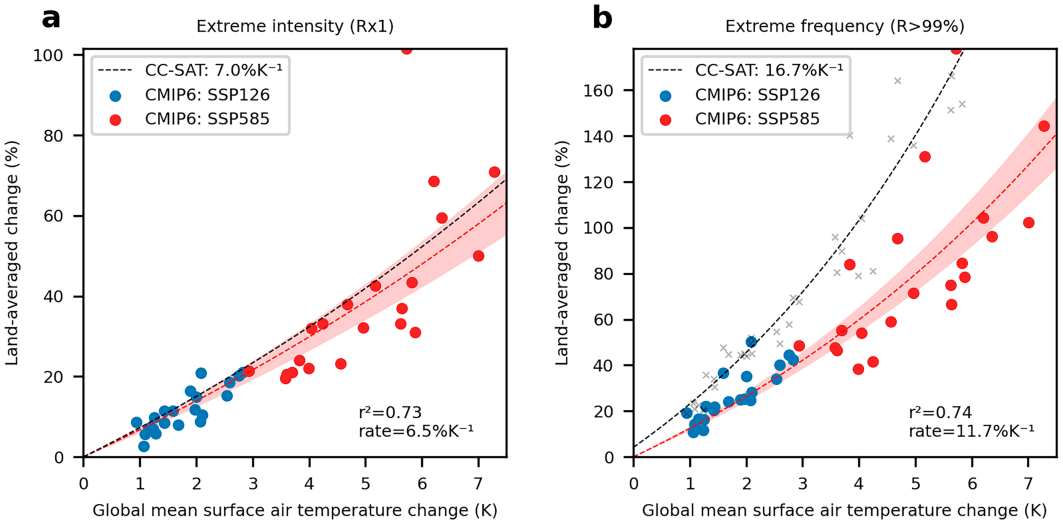

FIG. 3. The scaling of precipitation extremes over land with global mean surface air temperature (SAT) across CMIP6 models and scenarios. Low-frequency changes between 1950–60 and 2090–2100 are calculated from the lowest-frequency component of each precipitation metric (as in Fig. 1) and of annual mean temperature. Red and blue colors denote the SSP585 and SSP126 scenarios of future greenhouse forcing. The results of least squares linear regressions are shown in dashed red with the 5th and 95th confidence intervals shaded based on bootstrapped estimates of the regression (1000 climate model resamples with replacement). Changes in extreme metrics are calculated as a percentage of the historical baseline (1850–1950) for comparison with the theoretical expectations of the Clausius-Clapeyron (CC) relation, shown as dashed black lines. For the frequency of daily extremes ($P > 99\%$), a theoretical CC relation is estimated by scaling up each day of the historical precipitation distribution (1850–50) by the given level of SAT change and recalculating each index, following Fischer and Knutti (2016). Individual estimates from this method are shown in gray, with the black dashed line showing the result of an exponential regression to these estimates. The resulting theoretical scaling rate of this regression is displayed in the figure legend.

models. For the frequency of extremes, we assess a theoretical expectation from the CC relation by scaling up every day of the historical distribution of precipitation by 7% K^{-1} and re-estimating the exceedance of the original percentile-based thresholds [following the methods of Fischer and Knutti (2016)]. This theoretical CC relation predicts a scaling rate of 16.7% K^{-1}, and we find that the frequency of daily extremes does indeed increase at a higher rate of 11.7% K^{-1}, but one that falls considerably below this theoretical expectation. These results replicate many recent findings that changes in precipitation extremes scale with global SAT change across climate models, emphasizing the importance of limiting global warming to limit future impacts as well as the importance of the climate sensitivity as a dominant contributor to their uncertainty. Furthermore, results are robust to excluding regions with very low initial precipitation values, which could potentially bias our assessment using land-area averages of local percentage changes (see Fig. S5 in the online supplemental material).

While the results of Fig. 3 demonstrate that the intensification of precipitation extremes is largely determined by SAT changes, it is of further interest whether such temperature–precipitation scaling relations can be identified within individual models, and if so, how their rates may differ. We address this question by using the intertemporal changes in precipitation extremes and global SAT within a given ensemble member, taking the changes occurring between nonoverlapping decades separated by 30-yr periods within each of the two forcing scenarios for each climate model. This approach reveals robust temperature–precipitation scaling relations that explain a large proportion of the intertemporal changes in precipitation within each individual ensemble member (Fig. 4). The average R^2 of these relations is 0.82 and 0.78 for the intensity and frequency of precipitation extremes, respectively, demonstrating more robust scalings than those identified across climate models. This suggests that the physical processes that constrain the intensification of precipitation extremes to scale with SAT across climate models also hold within individual models across time.

When compared with using temporal averages over 10-, 20-, and 30-yr periods, we find that first applying LFCA to identify the lowest-frequency changes considerably improves the robustness of the identified scaling relations. Figures S6 and S7 in the online supplemental material show the scaling relations identified with LFCA for all ensemble members as compared with those in supplemental Figs. S8 and S9 without LFCA. When using temporal averages over longer periods (20 or 30 years) fewer observations are available to construct the scaling relations and the R^2 typically increases both with and without LFCA (results using 30-yr periods are shown in Figs. S10–S13 in the online supplemental material). Table 2 documents the improvements of using LFCA, showing that, even compared with 30-yr averages, LFCA improves the R^2 of the identified relations by 0.11 and 0.22 for the intensity and frequency of precipitation extremes

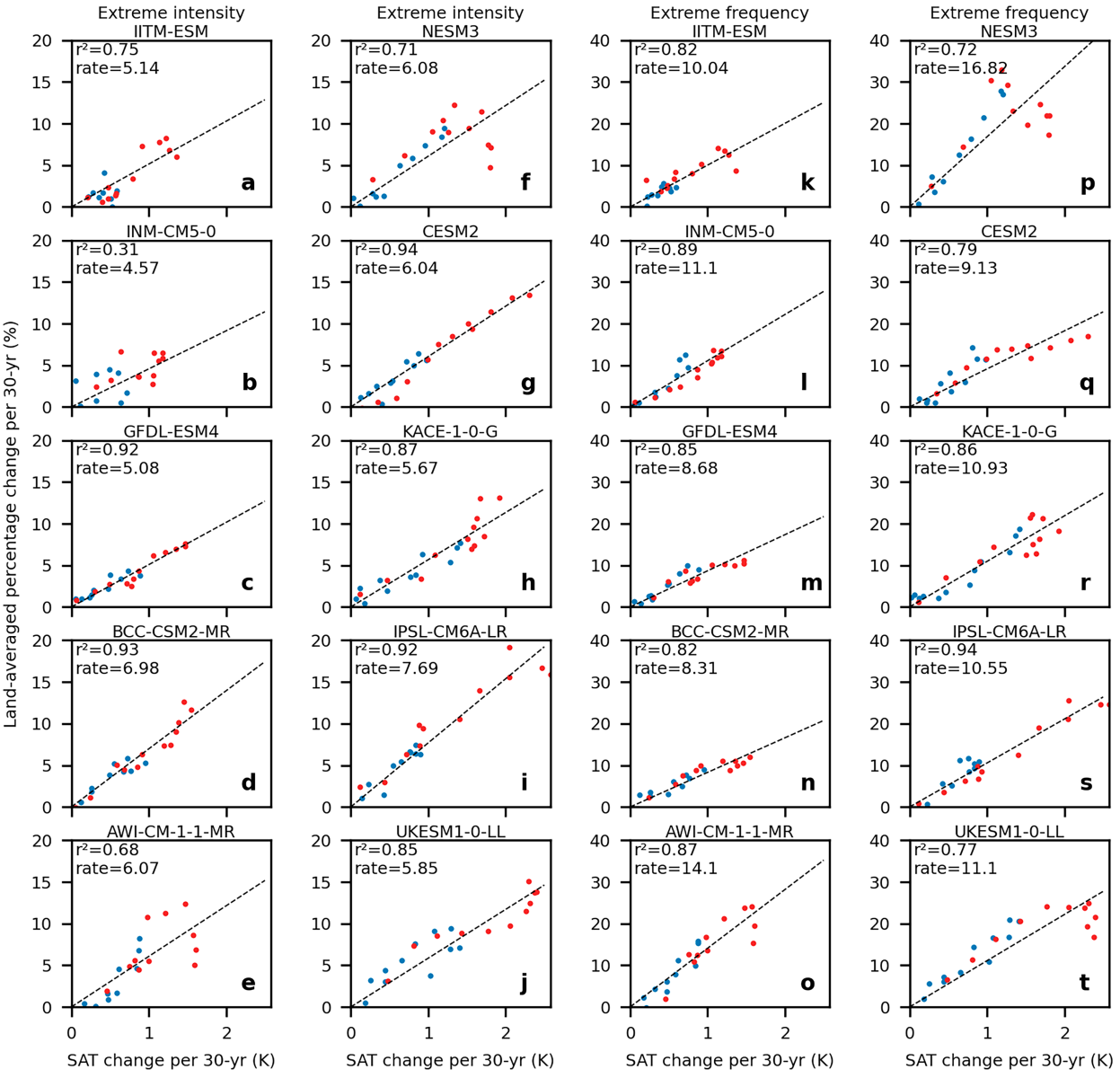

FIG. 4. The scaling of precipitation extremes over land with global mean SAT within individual CMIP6 climate models: (a)–(j) Land-averaged percentage changes in the lowest-frequency component of extreme intensity (Rx1) between pairs of nonoverlapping decades separated by 30 years under a low-greenhouse (SSP126; blue) and high-greenhouse (SSP585; red) forcing scenario. Note that percentage changes are here estimated with respect to the first of the two decades between which changes are estimated, in contrast to Fig. 3 in which they are estimated with respect to the baseline period 1850–1950. The results of a least squares regression are shown in black. (k)–(t) Equivalent results for changes in extreme frequency. Ten of the 21 CMIP-6 models are shown, ordered from lowest to highest SAT change, whereas the full ensemble is shown in Figs. S6 and S7 in the online supplemental material.

respectively. To assess the significance of these improvements in comparison with the uncertainty of the scaling relations, we re-estimate the scaling relations in each ensemble member under 1000 bootstrapped replacements of the different time periods used to estimate the changes. Given this methodological uncertainty, we find that 28% and 38% of climate models exhibit significant improvements in the R^2 of their scaling relation at the 10% level.

d. *Discrepancies and constraints on temperature–precipitation scaling rates*

Importantly, robustly identifying temperature–precipitation scaling relations within individual ensemble members facilitates a comparison of their rates across climate models (Fig. 5). On average, we find that the rates at which precipitation extremes scale with SAT within individual models is consistent with that identified between models (cf. Fig. 5 with Fig. 3), but

TABLE 2. An assessment of improvements in the identification of precipitation-temperature scalings when using LFCA rather than averages over 10-, 20-, and 30-yr periods, showing the multimodel average increase in the R^2 of the temperature–precipitation scaling relation when using LFCA rather than temporal averages, as well as the percentage of models with a significant (at the 10% level) increase in R^2 based on bootstrapped estimates of the scaling-relation uncertainties.

	LFCA vs 10-yr av.	LFCA vs 20-yr av.	LFCA vs 30-yr av.
Rx1 R^2 improvement	0.30	0.14	0.11
Percent of models with significant improvement	76	42	28
$R > 99\%$ R^2 improvement	0.38	0.23	0.22
Percent of models with significant improvement	76	57	38

considerable heterogeneity across models is clear. The coefficient of variation of these rates is 23% and 30% for the intensity and frequency of extremes respectively. Moreover, the intermodel differences in scaling rates are statistically significant at the 5% level for 59% and 65% of model pairs, when assessing uncertainty in the estimation of scaling rates as outlined above. These results demonstrate large biases between models, thus diagnosing a source of uncertainty in projections of future precipitation extremes.

To quantify the extent to which these biases contribute to uncertainty in the magnitude of projected precipitation change, we use a simple Monte Carlo simulation procedure that samples from and combines the different changes in global temperatures and different scaling rates projected by different members of the multimodel ensemble. By assessing the distribution of projected changes when uncertainty is propagated from both sources of errors, or when only one source of error is considered while the other is held fixed at its mean value, we can consider the relative extents to which these two sources contribute to overall uncertainty. Results indicate that under a high-greenhouse forcing scenario, biases in the scaling rates of the intensity of extremes contribute equally to uncertainty in projected change as does uncertainty in the projected level of warming (Fig. 6). For the frequency of extremes, biases in scaling rates contribute a considerably larger share to the overall uncertainty. For both metrics, biases in the scaling rates lead to larger right-tailed risks of large increases in precipitation extremes. These results indicate that addressing biases in scaling rates is at least equally important as constraining the climate sensitivity in reducing uncertainties in projections of future precipitation extremes.

While focused model development and physical theory are crucial avenues to limit these biases, comparison with observations may provide a helpful way to evaluate and select climate models that accurately reproduce observed scaling rates. We therefore compare the scaling rates identified within individual members of the CMIP6 ensemble with estimates from observations. Global

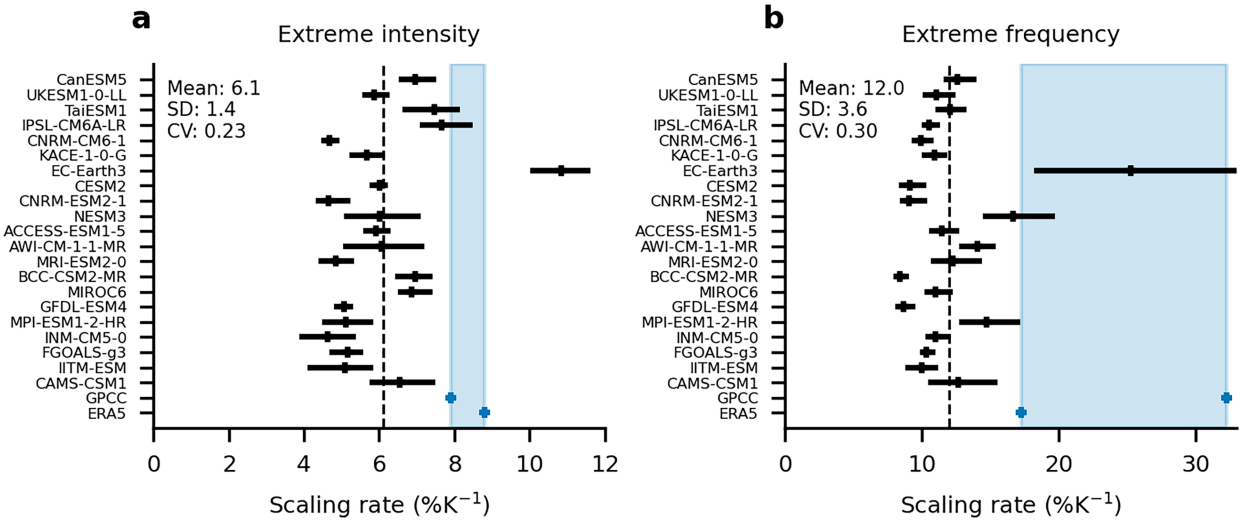

FIG. 5. A comparison of scaling rates between precipitation extremes and global temperatures across models and with estimates from observations: (a) Scaling rates between low-frequency changes in the intensity of daily precipitation extremes estimated for each individual climate model using LFCA, as shown in Fig. 4. The 90% confidence intervals on these scaling rates are obtained via bootstrapped replacements of the intertemporal changes used to estimate the scaling relations and are shown as horizontal bars. The multimodel mean of the precipitation–temperature scaling rates is shown by the vertical black line, with the mean, standard deviation, and coefficient of variation across models indicated in the upper left. Estimates of precipitation–temperature scaling rates from two observational products (ERA5 and GPCC) are made by taking the difference between averages over the first and final 20-yr period of the observational data period and are shown in blue. (b) Equivalent data for the frequency of daily precipitation extremes. In this case, precipitation percentiles are estimated over the whole historical period for the observational data and for a similar period (1985–2015) for the climate models (the scaling rates obtained are largely robust to the period used to estimate percentiles; see Fig. S15 in the online supplemental material).

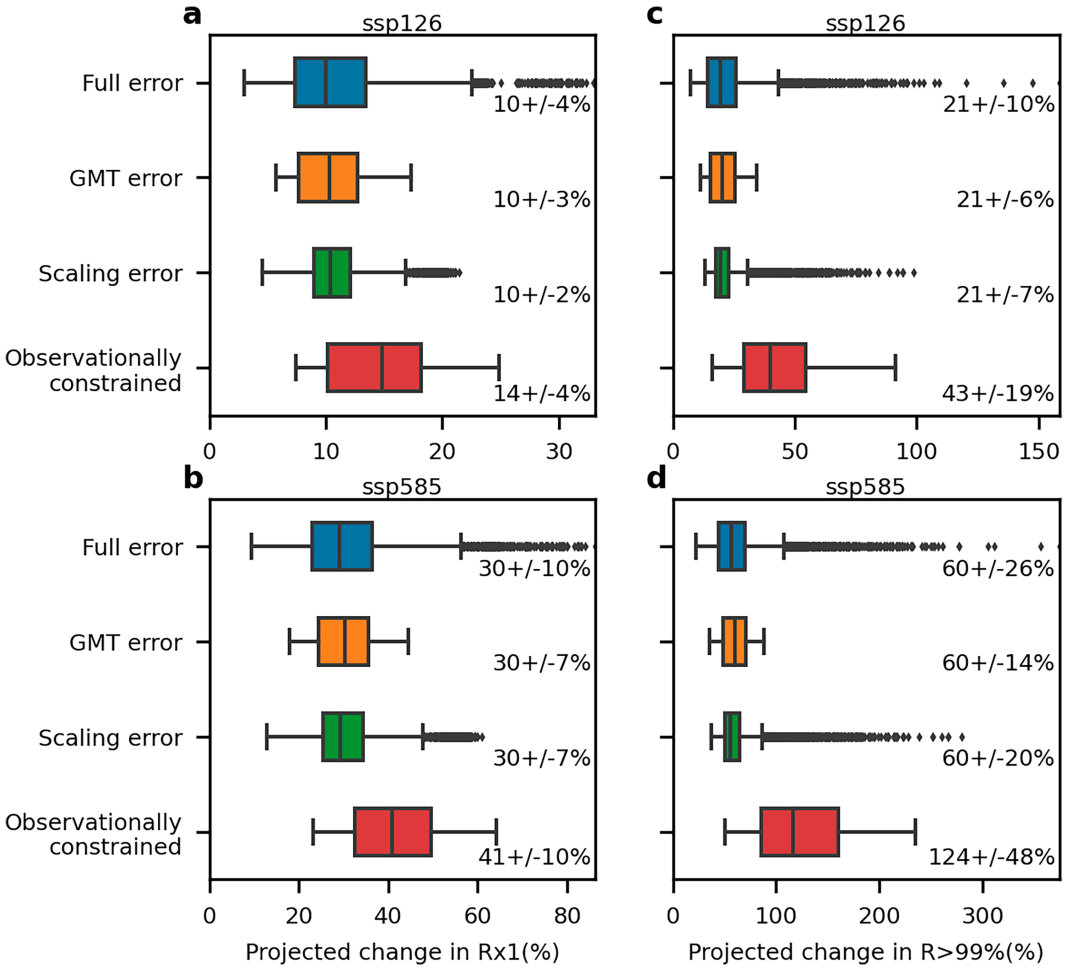

FIG. 6. Quantifying contributions to uncertainty in projected extreme precipitation change and constraining them with observations. Results show the distribution of projected changes in precipitation extremes when conducting a Monte Carlo simulation procedure that samples from the range of global SAT change and temperature–precipitation scaling rates projected by each member of the CMIP6 ensemble. "Full error" indicates simulations in which the uncertainty from both sources is propagated, whereas "GMT error" and "Scaling error" indicate cases in which uncertainty from only one source was combined with the mean-average of the other component. The "observationally constrained" case combines uncertainty from projected SAT change across climate models with the two observational estimates of the temperature–precipitation scaling relations shown in Fig. 5. Results are shown (a),(b) for the intensity of precipitation extremes (RX1) and (c),(d) for the frequency of extremes ($R > 99\%$) for both the weaker [SSP126; in (a) and (c)] and stronger [SSP585; in (b) and (d)] greenhouse forcing scenarios. Inset numbers indicate the mean and standard deviation of the distributions.

historical records of daily precipitation are challenging due to large spatial variability and limited ground-based networks, particularly across parts of the developing world, and we therefore use two observational products. We use data from the ERA5 reanalysis and Global Precipitation Climatology Centre (GPCC), which both focus on daily characteristics. While ERA5 uses land-station data and satellite observations combined with assimilation techniques from weather forecasting models (Hersbach et al. 2020), GPCC relies predominantly on land-station data and interpolation (Schneider et al. 2008). These characteristics enable an assessment of the robustness of observational estimates reflecting uncertainties from both data sources and methodological limitations. Given the relatively shorter time periods available in these observational products (approximately 40 years in each case), we use a simpler methodology to estimate the observed scaling of precipitation extremes with global SAT by taking the difference between averages in precipitation and SAT over the first and final 20 years of available data (note that SAT changes are in both cases derived from ERA5 but adjusted to reflect the time periods of the two precipitation datasets).

This comparison with observations indicates that the historical rates at which the intensification of precipitation extremes have scaled with global SAT exceed those of virtually all the ensemble members in CMIP-6 (observations shown in blue in Fig. 5). The two observational datasets agree closely on the observed scaling rates for the intensity of extremes (approximately 8%–9%), in

which case only 2 of 21 models exhibit scaling rates with uncertainty that falls within this range, and only one outlier that actually exceeds this range (EC-Earth3). Regarding the frequency of precipitation extremes, the two observational datasets exhibit a large range of scaling rates with ERA-5 producing estimates of 17% as compared with 32% in GPCC. These discrepancies indicate considerable challenges in the representation of the frequency of precipitation extremes even in historical observational data, a detailed analysis of which is beyond the scope of the present work. However, despite these uncertainties it is still clear that observed rates exceed those identified in CMIP6. Again, only two models exhibit scaling rates with uncertainties within this range while most models exhibit rates significantly below it. In addition to the assessments of methodological uncertainty already discussed, these results are robust to the exclusion of arid regions with very low initial precipitation values (Fig. S14 in the online supplemental material) as well as the use of different base periods to estimate percentiles in the CMIP6 ensemble (Fig. S15 in the online supplemental material).

Combining these observational estimates of temperature–precipitation scaling rates with projected changes in global temperatures from the CMIP6 ensemble produces an observationally constrained estimate of the distribution of projected precipitation extremes (Fig. 6). These estimates indicate an upward revision of projected changes in precipitation extremes, by a factor of approximately $\frac{1}{3}$ for the intensity of extremes and a doubling for the frequency of extremes. The consistent estimates of scaling rates for the intensity of extremes across the two observational products leads to reductions in the relative error (standard deviation divided by the mean) of the projections, from approximately $\frac{1}{3}$ to $\frac{1}{4}$ in the high forcing scenario. In the case of the frequency of extremes, larger discrepancies between the two observational products mean that relative error in the projections declines by a smaller margin, from 43% to 38%. While a simplified methodology with drawbacks discussed further in the following section, this approach indicates the potential for larger risk from intensifying precipitation extremes than that projected by the CMIP6 ensemble.

4. Discussion and conclusions

This paper uses a pattern-filtering methodology to assess low-frequency changes in the intensity and frequency of daily precipitation extremes in the multimodel ensemble CMIP6. We find that using LFCA improves agreement on the spatial patterns of projected change. Improvements on using 30-yr averages are marginal in the context of strong greenhouse forcing, but under weak greenhouse forcing significant improvements in agreement are found between nearly all model pairs. While this may indicate that LFCA improves the separation of the response of precipitation to human forcing from internal variability, the interpretation of the lowest-frequency change as an anthropogenically forced signal should be explored further. In particular, SMILEs (Kay et al. 2015) may provide a fruitful testing ground to evaluate the efficacy of LFCA in this task and could follow similar studies that demonstrate its strength at detecting forced responses within individual ensemble members from changes in surface temperature (Wills et al. 2020). Such tests could encourage the use of LFCA to detect forced changes in historical observations of precipitation where methods capable of detecting responses from single realizations are necessary.

Our work also sheds light on uncertainties in the magnitude of projected changes in precipitation extremes, in particular with regard to how these changes scale with global surface temperatures. Our finding that changes in precipitation extremes scale with global temperatures across members of multimodel ensembles such as CMIP-6 follows a number of other works (Kharin et al. 2013; Li et al. 2021). However, by using LFCA to robustly identify scaling relations between precipitation extremes and global surface temperature within individual ensemble members, we offer some new insights. First, the fact that such scaling relations are identifiable within individual ensemble members across time and scenarios emphasizes the robustness of such relations. The fact that on average, they follow the Clausius-Clapeyron relation (for the intensity of extremes) emphasizes the fact that thermodynamic drivers of precipitation extremes are dominant determinants of their global change.

Second, our approach reveals strong heterogeneity between different climate models in the rates at which precipitation extremes scale with global temperatures. This diagnoses a source of the uncertainty in projections of future precipitation extremes that we estimate to be equally large as contributions from uncertainty in the climate sensitivity. Our analysis using Monte Carlo simulations to combine global surface temperature changes and global precipitation scaling rates is naturally simplified and does not explicitly consider the driving physical processes. Nevertheless, it is justified by the fact that temperature–precipitation scaling rates appear largely independent of SAT change (Fig. A2 in the appendix), and its findings demonstrate that constraining these scaling rates is equally important as constraining climate sensitivity for limiting uncertainties in future projections. Large efforts have been placed in reducing uncertainties in climate sensitivity (Sherwood et al. 2020). Given the importance of precipitation extremes for impacts, our results may encourage investing similar efforts using a combination of physical theory, observations, and potentially even paleo-climatic evidence (Schmidt et al. 2013; Carmichael et al. 2018) to constrain the sensitivity of secondary climatic characteristics like precipitation extremes to global SAT change.

Third, we find that the rates at which precipitation extremes are projected to scale with global temperatures underestimate those observed historically in nearly all climate models. This is a concerning finding, indicating that there is risk of larger increases in precipitation extremes than those currently projected by models, particularly in terms of their frequency. These results are qualitatively consistent with earlier findings that the observed intensification of precipitation extremes was underestimated by older generations of climate models (Liu et al. 2009; Shiu et al. 2012), although different methods for measuring extremes make our findings difficult to compare quantitatively. While the short observational record may contribute to errors in the quantification of historical scaling rates, the fact that climate models indicate a robust scaling of changes across time and scenarios may add confidence to this assessment. Our combination of observed

scaling rates of precipitation extremes with projected changes in SAT allows an observationally constrained estimate of future changes in extremes, which reduces the uncertainty and increases the magnitude of projected change. Constraining projections with observations is a rapidly growing field, showing recent promise with precipitation (O'Gorman 2012; Zhang and Soden 2019; Shiogama et al. 2022; Chen et al. 2022; Zhang et al. 2022; Thackeray et al. 2022; Schewe and Levermann 2022). In particular, our methodology using temperature–precipitation scaling rates sheds complementary light on a recent method that uses observed historical changes to constrain future projections (Thackeray et al. 2022). While the method of Thackeray et al. (2022) reduces uncertainty in projected precipitation extremes, it does not find observed changes to be larger than those projected by models, and the observational constraint therefore does not increase the projected magnitude of change. This discrepancy to the results of our analysis and previous studies (Liu et al. 2009; Shiu et al. 2012) indicates that separating the different sources of uncertainty in projected precipitation change, namely uncertainty in projected warming as well as the rates at which precipitation extremes scale with global temperatures, is important not only for our qualitative understanding but also for the quantitative results of constraining projections with observations. Nevertheless, our analysis constraining projections with observations is primarily for conceptual purposes, and we recognize that combining our results with a constraint on the climate sensitivity would be necessary to arrive at a fully constrained projection of precipitation extremes as that provided by Thackeray et al. (2022). Indeed, discrepancies between observed and projected historical warming likely underlie the differences in projected precipitation when using these two methods to constrain them. Further analysis in this regard is an interesting avenue for future research but is beyond the scope of this paper.

Acknowledgments. Author Kotz and Wenz received funding from the Volkswagen foundation. Author Lange received funding from the German Research Foundation (DFG, project number 427397136) and from the German Federal Ministry of Education and Research (BMBF, project number 01LP1907A). Author Levermann received funding from the Horizon 2020 Framework Programme of the European Union (grant agreement number 820712).

Data availability statement. Raw CMIP6 data (https://esgf-node.llnl.gov/projects/cmip6/) and code for low-frequency component analysis (https://github.com/rcjwills/lfca) are available online. All other data and code are available from the authors upon request.

APPENDIX

Additional Figures

Scatterplots and correlations across the full ensemble between global surface air temperature change and metrics of the magnitude of forced change in extreme precipitation (Fig. A1), as well as global temperatures and the scaling rate of precipitation extremes (Fig. A2), are shown here.

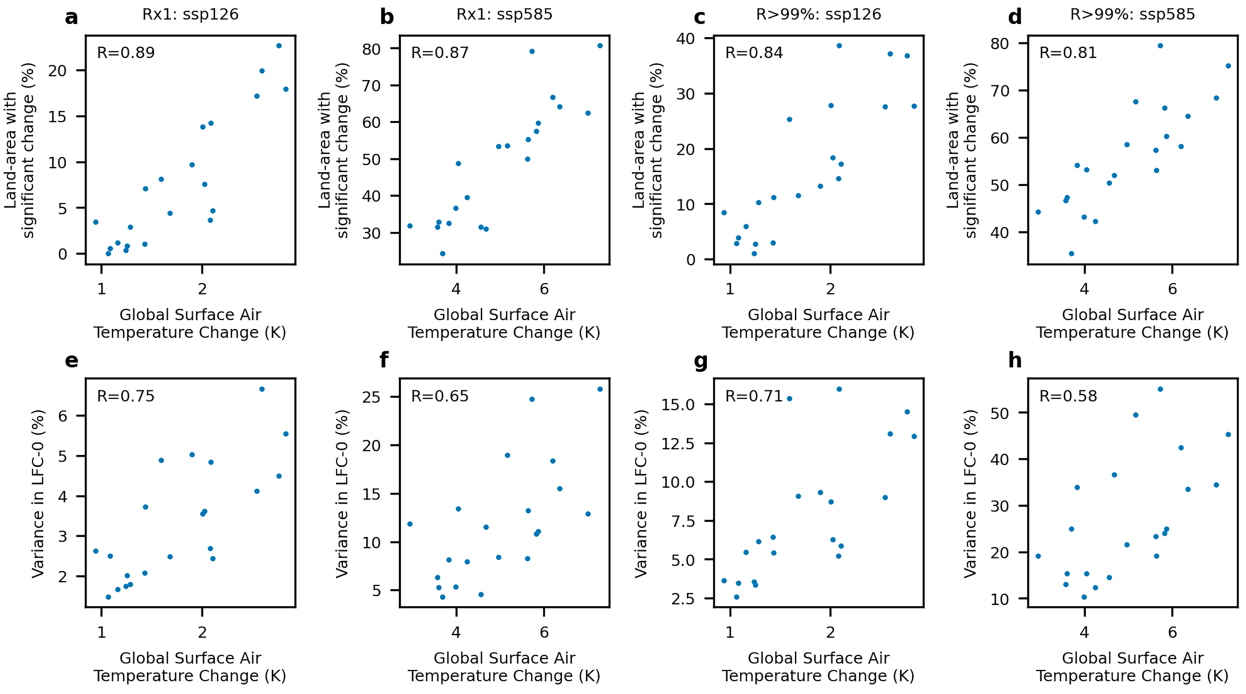

FIG. A1. Scatterplots and Pearson correlation values R for (top) the fraction of land area with significant changes and (bottom) the percentage of spatiotemporal variance explained by the lowest-frequency component against changes in global surface air temperatures. Results are shown for the (a),(b),(e),(f) intensity (Rx1) and (c),(d),(g),(h) frequency ($R > 99\%$) of daily precipitation extremes for the low-greenhouse [SSP126; in (a), (c), (e), and (g)] and high-greenhouse [SSP585; in (b), (d), (f), and (h)] forcing scenarios.

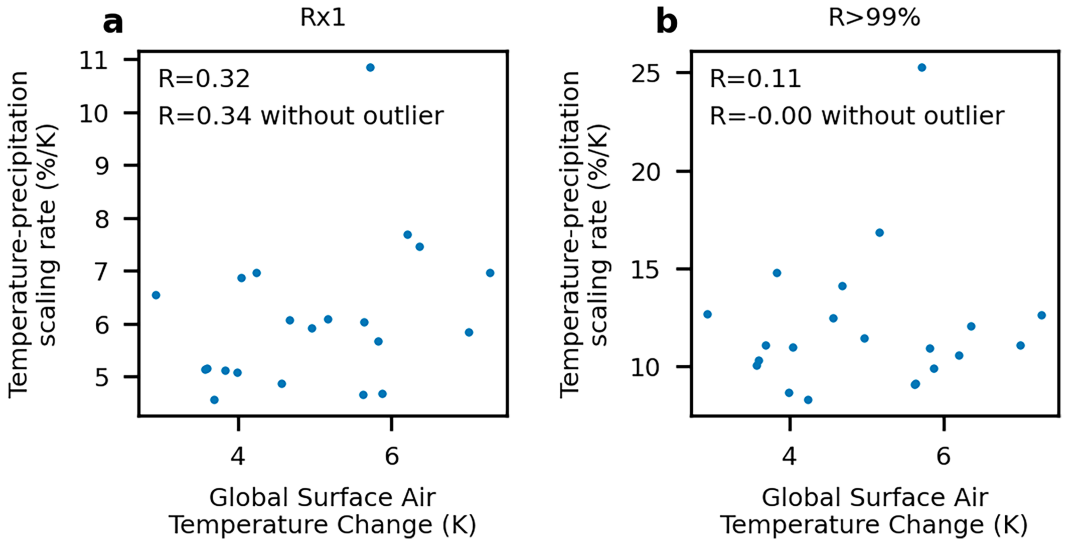

FIG. A2. Scatterplots and Pearson correlation values between the temperature–precipitation scaling rates and global SAT change in the high-forcing scenario (SSP585) identified in CMIP6 climate models.

REFERENCES

Aalbers, E. E., G. Lenderink, E. van Meijgaard, and B. J. J. M. van den Hurk, 2018: Local-scale changes in mean and heavy precipitation in western Europe, climate change or internal variability? *Climate Dyn.*, **50**, 4745–4766, https://doi.org/10.1007/s00382-017-3901-9.

Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 228–232, https://doi.org/10.1038/nature01092.

Blanusa, M. L., C. J. López-Zurita, and S. Rasp, 2023: Internal variability plays a dominant role in global climate projections of temperature and precipitation extremes. *Climate Dyn.*, **61**, 1931–1945, https://doi.org/10.1007/s00382-023-06664-3.

Boulange, J., N. Hanasaki, D. Yamazaki, and Y. Pokhrel, 2021: Role of dams in reducing global flood exposure under climate change. *Nat. Commun.*, **12**, 417, https://doi.org/10.1038/s41467-020-20704-0.

Carmichael, M. J., R. D. Pancost, and D. J. Lunt, 2018: Changes in the occurrence of extreme precipitation events at the Paleocene–Eocene thermal maximum. *Earth Planet. Sci. Lett.*, **501**, 24–36, https://doi.org/10.1016/j.epsl.2018.08.005.

Chen, H., and J. Sun, 2017: Contribution of human influence to increased daily precipitation extremes over China. *Geophys. Res. Lett.*, **44**, 2436–2444, https://doi.org/10.1002/2016GL072439.

Chen, Z., T. Zhou, X. Chen, W. Zhang, L. Zhang, M. Wu, and L. Zou, 2022: Observationally constrained projection of Afro-Asian monsoon precipitation. *Nat. Commun.*, **13**, 2552, https://doi.org/10.1038/s41467-022-30106-z.

Davenport, F. V., M. Burke, and N. S. Diffenbaugh, 2021: Contribution of historical precipitation change to US flood damages. *Proc. Natl. Acad. Sci. USA*, **118**, e2017524118, https://doi.org/10.1073/pnas.2017524118.

Deser, C., A. S. Phillips, M. A. Alexander, and B. V. Smoliak, 2014: Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *J. Climate*, **27**, 2271–2296, https://doi.org/10.1175/JCLI-D-13-00451.1.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Fischer, E. M., and R. Knutti, 2016: Observed heavy precipitation increase confirms theory and early models. *Nat. Climate Change*, **6**, 986–991, https://doi.org/10.1038/nclimate3110.

——, J. Sedláček, E. Hawkins, and R. Knutti, 2014: Models agree on forced response pattern of precipitation and temperature extremes. *Geophys. Res. Lett.*, **41**, 8554–8562, https://doi.org/10.1002/2014GL062018.

Frame, D. J., S. M. Rosier, I. Noy, L. J. Harrington, T. Carey-Smith, S. N. Sparrow, D. A. Stone, and S. M. Dean, 2020: Climate change attribution and the economic costs of extreme weather events: A study on damages from extreme rainfall and drought. *Climatic Change*, **162**, 781–797, https://doi.org/10.1007/s10584-020-02729-y.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hsiang, S. M., M. Burke, and E. Miguel, 2013: Quantifying the influence of climate on human conflict. *Science*, **341**, 1235367, https://doi.org/10.1126/science.1235367.

IPCC, 2021: *Climate Change 2021: The Physical Science Basis*. V. Masson-Delmotte, et al., Eds., Cambridge University Press, 2391 pp., https://doi.org/10.1017/9781009157896.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1.

Kharin, V. V., F. W. Zwiers, X. Zhang, and M. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345–357, https://doi.org/10.1007/s10584-013-0705-8.

King, A. D., and Coauthors, 2015: The timing of anthropogenic emergence in simulated climate extremes. *Environ. Res. Lett.*, **10**, 094015, https://doi.org/10.1088/1748-9326/10/9/094015.

Kirchmeier-Young, M. C., and X. Zhang, 2020: Human influence has intensified extreme precipitation in North America. *Proc.*

Natl. Acad. Sci. USA, **117**, 13 308–13 313, https://doi.org/10.1073/pnas.1921628117.

Knutti, R., and J. Sedláček, 2013: Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Climate Change*, **3**, 369–373, https://doi.org/10.1038/nclimate1716.

Kotz, M., L. Wenz, and A. Levermann, 2021: Footprint of greenhouse forcing in daily temperature variability. *Proc. Natl. Acad. Sci. USA*, **118**, e2103294118, https://doi.org/10.1073/pnas.2103294118.

——, A. Levermann, and L. Wenz, 2022: The effect of rainfall changes on economic production. *Nature*, **601**, 223–227, https://doi.org/10.1038/s41586-021-04283-8.

Lange, S., and Coauthors, 2020: Projecting exposure to extreme climate impact events across six event categories and three spatial scales. *Earth's Future*, **8**, e2020EF001616, https://doi.org/10.1029/2020EF001616.

Li, C., F. Zwiers, X. Zhang, G. Li, Y. Sun, and M. Wehner, 2021: Changes in annual extremes of daily temperature and precipitation in CMIP6 models. *J. Climate*, **34**, 3441–3460, https://doi.org/10.1175/JCLI-D-19-1013.1.

Liang, X.-Z., and Coauthors, 2017: Determining climate effects on US total agricultural productivity. *Proc. Natl. Acad. Sci. USA*, **114**, E2285–E2292, https://doi.org/10.1073/pnas.1615922114.

Liu, S. C., C. Fu, C.-J. Shiu, J.-P. Chen, and F. Wu, 2009: Temperature dependence of global precipitation extremes. *Geophys. Res. Lett.*, **36**, L17702, https://doi.org/10.1029/2009GL040218.

Madakumbura, G. D., C. W. Thackeray, and J. J. Norris, 2021: Anthropogenic influence on extreme precipitation over global land areas seen in multiple observational datasets. *Nat. Commun.*, **12**, 3944, https://doi.org/10.1038/s41467-021-24262-x.

Min, S.-K., X. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, **470**, 378–381, https://doi.org/10.1038/nature09763.

O'Gorman, P. A., 2012: Sensitivity of tropical precipitation extremes to climate change. *Nat. Geosci.*, **5**, 697–700, https://doi.org/10.1038/ngeo1568.

——, and T. Schneider, 2009: The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proc. Natl. Acad. Sci. USA*, **106**, 14 773–14 777, https://doi.org/10.1073/pnas.0907610106.

Pfahl, S., P. A. O'Gorman, and E. M. Fischer, 2017: Understanding the regional pattern of projected future changes in extreme precipitation. *Nat. Climate Change*, **7**, 423–427, https://doi.org/10.1038/nclimate3287.

Santer, B. D., K. E. Taylor, T. M. Wigley, J. E. Penner, P. D. Jones, and U. Cubasch, 1995: Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dyn.*, **12**, 77–100, https://doi.org/10.1007/BF00223722.

——, and Coauthors, 2007: Identification of human-induced changes in atmospheric moisture content. *Proc. Natl. Acad. Sci. USA*, **104**, 15 248–15 253, https://doi.org/10.1073/pnas.0702872104.

Schewe, J., and A. Levermann, 2022: Sahel rainfall projections constrained by past sensitivity to global warming. *Geophys. Res. Lett.*, **49**, e2022GL098286, https://doi.org/10.1029/2022GL098286.

Schmidt, G. A., and Coauthors, 2013: Using paleo-climate comparisons to constrain future projections in CMIP5. *Climate Past Discuss.*, **9**, 775–835, https://doi.org/10.5194/cpd-9-775-2013.

Schneider, U., T. Fuchs, A. Meyer-Christoffer, and B. Rudolf, 2008: Global precipitation analysis products of the GPCC. Global Precipitation Climatology Centre (GPCC), DWD Internet Publikation Rep. 112, 17 pp., https://opendata.dwd.de/climate_environment/GPCC/PDF/GPCC_intro_products_lastversion.pdf.

Schumacher, R. S., 2017: Heavy rainfall and flash flooding. *Oxford Research Encyclopedia of Natural Hazard Science*, https://doi.org/10.1093/acrefore/9780199389407.013.132.

Shepherd, T. G., 2014: Atmospheric circulation as a source of uncertainty in climate change projections. *Nat. Geosci.*, **7**, 703–708, https://doi.org/10.1038/ngeo2253.

Sherwood, S. C., and Coauthors, 2020: An assessment of Earth's climate sensitivity using multiple lines of evidence. *Rev. Geophys.*, **58**, e2019RG000678, https://doi.org/10.1029/2019RG000678.

Shiogama, H., M. Watanabe, H. Kim, and N. Hirota, 2022: Emergent constraints on future precipitation changes. *Nature*, **602**, 612–616, https://doi.org/10.1038/s41586-021-04310-8.

Shiu, C.-J., S. C. Liu, C. Fu, A. Dai, and Y. Sun, 2012: How much do precipitation extremes change in a warming climate? *Geophys. Res. Lett.*, **39**, L17707, https://doi.org/10.1029/2012GL052762.

Thackeray, C. W., A. Hall, J. Norris, and D. Chen, 2022: Constraining the increased frequency of global precipitation extremes under warming. *Nat. Climate Change*, **12**, 441–448, https://doi.org/10.1038/s41558-022-01329-1.

Thiery, W., and Coauthors, 2021: Intergenerational inequities in exposure to climate extremes. *Science*, **374**, 158–160, https://doi.org/10.1126/science.abi7339.

Thomas, B. F., and J. S. Famiglietti, 2019: Identifying climate-induced groundwater depletion in GRACE observations. *Sci. Rep.*, **9**, 4124, https://doi.org/10.1038/s41598-019-40155-y.

von Uexkull, N., M. Croicu, H. Fjelde, and H. Buhaug, 2016: Civil conflict sensitivity to growing-season drought. *Proc. Natl. Acad. Sci. USA*, **113**, 12 391–12 396, https://doi.org/10.1073/pnas.1607542113.

Warszawski, L., K. Frieler, V. Huber, F. Piontek, O. Serdeczny, and J. Schewe, 2014: The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework. *Proc. Natl. Acad. Sci. USA*, **111**, 3228–3232, https://doi.org/10.1073/pnas.1312330110.

Willner, S. N., A. Levermann, F. Zhao, and K. Frieler, 2018: Adaptation required to preserve future high-end river flood risk at present levels. *Sci. Adv.*, **4**, eaao1914, https://doi.org/10.1126/sciadv.aao1914.

Wills, R. C., T. Schneider, J. M. Wallace, D. S. Battisti, and D. L. Hartmann, 2018: Disentangling global warming, multidecadal variability, and El Niño in Pacific temperatures. *Geophys. Res. Lett.*, **45**, 2487–2496, https://doi.org/10.1002/2017GL076327.

——, D. S. Battisti, K. C. Armour, T. Schneider, and C. Deser, 2020: Pattern recognition methods to separate forced responses from internal variability in climate model ensembles and observations. *J. Climate*, **33**, 8693–8719, https://doi.org/10.1175/JCLI-D-19-0855.1.

Wood, R., and R. Ludwig, 2020: Analyzing internal variability and forced response of sub-daily and daily extreme precipitation over Europe. *Geophys. Res. Lett.*, **47**, e2020GL089300, https://doi.org/10.1029/2020GL089300.

Zhang, B., and B. J. Soden, 2019: Constraining climate model projections of regional precipitation change. *Geophys. Res. Lett.*, **46**, 10 522–10 531, https://doi.org/10.1029/2019GL083926.

Zhang, W., K. Furtado, T. Zhou, P. Wu, and X. Chen, 2022: Constraining extreme precipitation projections using past

precipitation variability. *Nat. Commun.*, **13**, 6319, https://doi.org/10.1038/s41467-022-34006-0.

Zhang, X., H. Wan, F. W. Zwiers, G. C. Hegerl, and S.-K. Min, 2013: Attributing intensification of precipitation extremes to human influence. *Geophys. Res. Lett.*, **40**, 5252–5257, https://doi.org/10.1002/grl.51010.

Ziegler, A. D., J. Sheffield, E. P. Maurer, B. Nijssen, E. F. Wood, and D. P. Lettenmaier, 2003: Detection of intensification in global- and continental-scale hydrological cycles: Temporal scale of evaluation. *J. Climate*, **16**, 535–547, https://doi.org/10.1175/1520-0442(2003)016<0535:DOIIGA>2.0.CO;2.