



A framework for detection and attribution of regional precipitation change: Application to the United States historical record

Mark D. Risser¹ · William D. Collins^{1,2} · Michael F. Wehner¹ · Travis A. O'Brien^{1,3} · Christopher J. Paciorek⁴ · John P. O'Brien^{1,5} · Christina M. Patricola-DiRosario^{1,6} · Huanping Huang¹ · Paul A. Ullrich⁷ · Burlen Loring¹

Received: 4 August 2021 / Accepted: 24 April 2022 / Published online: 6 June 2022
© The Author(s) 2022, modified publication 2026

Abstract

Despite the emerging influence of anthropogenic climate change on the global water cycle, at regional scales the combination of observational uncertainty, large internal variability, and modeling uncertainty undermine robust statements regarding the human influence on precipitation. Here, we use output from global climate models in a perfect-data sense to develop a framework for conducting regional detection and attribution (D&A) for precipitation, starting with the contiguous United States (CONUS) where observational uncertainty is lower than in other regions. Our unified approach can simultaneously detect systematic trends in mean and extreme precipitation, attribute trends to anthropogenic forcings, compute the effects of forcings as a function of time, and map the effects of individual forcings. Model output is used to conduct a set of tests that yield a parsimonious representation for characterizing seasonal precipitation over the CONUS for the historical record (1900 to present day), which ensures our D&A is insensitive to structural uncertainty. Our framework is developed using synthetic data in a Pearl-causal perspective wherein causality can be identified using intervention-based simulations. While the hypothesis-based framework and accompanying generalized D&A formula we develop should be widely applicable, we include a strong caution that the hypothesis-guided simplification of the formula for the historical climatic record of CONUS as described in this paper will likely fail to hold in other geographic regions and under future warming.

Keywords Extreme precipitation · Natural variability · Local impacts · Anthropogenic aerosols · CMIP6 · DAMIP · Pearl causality

1 Introduction

Even though anthropogenic influence has been identified for many variables, including surface air temperature (Hegerl et al. 1997; Tett et al. 1999), sea level pressure (Gillett et al. 2003), tropopause height (Santer et al. 2003), free atmospheric temperature (Jones et al. 2003), and ocean heat content (Barnett et al. 2005), robust and conclusive statements regarding human influence on regional precipitation remain elusive. Changes at the global scale are difficult to detect due to offsetting increases and decreases in different parts of the globe, but Hartmann et al. (2013) find it *likely* that human influence has affected the global water cycle since the middle of the 20th century. There is a detectable and attributable human influence on zonal land-averages of mean precipitation (Zhang et al. 2007; Sarojini et al. 2012; Noake et al. 2012), with increases in precipitation for the Northern Hemisphere (NH) midlatitudes, Southern Hemisphere (SH) subtropics, and SH deep tropics but decreases for the NH

✉ Mark D. Risser
mdrisser@lbl.gov

¹ Lawrence Berkeley National Laboratory, Berkeley, CA, USA

² Department of Earth and Planetary Science, University of California, Berkeley, CA, USA

³ Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, USA

⁴ Department of Statistics, University of California, Berkeley, CA, USA

⁵ National Center for Atmospheric Research, Boulder, CO, USA

⁶ Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA, USA

⁷ Department of Land, Air, and Water Resources, University of California, Davis, CA, USA

subtropics. Detection of changes in mean precipitation for individual grid boxes has recently been attempted (Knutson and Zeng 2018), although the anthropogenic signal is relatively weak at these scales. For extreme precipitation, attribution statements have generally only been attempted for continental-scale averages (Min et al. 2011; Zhang et al. 2013; Paik et al. 2020; Dong et al. 2021), with human-induced increases in greenhouse gases leading to an intensification of heavy precipitation events for much of the NH (with a few exceptions, e.g., Groisman et al. 2005). Two recent studies (Kirchmeier-Young and Zhang 2020; Huang et al. 2021) detect changes in extreme precipitation at sub-continental scales, but the changes are only attributable for a subset of the models considered. All the aforementioned studies rely on global climate models to simulate the anthropogenically-forced signal and natural variability in precipitation and then project observations onto these patterns to assess their agreement with the climate models.

Unfortunately, continental-scale statements about the human influence on precipitation do not provide the spatially resolved conclusions needed to inform assessments of how much climate change is happening locally. Such information, including both the magnitude of the change and its nature (i.e., is precipitation increasing or decreasing), are critical for resource managers who are trying to plan for the impacts of climate change. Furthermore, since precipitation is an episodic variable, in some places magnitudes and frequencies can have conflicting signs resulting in confusing trends in total precipitation (see, e.g., Polade et al. 2014, 2017; Gershunov et al. 2019). In these cases, it is important to understand regional (sub-continental) changes as well as the spatial patterns of change. As an example, when considering the contiguous United States (CONUS), even though there are well-documented 50- to 100-year trends in seasonal mean and extreme precipitation (Kunkel 2003; Easterling et al. 2017; Risser et al. 2019a), existing model-based detection and attribution (D&A) studies that explore local changes in precipitation find essentially no detectable signal. For example, over 1901–2010, Knutson and Zeng (2018) find that the observations are consistent with natural-forcings runs for 58% of the global grid boxes with sufficient observational coverage (this percentage increases to 71% for 1951–2010 and 78% for 1981–2010). Knutson and Zeng (2018) further show that approximately 85% of the CONUS grid cells have non-detectable trends and the other 15% cannot be attributed to human influence. Guo et al. (2019) use a novel “dynamical adjustment” technique to remove the influence of unforced atmospheric circulation variability on wintertime Northern Hemisphere observations, but again only find significant trends for a small fraction of the CONUS. No such grid-cell level statements are available for extreme precipitation.

The lack of confidence in regional attribution statements regarding precipitation is the result of three factors: observational uncertainty, large internal variability, and modeling uncertainty (Sarojini et al. 2016). Globally, the limited existence of high-quality, century-length records is the primary driver of observational uncertainty: Menne et al. (2012) show that dense, long-term measurements of daily precipitation are limited to CONUS, southern Canada, Western Europe, and parts of Australia. Measurement uncertainty translates directly into attribution uncertainty, e.g., attribution statements for zonal averages are sensitive to the observational data set used (Noake et al. 2012) and model-simulated changes are significantly diminished when considering only grid cells with sufficient observational coverage (Sarojini et al. 2012). Otherwise, the magnitude of internal variability is confounded with modeling uncertainty, particularly for regional scales. At these scales, human-induced change is the result of complicated interactions between natural variability and anthropogenic forcing, such that the structural uncertainty of models remains a significant obstacle to attributing regional changes (Sarojini et al. 2016). Furthermore, the aforementioned D&A studies rely on relatively small ensembles, which likely yield insufficient sampling of the internal variability (DelSole et al. 2019). Model uncertainty is exacerbated for precipitation attribution since a treatment of anthropogenic aerosols is critical (Hegerl et al. 2015). In this case, models include two sources of uncertainty: the representation of aerosols internally (e.g., prescribed versus parameterized) as well as the precipitation response to aerosols. For all of these reasons, the comparison of uncertain observations (at least for much of the globe) with uncertain models makes robust attribution statements for regional precipitation extremely difficult.

Fortunately, Sarojini et al. (2016) suggest a path forward for regional detection and attribution of precipitation. First, they argue that new methods can “facilitate the identification of significant changes in precipitation even in the presence of substantial modeling and observational uncertainty” such that we can make “faster progress ...than would be possible by simply waiting for models or observations to improve or by simply waiting for the signal of climate change to strengthen sufficiently to emerge from the noise of internal variability.” Importantly, new methods should analyze changes in the processes that control natural variability (i.e., the noise) by explicitly modeling the drivers of precipitation variability. Like Hegerl et al. (2015), they emphasize the importance of accounting for aerosols, highlighting model-based studies that have identified the counteracting influence of anthropogenic aerosols on expected increases from greenhouse gas emissions (Wu et al. 2013; Wilcox et al. 2013). Finally, they illuminate the importance of disentangling the complex causes of regional changes in precipitation, including the evaluation of a diverse set of external drivers that may influence the precipitation response as well as possible interactions between external forcing and natural variability.

In light of these suggestions, in this paper we propose a novel framework for conducting regional D&A for seasonal mean and extreme precipitation wherein we detect systematic trends, attribute trends to specific forcings, compute the effects of forcings as a function of time, and map the spatial patterns of each forcings' effect. Our initial focus is on the CONUS, since the influence of observational uncertainty is minimized relative to other global land regions. Our approach to detection and attribution provides two novel aspects relative to existing methods. First, we develop a parsimonious formula that clearly specifies how both natural drivers of climate variability and anthropogenic forcing influence precipitation in a single framework. The implication is that we ensure any uncertainty is in the response of precipitation and not in climate models' representation of the drivers and forcings. Second, while traditional D&A studies are subject to climate models' uncertainties regarding short-lived climate forcings (SLCFs) and their influence on precipitation, our philosophy uses model output in a perfect data sense. Importantly, this means that we can turn an apparent limitation into an opportunity: by evaluating the performance of our methodology across the broad diversity of responses to SLCFs in the CMIP6 multi-model ensemble, we ensure our D&A is insensitive to structural uncertainty. In other words, we can evaluate our methodology regardless of the quality of an individual model's characterization of precipitation and its response to external forcing.

Additionally, our proposed framework can be rigorously tested to ensure that it is applied appropriately to an observational data set. First, we derive a physically defensible characterization of the influence of anthropogenic warming that captures theoretical global scaling relationships for mean and extreme precipitation, such that the estimation of scaling relationships is effective in both single-forcing experiments and historical runs with all forcings. We furthermore identify individual external forcings that are important for seasonal precipitation over the CONUS in the historical record, as well as those which can be safely ignored regarding their influence on secular trends. Given the importance of anthropogenic aerosols, we use the broad diversity of responses across the CMIP6 multi-model ensemble to arrive at a method for characterizing aerosols that is insensitive to structural uncertainty. Our framework also assesses the influence of external forcings on the relationships between known large-scale modes of climate variability and precipitation, and can separate the counteracting influence of well-mixed greenhouse gases and anthropogenic aerosols. Finally, we find that the signal-to-noise ratio for secular trends in precipitation is unaffected by warming, i.e., the challenge of detecting trends is not complicated by increased warming to the global climate system. While the analysis in this paper utilizes a Pearl-causal framework (Hannart et al. 2015) for D&A (i.e., using climate models that involve an

intervention), we plan to apply the same formula to in situ measurements from the historical record in a Granger-causal framework (Granger 1969).

The paper proceeds as follows: in Sect. 2, we outline a general formula for modeling a time series of annual mean or extreme precipitation as well as a set of testable hypotheses for how the general formula can be simplified for regional D&A over the historical period. In Sect. 3, we describe the climate model data sets, forcing data sets, and modes of natural climate variability used to test the hypotheses proposed in Sect. 2. Section 4 presents the main results of the paper, wherein we systematically examine each hypothesis and draw conclusions for developing an appropriate formula for detection and attribution for precipitation over the CONUS in the historical record. Section 5 concludes the paper. A summary figure (Fig. 1) and table (Table 3) provide high-level overviews of the analysis and subsequent conclusions made.

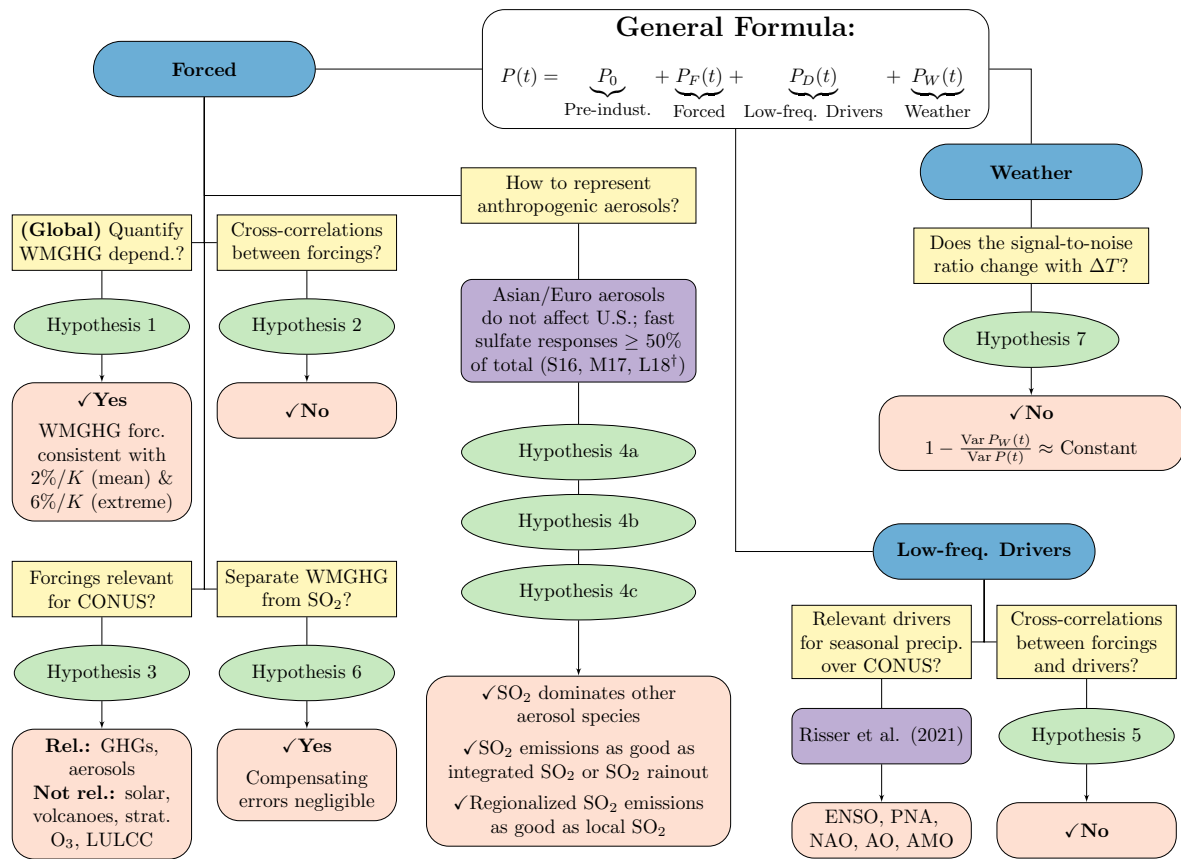
2 The D&A formula

We first outline a general framework for modeling a time series of precipitation, denoted $\{P(t) : t = 1, \dots, T\}$, where the temporal units t refer to a year and $P(t)$ is either the seasonal mean precipitation rate (mm day^{-1}) or seasonal maximum daily precipitation total (mm , often referred to as Rx1Day). To account for seasonal differences in the characteristics of precipitation, the following framework will be applied separately to each three-month season (DJF, MAM, JJA, and SON). Suppressed in the notation throughout is geographic location; this is (for now) generic but will generally refer to geospatial locations (e.g., model grid cells or weather station locations). The formula is as follows: for time $t = 1, \dots, T$, we define

$$P(t) = \underbrace{P_0}_{\text{Pre-ind.}} + \underbrace{P_F(t)}_{\text{Forced}} + \underbrace{P_I(t)}_{\text{Internal}} \quad (1)$$

where the time series is decomposed into a pre-industrial term P_0 that represents an overall average uninfluenced by external forcing, a forced component $P_F(\cdot)$, and an internal variability component $P_I(\cdot)$. The forced component of this time series is further decomposed as

$$P_F(t) = \underbrace{\sum_{\text{Forcing } i} \beta_i F_i(t, \tau_i)}_{\text{Forced Response}} + \underbrace{\sum_{\text{Forcing } i} \sum_{\text{Forcing } j \neq i} \frac{\partial [\beta_i F_i(t, \tau_i)]}{\partial F_j} F_j(t, \tau_j)}_{\text{Cross-correlation}} + \dots, \quad (2)$$



† S16 = Samset et al. [2016], M17 = Myhre et al. [2017], L18 = Liu et al. [2018]

Fig. 1 Flow chart schematic outlining the various hypotheses considered and the conclusions obtained in Sect. 4 regarding the general formula proposed in Eq. 1 for modeling an annual time series of precipitation, denoted $P(t)$. The blue ovals summarize the main components of the general formula (forced, driven, and weather variabil-

ity). The yellow boxes correspond to the general research questions explored in the paper, purple boxes indicate relevant conclusions from existing literature, and the green circles denote the methods for making conclusions about each research question. The red boxes contain the resulting conclusions

including a sum of forced responses, anthropogenic or otherwise, and a set of cross-correlation terms between pairs of forcings. (Note: the cross-correlation term is another way of writing a statistical interaction.) The forced term $P_F(\cdot)$ could also involve higher-order nonlinear terms (e.g., quadratic) or interactions (e.g., terms involving three or more forcings). The terms $F_i(t, \tau_i)$ are defined by Eq. 14 (Appendix A) and represent forcings modified by the lagged response of the climate system on characteristic timescales τ_i determined by the thermal inertia of the ocean.

The internal variability component $P_I(\cdot)$ is further decomposed as

$$P_I(t) = \underbrace{P_D(t)}_{\text{Low-freq. Drivers}} + \underbrace{P_W(t)}_{\text{Weather}}, \quad (3)$$

where $P_D(\cdot)$ is a “driven” term that describes year-to-year variability due to known modes of large-scale ocean (e.g., the El Niño Southern Oscillation or ENSO) or atmospheric (e.g., the Pacific North American pattern) drivers, and $P_W(\cdot)$ is a term associated with short-term weather variability. We assume that the driven component can be represented as

$$P_D(t) = \sum_{d=\text{ELI, AO, NAO, PNA, AMO}} \beta_d d(t) + \sum_{\text{Driver } d} \sum_{\text{Forcing } j} \frac{\partial [\beta_d d(t)]}{\partial F_j} F_j(t, \tau_j) + \dots, \quad (4)$$

where the relevant drivers for seasonal precipitation over CONUS are posited to be the ENSO Longitude Index (ELI), the Arctic Oscillation (AO), the North Atlantic Oscillation (NAO), the Pacific-North American pattern (PNA), and the Atlantic Multidecadal Oscillation (AMO), as explored in Risser et al. (2021) (see Sect. 3.3 for more information). Note that, as described in Risser et al. (2021), the AO is excluded from the DJF analysis due to its strong coupling (and high correlation) with the NAO in this season. The driven component also includes cross-correlation terms (which again can be thought of as statistical interactions) between the drivers and the external forcing agents, wherein (for example) the relationship between seasonal precipitation and ENSO might be influenced by the well-mixed greenhouse gas (WMGHG) forcing. Equation 4 could also include higher-order terms such as interactions between modes of variability or stochastic modulation of teleconnections (see, e.g., Gershunov et al. 2001).

Finally, the internal variability component includes an error term $P_W(\cdot)$ that represents residual variability in the time series due to weather. We suppose that this term is, on average, zero, but has a variance described by

$$\text{Var } P_W(t) = \sigma^2(t) \approx V_0 \times \exp\{V_1 t\}, \quad (5)$$

i.e., the magnitude of the weather variability depends on the time t and can be approximately modeled as the product of a “baseline” or pre-industrial variance V_0 and a time-varying quantity $\exp\{V_1 t\}$ where $1/|V_1|$ is a characteristic timescale for the response of the variance. This framework allows the weather variability to change (either increase or decrease) over the historical record. Comparison of the all-hist (all historical forcing) and nat-hist (natural-only forcing) ensembles from C20C+ indicates that precipitation variability increases in the presence of anthropogenic forcing (Pendergrass et al. 2017; O’Brien 2019), so our working assumption is that $V_1 > 0$. Since the time units t refer to annual quantities, we assume that all the temporal autocorrelation in the time series is fully captured by $P_F(\cdot)$ and $P_D(\cdot)$, i.e., the residual weather variability terms for each year are temporally independent. For the seasonal mean analysis, we assume that the weather variability follows a normal distribution (with mean zero and variance $\sigma^2(t)$), while for the seasonal Rx1Day analysis, we assume that the weather variability follows a Generalized Extreme Value distribution (centered on zero such that the variability is described by $\sigma^2(t)$). Note that these assumptions on the distribution of weather variability refer to seasonal aggregates (mean and maximum) of daily precipitation and not the total distribution of daily precipitation.

The general formula defined by Eqs. 2–5 contains many degrees of freedom, such that applying the full equation to an observed time series will likely result in over-fitting. We

therefore propose a series of tests or falsifiable hypotheses that will allow us to simplify Eqs. 2–5 for use in the special case of seasonal precipitation over the United States. Two notes: first, we set out to investigate a variety of complicated factors using the analyses described in Sect. 4 and found that we could simplify Eq. 1 (the flow chart diagram in Fig. 1 summarizes the various analyses considered); the hypotheses are organized in terms of these discoveries. Second, it is important to emphasize that our analyses focus on the contiguous United States over the historical record (1900 to present), hence all conclusions regarding simplifications to the general formula apply only to a specific geographic region and time period. While we do not claim that any of these hypotheses are satisfied in other parts of the globe or under future warming to the global climate system, we assert that our hypothesis-driven framework and the accompanying general D&A formula are broadly applicable to other regions and timeframes. To summarize the wide range of science questions considered for each component of the general formula in Eq. 1, we have developed the flow chart shown in Fig. 1.

Our first test addresses whether we can correctly identify the influence of WMGHG forcing on precipitation using a linear regression framework:

Hypothesis 1 (H1). Globally, the coefficient for WMGHG forcing (β WMGHG) in Eq. 2 is consistent with 2%/K for mean precipitation and 6%/K for extreme precipitation.

While the focus of this paper is CONUS, we do not have any *a priori* expectations about the magnitude of the WMGHG signature regionally (e.g., over CONUS) and seasonally. We do know what to expect in terms of global-average annual temperature changes, and hence our first test is conducted at the global scale. Next, again looking at precipitation globally, we test whether there are meaningful cross-correlation terms between external forcings:

Hypothesis 2 (H2). Globally, over the historical period, all cross-correlation terms in the forced component in Eq. 2 are negligible, i.e.,

$$\sum_{\text{Forcing } i} \sum_{\text{Forcing } j \neq i} \gamma_{ij} \frac{\partial [\beta_i F_i(t, \tau_i)]}{\partial F_j} F_j(t, \tau_j) + \dots \approx 0$$

As mentioned previously, there could also be nonlinear components in the forced term given in Eq. 2; however, for the current analysis we assume that these terms are negligible. The climate state over the historical period has not changed enough to warrant strongly nonlinear behavior (e.g., step changes due to the loss of sea ice), and numerous climate model studies have demonstrated the additivity of responses to forcing agents (e.g. Kirkevåg et al. (2008); Shiogama et al. (2013); Marvel et al. (2015)). Shifting our focus to CONUS, we next test for the presence of meaningful trends in seasonal precipitation due to individual forcing agents:

Hypothesis 3 (H3). For CONUS and over the historical record, the effects of individual forcings are non-negligible for secular trends in precipitation.

In other words, this hypothesis seeks to determine the set of forcings that need to be included in the forced component $P_F(\cdot)$ of seasonal precipitation over CONUS and over the historical period. Given that one of these non-negligible forcings is tropospheric aerosols (see Sect. 4.2), we next explore a set of tests to assess structural uncertainties in our representation of anthropogenic aerosols:

Hypothesis 4a (H4a). For CONUS and over the historical record, the effect of SO_2 on seasonal precipitation dominates the influence of other anthropogenic aerosol species (black and organic carbonaceous aerosols and ammonia).

Hypothesis 4b (H4b). For CONUS and over the historical record, SO_2 emissions correlate with changes in precipitation as well as deposition of SO_2 by rainfall and column integrated SO_2 mass.

Hypothesis 4c (H4c). For CONUS and over the historical record, a regionalized time series of SO_2 emissions yield the same result as using either local or CONUS-wide estimates of SO_2 emissions.

We next test whether individual forcing agents influence the relationships between the modes of natural variability and seasonal precipitation:

Hypothesis 5 (H5). For CONUS and over the historical record, all higher-order and interaction terms between the climate drivers and external forcing agents in Eq. 4 are negligible, i.e.

$$\sum_{\text{Driver } d} \sum_{\text{Forcing } j} \gamma'_{dj} \frac{\partial [\beta_d d(t)]}{\partial F_j} F_j(t, \tau_j) + \dots \approx 0.$$

Because externally forced trends in seasonal precipitation over the historical record are small, we next test whether we can distinguish the influence of WMGHGs from the influence of anthropogenic aerosols:

Hypothesis 6 (H6). For CONUS and over the historical record, the compensating errors in the effects of WMGHG and anthropogenic aerosol forcing are negligible, i.e., the estimated effects of WMGHGs and aerosols are (1) unbiased and (2) do not involve any compensating trade-offs.

Finally, we conduct a test of the influence of warming on the signal-to-noise ratio in seasonal precipitation:

Hypothesis 7 (H7). For CONUS and over the historical record, the signal-to-noise ratio of seasonal precipitation is a constant over time, i.e.,

$$1 - \frac{\text{Var } P_W(t)}{\text{Var } P(t)} \approx 1 - \frac{\text{Var } P_W(0)}{\text{Var } P(0)} \approx \text{Constant}.$$

Each of these hypotheses will be evaluated using a large set of climate model output (most from the CMIP6 multi-model ensemble) in a perfect data sense to test and guide

fits that will eventually be applied to observations and also ensure our D&A is insensitive to structural uncertainty. The result of our explorations (described in the remainder of this paper) is that we can simplify the general formula defined by Eqs. 2–5 as follows:

$$\begin{aligned} P(t) &= P_0 + P_F(t) + P_D(t) + P_W(t), \text{ where} \\ P_F(t) &\approx \beta \text{ WMGHG } F \text{ WMGHG}(t, \tau \text{ WMGHG}) \\ &\quad + \beta \text{ SO}_2 F \text{ SO}_2(t, \tau \text{ SO}_2) \\ P_D(t) &\approx \sum_{d=\text{ELI, AO, NAO, PNA, AMO}} \beta_d d(t), \\ 1 - \frac{\text{Var } P_W(t)}{\text{Var } P(t)} &\approx \text{Constant}. \end{aligned} \quad (6)$$

The time lag $\tau \text{ WMGHG}$ is estimated in Appendix A. For reasons discussed in Sect. 4.3, we assume that the response of precipitation to SO_2 is dominated by fast processes and therefore set $\tau \text{ SO}_2 \approx 0$.

3 Data sources

3.1 Climate model output

The hypotheses proposed in Sect. 1 require a significant amount of climate model output to test each component in a perfect data sense. The various experiments used are described in the following sections. Across the different models, the primary focus is on mean and extreme precipitation, defined as grid-cell average precipitation rate and grid-cell maximum daily precipitation (denoted Rx1Day). Mean and maximum daily quantities are calculated from the GCM-output precipitation rate (where we convert the raw output in $\text{kg m}^{-2} \text{s}^{-1}$ to mm day^{-1}). Hypotheses H1 and H2 involve annual precipitation (in light of theoretical expectations) while all other tests involve seasonal quantities.

CMIP6 historical, DAMIP, LUMIP experiments. First, we utilize output from the Detection and Attribution Model Intercomparison Project (DAMIP; Gillett et al. 2016), which seeks to enable estimation of the influence of various anthropogenic and natural forcings on the global climate system by providing sets of single-forcing or single-group forcing experiments; we use the hist-GHG (greenhouse gases only), hist-aer (aerosol only), hist-nat (solar and volcanoes only), and hist-stratO3 (stratospheric ozone only) simulations. Next, we use output from the Land Use Model Intercomparison Project (LUMIP; Lawrence et al. 2016) which seeks to quantify the influence of anthropogenic land-use and land-cover change (LULCC) on climate by way of the “hist-noLu” simulation, which takes the historical runs and turns off the LULCC. For comparison, we also examine the

CMIP6 historical simulations (Eyring et al. 2016) that span 1850 to 2014 and use observationally based, evolving, and externally imposed forcings, specifically solar variability, volcanic aerosols, and changes in atmospheric composition caused by human activities. The Global Circulation Models (GCMs) from DAMIP, LUMIP, and CMIP6-historical used here are shown in Table F.1 in Supplementary material, along with the number of ensemble members and number of years available.

CMIP6 DECK experiments: *piControl*, *1pctCO2*, and *abrupt4xCO2*. The various historical experiments only consider “moderate” warming to the global climate system, in that they simulate pre-industrial (1850) to present-day conditions. To further assess the influence of WMGHGs specifically under greater levels of warming, we utilize three of the CMIP6 Diagnostic, Evaluation, and Characterization of Klima (DECK) experiments, namely the pre-industrial control simulation (*piControl*); a simulation forced by an abrupt quadrupling of CO_2 (*abrupt4xCO2*); and a simulation forced by a $1\% \text{ yr}^{-1} \text{ CO}_2$ increase (*1pctCO2*); see Eyring et al. (2016). The experimental design of the *1pctCO2* runs yields simulations of the climate under a very large range of WMGHG forcing, from approximately $\approx 0.32 \text{ W m}^{-2}$ (the 1850 level of WMGHG forcing) to more than 10 W m^{-2} (present day WMGHG forcing is approximately 3.6 W m^{-2}). The *abrupt4xCO2* experiment is branched from the *piControl* and involves an immediate quadrupling of the global annual mean 1850 CO_2 value that is used in *piControl*. Table F.2 in Supplementary material contains a tally of the GCMs and ensemble members that provide the requisite variables for *abrupt4xCO2* and *1pctCO2*.

AerChemMIP experiments. The Aerosol Chemistry Model Intercomparison Project (Collins et al. 2017, AerChemMIP;) is a CMIP6-endorsed initiative that seeks to quantify the various impacts of aerosols and chemically reactive gases on climate and air quality, with a specific focus on near-term climate forcers (SLCFs: methane, tropospheric ozone and aerosols, and their precursors), nitrous oxide and halocarbons. We use the targeted AerChemMIP experiments with CMIP6 climate models to assess a variety of questions relating to structural uncertainty in our representation of anthropogenic aerosols. First, we utilize a set of AerChemMIP time-slice experiments (see Table 6 of Collins et al. 2017) to assess the influence of individual SLCFs by comparing simulations with pre-industrial versus present-day values of each SLCF. The time-slice control experiment (*piClim-control*) uses 1850 concentrations of WMGHGs and all SLCFs and is run in an atmosphere-only configuration with prescribed sea-surface temperatures (SSTs) and sea ice. Otherwise, we use a set of experiments that mirror the control but change a single aerosol precursor from 1850 to 2014 values: *piClim-SO2*

(anthropogenic emissions of SO_2), *piClim-BC* (black carbonaceous aerosols), *piClim-OC* (organic carbonaceous aerosols), *piClim-NH3* (anthropogenic emissions of ammonia), and also *piClim-aer* (all aerosol precursors). See Table F.3 in Supplementary material for the models and ensemble members used. Next, we utilize the AerChemMIP “histSST” simulations of the historical record (1850 to 2014), which are atmosphere-only runs with prescribed SSTs and sea ice based on the historical record (see Table 2 of Collins et al. 2017). The models that participate in this experiment must include an interactive aerosols component, and (depending on the climate model) may or may not additionally include interactive tropospheric or stratospheric chemistry; see Table F.4 in Supplementary material for the models and ensemble members used.

C20C+ and HAPPI Experiments. First, we use the “all historical” (All-Hist) ensembles from the Climate of the 20th Century Plus Detection and Attribution (C20C+ D&A) project which are uncoupled, atmosphere- and land-only simulations conducted using observed historical climate forcings, sea surface temperatures, and sea ice coverage. The All-Hist simulations correspond to approximately $+0.85^\circ\text{C}$ of warming relative to the pre-industrial period. We furthermore utilize the Plus15-Future and Plus20-Future simulations from the Half a Degree Additional Warming, Prognosis and Projected Impacts (HAPPI; Mitchell et al. 2017), which explore stabilized warming experiments of $+1.5^\circ\text{C}$ and $+2.0^\circ\text{C}$, respectively. Like the All-Hist, these are atmosphere-only simulations with 2006–2015 boundary conditions modified to reflect stabilized increases in near-surface air temperatures. For the purposes of our analyses (see Sect. 4.6), we require a large ensemble (at least 50 members) that spans at least 10 years for each of All-Hist, Plus15-Future, and Plus20-Future as well as daily precipitation rate (see Table F.5 in Supplementary material).

3.2 Characterizing external forcings

At the global scale, we assume that the effects of solar forcings and stratospheric ozone on mean and extreme precipitation are small and negligible; we will later test this claim for the CONUS analysis. Otherwise, we need to account for well-mixed greenhouse gases (WMGHG), volcanic forcing, and anthropogenic aerosols. Time series of the following sources are shown in Fig. 2.

WMGHG forcings. The five WMGHGs we need to account for are carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and the CFC-11 and CFC-12 halocarbons. The atmospheric concentration values for each WMGHG over 1850–2014 come from the historical boundary conditions (Meinshausen and Vogel 2016), while the 2015–2020 values come from the SSP585 boundary condition files (Meinshausen and Nicholls 2018). The forcing formulae for CO_2 ,

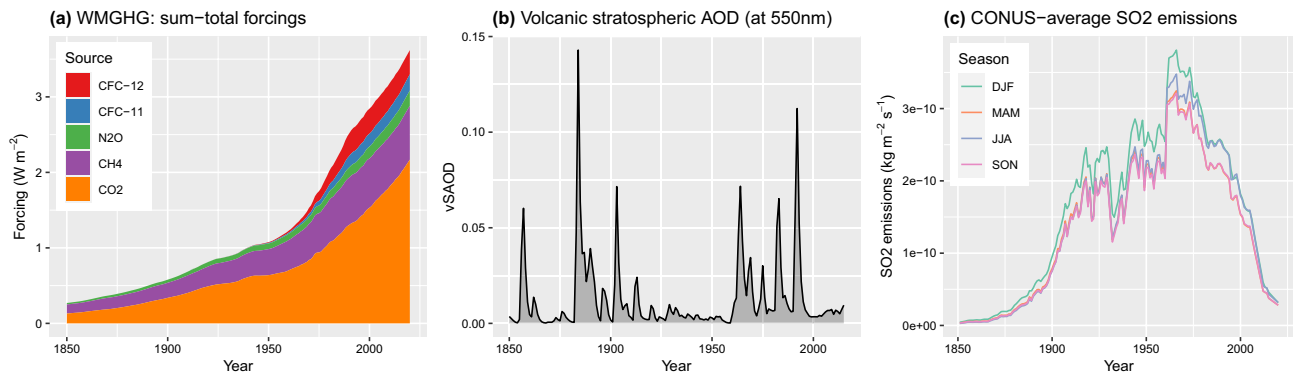


Fig. 2 Observational time series of external forcings used for the various model analyses in this paper over 1850 to present day for well-mixed greenhouse gases (sum-total WMGHG forcings in W m^{-2} ; **a**

stratospheric aerosol optical depth due to volcanoes at 550 nm (**b**), and CONUS-average SO_2 emissions for each season (**c** MAM and SON are nearly overlapping)

CH_4 , and N_2O are in Etminan et al. (2016), Table 1; the forcing formulae for CFC-11 and CFC-12 are given in Table 3 of Hodnebrog et al. (2013). To characterize the WMGHG forcing overall, we will simply compute a total WMGHG forcing ($\text{CO}_2 + \text{CH}_4 + \text{N}_2\text{O} + \text{CFC-11} + \text{CFC-12}$), all in one value as a function of time; see Fig. 2a.

Since the response of precipitation to WMGHG forcings is dominated by the lagged response of SSTs to these forcings (Samset et al. 2016; Douville and John 2021), and since these time lags range between 1 and 3 decades (Ricke and Caldeira 2014) due to the thermal inertia of the upper ocean, the function for total WMGHG forcing includes time lags as discussed in Appendix A.

Volcanic aerosols. For this forcing, we will simply use the observational time series of volcanic stratospheric aerosol optical depth at 550 nm (vSAOD). As with our previous observational analysis (Risser et al. 2021), we use a hybrid time series derived from Sato et al. (1993) (which covers 1850–2012) as well as an updated time series of volcanic SAOD from Schmidt et al. (2018) and Mills et al. (2016) (data provided via personal communication with Dr. Anja Schmidt), which covers January 1975 to December 2015 along with estimated values for 2016–2020. See Risser et al. (2021) for more details.

Anthropogenic aerosols. At the CONUS scale, we require a time series of anthropogenic aerosols that can be applied to the observational record. A significant amount of the analysis in this paper specifically addresses this question (see Sect. 4.3), for which we utilize the CMIP6 forcing datasets (Hoesly et al. 2018; Gidden et al. 2018), which provide monthly measurements of anthropogenic SO_2 emissions (in $\text{kg m}^{-2} \text{s}^{-1}$) across various sectors (e.g., agriculture, industry, etc.) on a $0.5^\circ \times 0.5^\circ$ longitude/latitude grid. The monthly data are summed across sectors, averaged seasonally, and

then averaged across all $0.5^\circ \times 0.5^\circ$ grid cells that lie within the CONUS boundary. These seasonal time series aggregated to the CONUS scale, shown in Fig. 2c, agree well with various other estimates of total CONUS emissions (see Supplemental Figure G.1).

3.3 Drivers of natural climate variability

A recent paper by Risser et al. (2021) forms the basis for our forthcoming observational analysis and identifies a set of drivers of natural climate variability for describing year-to-year variability in seasonal precipitation over CONUS in the observational record. Given that the motivation of this paper is to set the stage for an observationally based attribution statement regarding seasonal precipitation in CONUS, we must similarly account for modes of natural variability in an analysis of climate model output. Section 2.2 of Risser et al. (2021) contains a thorough literature review of the different sources of natural climate variability considered (as well as those that are not considered); Sect. 4.1 of Risser et al. (2021) demonstrates that this set of “drivers” (as we will henceforth refer to them in this paper) is complete insofar as they sufficiently account for year-to-year variability in seasonal precipitation that can be explained by known modes of large-scale oceanic and atmospheric variability. Here, we explore the ENSO Longitude Index (ELI, Williams and Patricola 2018); the Pacific/North American (PNA) pattern (Wallace and Gutzler 1981; Barnston and Livezey 1987); the Arctic Oscillation (AO) (Thompson and Wallace 2000); the North Atlantic Oscillation (NAO) (Hurrell et al. 2003); and the Atlantic Multidecadal Oscillation (AMO) index (Schlesinger and Ramankutty 1994; Kerr 2000) to quantify multidecadal variability in Atlantic sea surface temperatures (Mann et al. 2020, 2021). Unlike Risser et al. (2021), we do not include volcanic stratospheric aerosol optical depth as a result of the analysis of the DAMIP hist-nat ensemble in

Sect. 4.2. We use the Climate Variability Diagnostics Package (CVDP; Phillips et al. 2014) to calculate all of these indices (except ELI) from the various coupled climate model runs examined throughout this paper, and implement the ELI calculation as a Python-based analysis pipeline in the Toolkit for Extreme Climate Analysis (Loring et al. 2016). All indices are calculated based on monthly data and then averaged seasonally (following Risser et al. 2021).

4 Results

We now describe a set of analyses that test each of the hypotheses outlined in Sect. 1. Figure 1 provides an overview the motivation and testing of each hypothesis and Table 3 re-states the hypotheses and our conclusions. Table 3 also provides a confidence assessment for each conclusion (derived in Appendix E). These confidence statements use language as precisely defined by Mastrandrea et al. (2010); note, however, that these statements do not reflect a measure of expert assessment but instead summarize uncertainty across the large CMIP6 multi-model ensemble.

4.1 Global analysis (H1 and H2)

Our first investigation involves hypotheses H1 and H2, which ask whether we can correctly identify the magnitude of the WMGHG effect on precipitation via a linear regression framework (H1), and if so, can we further isolate this WMGHG dependence in a noisy climate system with all forcings (H2)? The hist-GHG ensembles provide a clean test of hypothesis H1, since the only external influence on precipitation in these simulations is the WMGHG forcing. For H2, we subsequently use the historical runs to assess whether the WMGHG dependence is the same in both scenarios or the WMGHG dependence is meaningfully influenced by other external forcings. For both of these tests, we focus on the global average of annual mean and extreme precipitation; here, the global average for extreme precipitation involves the grid cell average of annual maximum daily precipitation (Rx1Day).

To test H1, we conduct three regression analyses: first, we regress

$$Y_{m,e}^{\text{hist-GHG}}(t) = \beta_{m,0}^{\text{hist-GHG}} + \beta_{m,1}^{\text{hist-GHG}} F \text{ WMGHG}(t, \tau \text{ WMGHG}_m) + \varepsilon_{m,e}(t), \quad (7)$$

where $Y_{m,e}^{\text{hist-GHG}}(t)$ is the global average, grid-cell area weighted, annual time series for each hist-GHG ensemble member $e = 1, \dots, N_m$ and year $t = t_0^m, \dots, t_T^m$ for model $m = 1, \dots, 11$. In Eq. 7, $F \text{ WMGHG}(t, \tau \text{ WMGHG}_m)$ is the total lagged WMGHG forcing in year t using the model-specific lag $\tau \text{ WMGHG}_m$, the error term $\varepsilon_{m,e}(t)$ is distributed

as $N(0, \sigma^2(t))$, and the error variance is modeled as in Eq. 5. Second, we model the global average of extreme daily precipitation, i.e., Rx1Day (mm), denoted $Z_{m,e}^{\text{hist-GHG}}(t)$ as arising from a GEV distribution (see, e.g., Coles 2001) with time-varying location

$$\mu_m^{\text{hist-GHG}}(t) = \alpha_{m,0}^{\text{hist-GHG}} + \alpha_{m,1}^{\text{hist-GHG}} F \text{ WMGHG}(t, \tau \text{ WMGHG}_m), \quad (8)$$

time-varying scale $\sigma_m^{\text{hist-GHG}}(t)$ (modeled as in Eq. 5), and time-invariant shape $\xi_m^{\text{hist-GHG}}$. Finally, we regress the global average of mean surface temperature t_s (K), denoted $S_{m,e}^{\text{hist-GHG}}(t)$, on WMGHG as in Eq. 7. In each case, we aggregate across the ensemble and estimating a single set of coefficients for each model. We use these fitted regression equations to estimate the precipitation rate, 20-year return value, and global average surface temperature for 1850 and 2020 and then calculate the model-specific scaling rates (percent change in precipitation per K), denoted $R_{\text{mean},m}^{\text{hist-GHG}}$ and $R_{\text{GEV},m}^{\text{hist-GHG}}$, by comparing 2020 estimates with 1850 estimates. To propagate the uncertainty through from the various regressions to the final scaling factors $R_{\text{mean},m}^{\text{hist-GHG}}$ and $R_{\text{GEV},m}^{\text{hist-GHG}}$, we utilize the bootstrap to calculate basic bootstrap confidence intervals (see, e.g., Paciorek et al. 2018), here and throughout at the 95% confidence level.

The estimated hist-GHG percent changes in precipitation (mean and return values) and scaling rates as a function of the warming are shown in Fig. 3 for each model along with 95% confidence intervals along both axes. The expected 2%/K and 6%/K for means and extremes, respectively, are shown with black lines. Interestingly, the mean scaling rates seem to be systematically less than 2%/K (although in some cases the confidence interval covers 2%/K); the extreme scaling rate also is generally less than 6%/K, except for the IPSL runs. The scaling factors do not appear to be influenced by the magnitude of the warming, although all but two of the models considered have less than $\approx 1.6\text{K}$ warming over 1850–2020. While the scaling rates for both mean and extreme precipitation generally fall short of their theoretical values, these results are consistent with related explorations of precipitation scaling. For example, Collins et al. (2013) find that the rates of change in means were $1\text{--}3\text{ K}^{-1}$, and changes in daily extreme precipitation ranged from 4 K^{-1} (CMIP3) to 5.3 K^{-1} (CMIP5 models). Furthermore, Kharin et al. (2013) find rates of change for mean and extreme precipitation in CMIP5 models to be, on average, 1.8 K^{-1} and 5.8 K^{-1} , respectively, although these results explore changes for future projections (RCP 2.6, 4.5, and 8.5), which have up to 5K of warming. Furthermore, Kharin et al. (2013) find that these rates are indeed dependent on the level of warming, such that the rate was lower for lower warming amounts. Both of these results are consistent with what we see in Fig. 3, particularly given the moderate levels of warming over the historical period in the hist-GHG runs (note

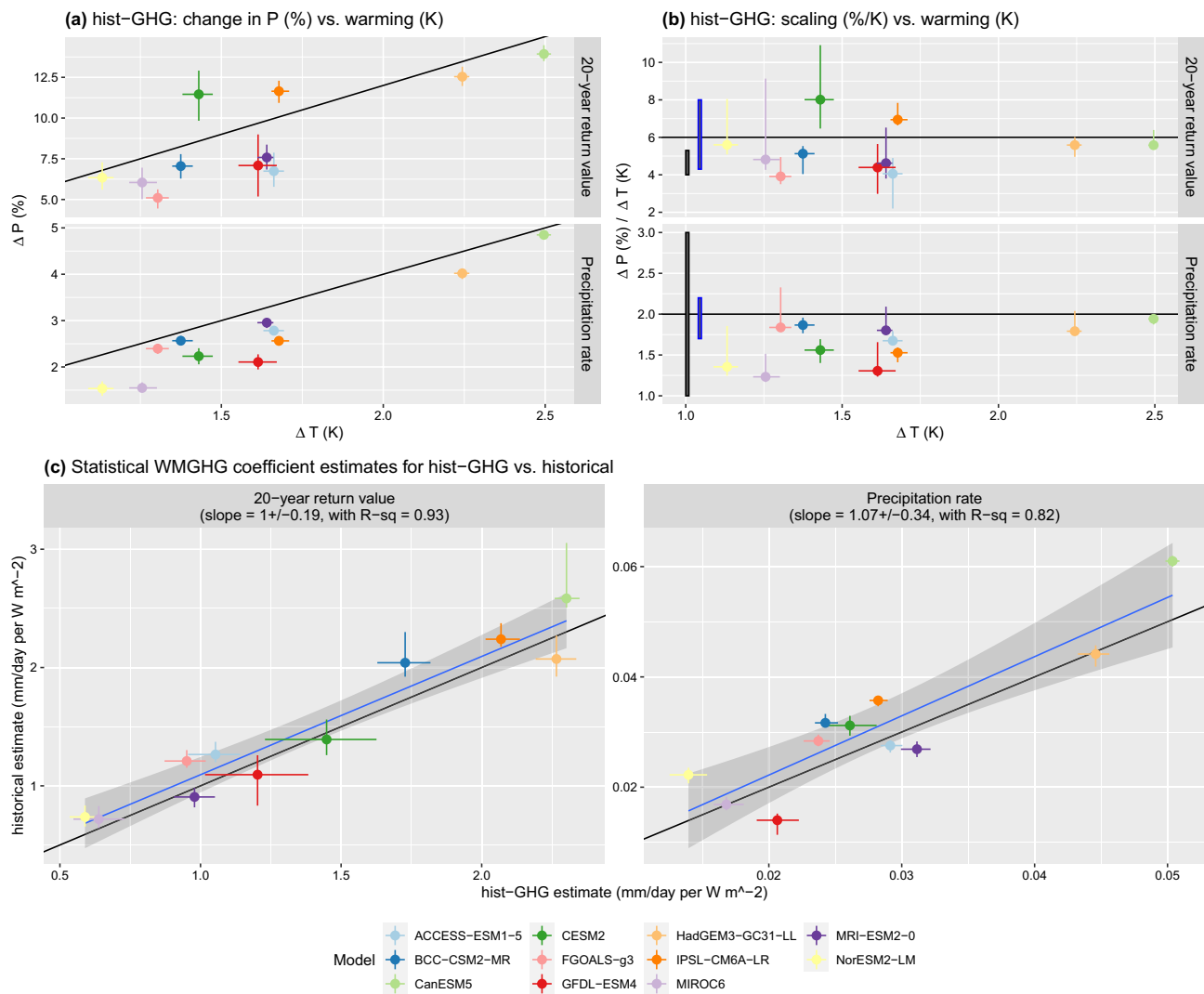


Fig. 3 Estimated hist-GHG changes in globally averaged annual mean and extreme precipitation versus warming (a), the corresponding hist-GHG scaling factors $R_{\text{mean},m}^{\text{hist-GHG}}$ and $R_{\text{GEV},m}^{\text{hist-GHG}}$ versus warming (b), and statistical WMGHG coefficient estimates calculated from hist-GHG versus historical for mean and extreme precipitation (c). Each point contains coordinate-wise 95% basic bootstrap confidence

intervals. In a and b, the black lines denote the theoretically expected responses; the range of scaling rates from Collins et al. (2013) and Kharin et al. (2013) are shown in panel (b) with black and blue bars, respectively. The black lines in c are the 1-1 line, and the blue lines in c represent the linear fit for the historical relative to hist-GHG (95% uncertainty band in light gray)

that Fig. 3b also shows the range of scaling rates from Collins et al. (2013) and Kharin et al. (2013) for reference). Similar results are obtained when analyzing the CMIP6 DECK experiments piControl and 1pctCO2 introduced in Sect. 3.1, which involve significantly more warming to the global climate system; see Appendix B. These results support our confidence assessment in Appendix E that it is *likely* (as precisely defined by Mastrandrea et al. (2010)) that the WMGHG forcing as quantified via our regression framework is consistent with theoretical expectations for mean and extreme precipitation.

Next, we outline a formal framework for testing hypothesis H2, where we assess whether we can isolate the WMGHG influence in the historical runs at the global scale and how this compares with the hist-GHG runs. We focus on the global scale so that, for the purposes of this test, we can ignore the substantial and demonstrable regional effects of long-term unforced variability and weather noise (McKinnon and Deser 2018; Deser 2020) on the forced response of the climate system (Kay et al. 2015, Fig. 2). For this test, we assume that the effects of solar forcings and stratospheric O_3 are small, and so we just have to account for the volcanic SAOD, WMGHGs, and aerosols in the historical runs. As

previously mentioned, accounting for the aerosol forcing in the historical runs is nontrivial. To account for the different ways each modeling center includes aerosols and to account for the spatially heterogeneous influence of aerosols globally, as a pre-processing step we first calculate a smoothed aerosol trend in mean and extreme precipitation in the hist-aer runs (specifically, fitting a natural spline curve with 7 degrees of freedom over 1850–2020). Then, we subtract this smoothed trend from the historical annual time series at each grid cell. Once we have removed the aerosol trends, we calculate two area-weighted global average time series for each historical model/ensemble member: the global average of mean precipitation (mm/day) minus the aerosol trend, denoted $Y_{m,e}^{\text{hist, no-aer}}(t)$; and the global average of extreme daily precipitation, i.e., Rx1Day (mm) minus the aerosol trend, denoted $Z_{m,e}^{\text{hist, no-aer}}(t)$. Similar to the hist-GHG analysis, we then conduct the following regressions:

$$Y_{m,e}^{\text{hist, no-aer}}(t) = \beta_{m,0}^{\text{historical}} + \beta_{m,1}^{\text{historical}} F \text{ WMGHG}(t, \tau \text{ WMGHG}_{m,m}) + \beta_{m,2}^{\text{historical}} \text{vSAOD}(t) + \varepsilon_{m,e}(t) \quad (9)$$

for the mean (again assuming time-varying Normal error), and modeling the Rx1Day measurements $Z_{m,e}^{\text{hist, no-aer}}(t)$ as arising from a GEV distribution with location parameter

$$\mu_m^{\text{historical}}(t) = \alpha_{m,0}^{\text{historical}} + \alpha_{m,1}^{\text{historical}} F \text{ WMGHG}(t, \tau \text{ WMGHG}_{m,m}) + \alpha_{m,2}^{\text{historical}} \text{vSAOD}(t) \quad (10)$$

with time-varying scale $\sigma_m^{\text{historical}}(t)$ and time-invariant shape $\xi_m^{\text{historical}}$. Our test then concerns whether $\beta_{m,1}^{\text{hist-GHG}} \approx \beta_{m,1}^{\text{historical}}$ and $\alpha_{m,1}^{\text{hist-GHG}} \approx \alpha_{m,1}^{\text{historical}}$ (where the hist-GHG quantities are as in Eqs. 7 and 8). Uncertainty is once again quantified using the bootstrap.

For the 11 models in Table F.1 in Supplementary material that have both hist-GHG and historical runs, we compare the statistical WMGHG coefficient estimates from these two experiments along with their 95% basic bootstrap confidence intervals; see Fig. 3. In these plots, the x -axis shows the hist-GHG coefficients ($\beta_{m,1}^{\text{hist-GHG}}$, panel a.; $\alpha_{m,1}^{\text{hist-GHG}}$, panel b.) with the historical coefficients on the y -axis ($\beta_{m,1}^{\text{historical}}$, panel a.; $\alpha_{m,1}^{\text{historical}}$, b.). The plots also show the 1–1 line (black) and the linear fit for the historical relative to hist-GHG (blue line; 95% uncertainty band in light gray); the caption of each figure shows the slope of the blue linear fit ($\pm$ the standard error) and the R^2 value for the linear fit. While the individual models' confidence intervals often do not intersect the 1–1 line, across the multi-model ensemble the regression coefficients for hist-GHG and historical generally agree quite well. In other words, we can indeed isolate the WMGHG dependence in the historical runs consistently with what we get from the hist-GHG results. This result supports our confidence assessment in Appendix E that it is *very likely* (as

precisely defined by Mastrandrea et al. (2010)) that there are no meaningful cross-correlation terms in the WMGHG influence in precipitation over the historical record.

4.2 CONUS trends from individual forcing agents (H3)

Next, we set out to investigate if there are meaningful trends in seasonal precipitation at the scale of CONUS due to individual forcing agents. Using the DAMIP and LUMIP experiments, we can explicitly quantify trends due to solar forcing, volcanoes, stratospheric ozone, land use/land cover change, WMGHGs, and anthropogenic aerosols. For this hypothesis, to be as general as possible, we simply assess changes in seasonal precipitation (mean and extreme) due to changes in global average surface temperature for each model and each experiment. For each model grid cell over CONUS, we first extract the seasonal (DJF, MAM, JJA, and SON) mean precipitation rate and maximum daily precipitation (Rx1Day) in each year and ensemble member, and then take area-weighted averages of the seasonal quantities. Next, we regress the time series of seasonal precipitation versus the global average surface temperature (using OLS for precipitation rate and GEV regression for Rx1Day and in both cases using a time-varying weather variance as in Sect. 4.1), aggregating over the ensemble members to obtain a single trend estimate. Next, we convert the trends in global surface temperature to a change in precipitation per century and then use the block bootstrap (as in, e.g., Risser et al. 2019b) to quantify uncertainty and provide a 95% confidence interval. Note that to assess the influence of land use/land cover change (LULCC), we generate a “hist-Lu” scenario by subtracting the hist-noLu ensemble average time series from the historical ensemble average time series, which implicitly assumes linearity in the relationship between LULCC and seasonal precipitation. For comparison, we also calculate and plot corresponding trends from the CMIP6 historical and hist-noLu experiments.

The CONUS-average individual forcing trends are shown in Fig. 4 for each model, experiment, and season for both mean precipitation and Rx1Day. Across essentially all models and seasons, it is apparent that WMGHG forcing yields strong positive trends in mean and extreme precipitation at the CONUS scale as evidenced by the fact that the hist-GHG confidence intervals are generally positive and do not include zero in most cases. Similarly, the anthropogenic aerosol forcing yields strong negative trends in both mean and extreme precipitation since the hist-aer confidence intervals are negative and again do not include zero. Both of these results are very much expected, particularly for precipitation globally, but it is interesting that we verify this result at the CONUS scale. Conversely, the hist-nat and hist-stratO3 trends are in most cases non-distinguishable

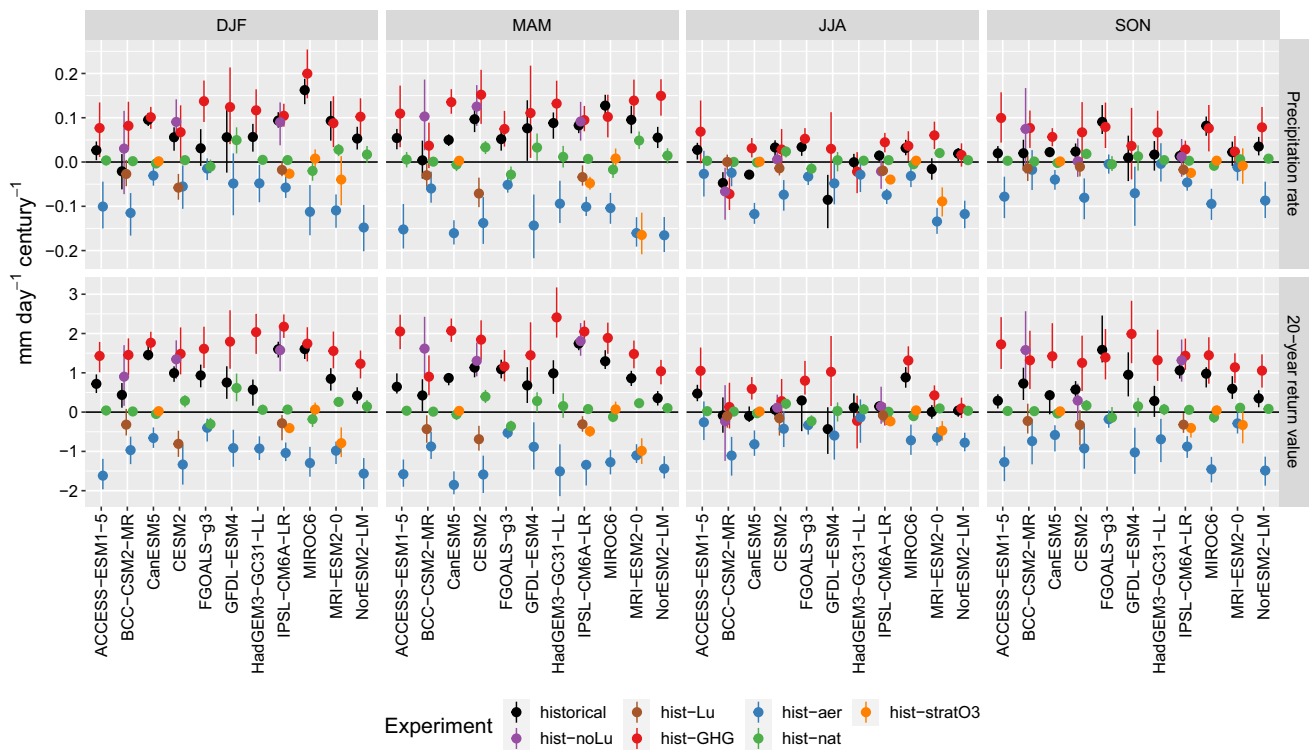


Fig. 4 CONUS-average single-forcing trends in seasonal mean (top) and extreme (bottom) precipitation, calculated via a regression versus global average surface temperature, aggregated across ensemble

from zero, meaning that these forcing agents do not significantly influence seasonal precipitation at the CONUS scale. In other words, this implies that neither stratospheric ozone nor the natural forcings (solar and volcanoes) have a meaningful impact on seasonal precipitation over CONUS. Finally, while there are only a few models that provide the hist-noLu (and hence the derived hist-Lu) experiment, in most cases the trends due to LULCC as represented by our derived hist-Lu time series overlap zero. It is also generally the case that the hist-noLu trends are non-distinguishable from the historical trends. As such, we also conclude that LULCC does not significantly influence seasonal precipitation over CONUS.

In summary, we have found that WMGHG forcing tends to increase seasonal precipitation, anthropogenic aerosol forcing tends to decrease seasonal precipitation, and the influence of natural forcings, stratospheric ozone, and LULCC is minimal and can be, for our analysis, ignored. These results support our confidence assessments in Appendix E regarding the likelihood that individual forcings are non-negligible for secular trends in precipitation over CONUS: *likely* for WMGHGs and anthropogenic aerosols, *about as likely as not* for stratospheric ozone and LULCC, and *unlikely* for natural forcings (italicized terms are precisely defined by Mastrandrea et al. (2010)).

members, and converted to a change in precipitation per century. Each point includes a 95% basic bootstrap confidence interval

4.3 Structural uncertainty in representing anthropogenic aerosols (H4)

Given the clear influence of tropospheric aerosols on seasonal precipitation identified in Fig. 4, we must account for this external forcing in an analysis of the historical record. The obvious question then becomes: how exactly should this be done appropriately? The natural choice would be to use atmospheric SO_2 concentrations because it is atmospheric SO_2 that alters clouds via oxidation to sulfate cloud-condensation nuclei (Lohmann and Feichter 2005; Chen et al. 2008). However, this is problematic because there are no detailed long-term, spatially resolved records of SO_2 concentrations, and we are hence reliant on models. To make matters worse, atmospheric concentrations and surface deposition rates of SO_2 differ considerably among models due to their chemistry parameterizations and underlying atmospheric codes (Lamarque et al. 2013). This diversity is evident in the CONUS-mean atmospheric burdens and deposition rates of SO_2 from AerChemMIP histSST runs, such that the only field with minimal spread across the multi-model ensemble are the prescribed anthropogenic emissions (Fig. G.3 in Supplementary material).

To this end, we now conduct a set of tests to address structural uncertainty in anthropogenic aerosols and justify our conclusions using climate model output from the AerChemMIP experiments described in Sect. 3.1. Our exploration of how to characterize anthropogenic aerosols is simplified somewhat due to results from the Precipitation Driver and Response Model Intercomparison Project (PDRMIP; Myhre et al. 2017). Using PDRMIP simulations, (Liu et al. 2018) show that Asian and European sulfate aerosols have negligible effects on precipitation over CONUS, and therefore we can use aerosol measures local to CONUS. This conclusion is further supported by source apportionments of the SO_2 over CONUS, which show that until 2000 that $\leq 10\%$ of the SO_2 burden over any of the four quadrants (northwestern, southwestern, northeastern, and southeastern) of US national territory had a non-domestic origin (Yang et al. 2018). Over the global land surface, PDRMIP finds that the fast precipitation response to increased sulfate aerosol burdens is approximately three times that of the slow (SST-mediated) response, although at continental scales (e.g., CONUS) the contributing models do not consistently agree on the partitioning between fast and slow responses (Samset et al. 2016). We have used this finding to support omission of lagged (slow) responses of CONUS precipitation to SO_2 . Supported by the PDRMIP findings and the source apportionment by Yang et al. (2018), we hypothesize that the response of CONUS to SO_2 is dominated by fast responses to domestic emissions of this species.

4.3.1 Dominant aerosol species for changes in precipitation (H4a)

First, we show that SO_2 is, in fact, the dominant aerosol species for changes in precipitation. To support this conclusion, we utilize the AerChemMIP time-slice experiments to assess the influence of each individual near-term climate forcer (SO_2 , black carbonaceous aerosols, organic carbonaceous aerosols, and ammonia) on seasonal precipitation. This test involves looking at changes in seasonal precipitation (both precipitation rate and Rx1Day) for each of piClim- SO_2 , piClim-BC, piClim-OC, and piClim- NH_3 relative to piClim-control averaged over the climatology of each experiment. In addition to comparing each “single forcing” experiment versus the control, we further quantify the combined effects of black carbonaceous (BC) aerosols, organic carbonaceous aerosols, and ammonia relative to the control simulation, since the sum of the global-mean effective radiative forcings (ERFs) by these species is much less than that of SO_2 for 1850–2014 (Thornhill et al. 2021) and the globally averaged effects of BC aerosols on historical land-mean precipitation are negligible during 1850 to 2005 (Richardson et al. 2018). All differences are calculated for each model grid cell, and we then summarize the differences for several land

regions: all Northern Hemisphere land regions (NH land), CONUS, China, and India. Our exploration here concerns regions outside of CONUS due to the relatively small aerosol forcing over CONUS in 2014 (the comparison year in AerChemMIP).

The results of this analysis are summarized in Table 1, which shows a tally of the model-season combinations (out of 16 total) for which the sum of black carbonaceous aerosols, organic carbonaceous aerosols, and ammonia (BC + OC + NH_3) cannot be reliably distinguished from control run for each region. Focusing on the NH land region for seasonal Rx1Day, the changes for BC + OC + NH_3 cannot be reliably distinguished from 0 for all but one model-season combination; for seasonal precipitation rate, this is also true for 9/16 of the model-season combinations. Furthermore, for all seasons and models (and regions), it is possible that the decreases in both precipitation rate and Rx1Day for SO_2 are larger in magnitude (more negative) than those for BC + OC + NH_3 to within the full 95% CI intervals, i.e., this possibility is not excluded. This result is shown for NH land in Supplemental Figure G.2. This result supports our confidence assessment in Appendix E that it is *likely* (as precisely defined by Mastrandrea et al. (2010)) that SO_2 is the aerosol species that dominates historical changes in precipitation over CONUS.

4.3.2 Characterization of SO_2 influence on precipitation (H4b)

Because SO_2 is the dominant aerosol species for driving changes in seasonal precipitation, the next question involves how to quantify the influence of SO_2 . There are (at least) three candidates: SO_2 emissions (denoted “emiso2”), deposition of SO_2 by rainfall (denoted “wetso2”), and column integrated SO_2 mass (denoted “iso2”). Each of these quantities are resolved within the model to the grid cell, either as a boundary condition (emiso2) or as calculated internally within the model via parameterizations (wetso2 and iso2). In principle, the ultimate causal variable relating aerosols to precipitation is emissions, but the relationship between emissions and precipitation is quite complicated because of transport of SO_2 and transformations in

Table 1 Number of model-season pairs (16 total) for which the sum of black carbonaceous aerosols, organic carbonaceous aerosols, and ammonia cannot be reliably distinguished from the control run

Region	Precipitation rate	Rx1Day
Northern Hemisphere land	9	15
CONUS	14	15
China	11	12
India	13	14

Here, “does not differ” is determined when the 95% confidence interval includes 0

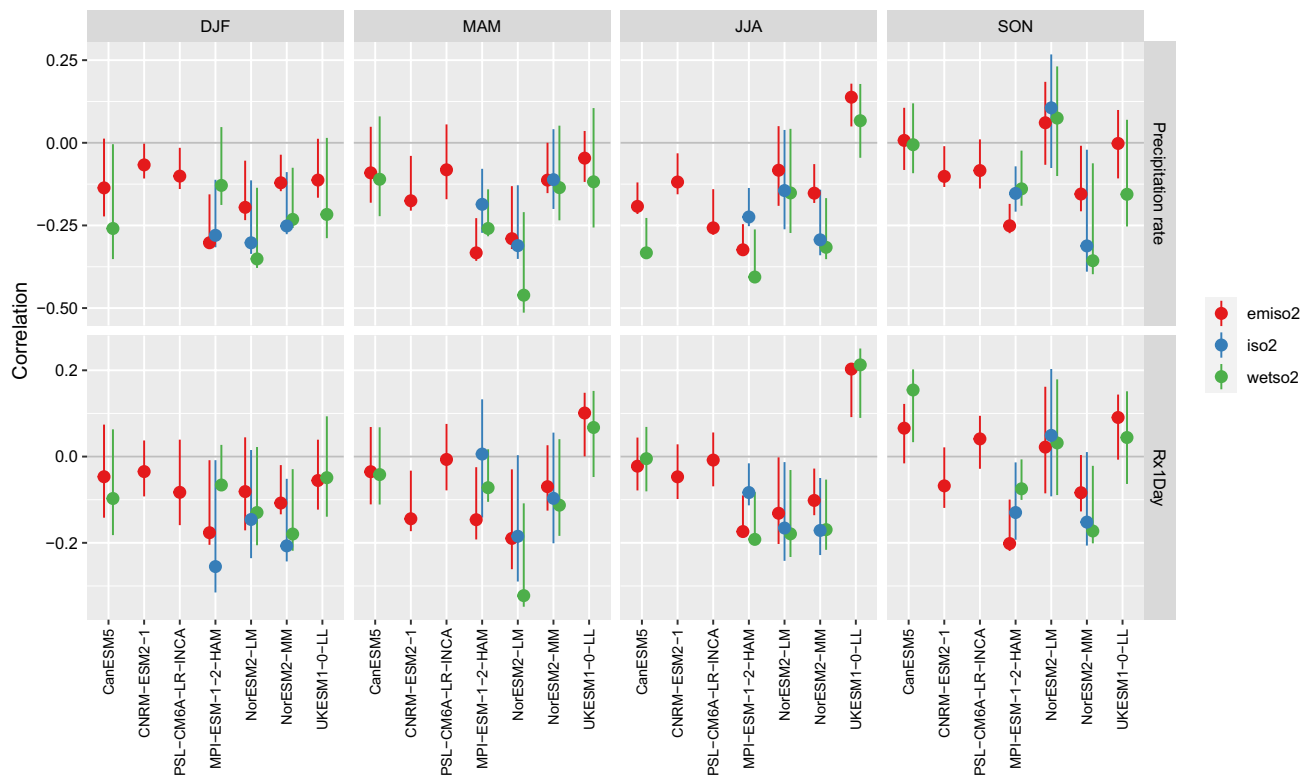


Fig. 5 Correlation between change in precipitation and change in SO_2 (piClim-aer minus piClim-control) with 95% basic bootstrap confidence interval for all model grid cells over Northern Hemisphere land regions. Across all models/seasons and mean/extreme precipitation,

the correlations between emiso2 (SO_2 emissions) and precipitation are indistinguishable from correlations between precipitation and either wetso2 (deposition of SO_2 by rainfall) or iso2 (column integrated SO_2 mass)

the atmosphere. Therefore, we seek empirical support for a simple functional form for the effect of emissions on precipitation. Ideally, emiso2 does as well as either of the other two measures since it is the only field we know reliably in the observational record. To test this claim, we again use the piClim-aer and piClim-control AerChemMIP time-slice experiments to assess the correlation over grid cells between changes in seasonal precipitation (piClim-aer minus piControl) and changes in the SO_2 sources (again piClim-aer minus piControl).

Figure 5 shows the correlations over grid cells for NH land for all available data across precipitation type (mean or extreme), season, and SO_2 source along with a 95% basic bootstrap confidence interval. For the NH land regions, the correlations of both precipitation rate and Rx1Day with emiso2 across all combinations of seasons and models are not meaningfully different from the corresponding correlations with iso2 and wetso2 when considering the full 95% CI intervals. The same is true for CONUS and China (not shown); in India the same is true except for MAM precipitation rate in three of the models (again not shown). In other words, we show that the relationship between emissions and precipitation is very similar to that between precipitation and variables more directly connected to cloud microphysics (integrated SO_2 or SO_2 rainout).

This result supports our confidence assessment in Appendix E that it is *very likely* (as precisely defined by Mastrandrea et al. (2010)) emiso2 correlates with changes in precipitation as well as wetso2 and iso2; hence we can safely use emiso2 to characterize the relationship between seasonal precipitation and anthropogenic SO_2 .

4.3.3 Regionalized vs. local estimates of SO_2 emissions (H4c)

Finally, given the results of Hypotheses 4a and 4b, we evaluate whether the influence of anthropogenic aerosols should be determined using local (model grid-cell) estimates of emiso2, or whether regionally-averaged emissions could be used instead. In principle, a localized measure would be appropriate given the dominance of fast (microphysics-mediated) effects of SO_2 on rainfall (Samset et al. 2016; Yang et al. 2018), the large geographic gradients in anthropogenic emissions and column-integrated atmospheric burdens of SO_2 across CONUS (Yang et al. 2018) as well as in the frequency of different precipitating weather systems (Kunkel et al. 2012), and the short characteristic lifetimes of

both SO₂ (Lee et al. 2011) and cloud systems (e.g., Esmaili et al. 2016; Dagan et al. 2018)).

Unfortunately, for the time period over which we have sufficient in situ precipitation data (1900 to present), the seasonally-averaged time series of grid cell SO₂ emissions is highly correlated with the WMGHG forcing time series (exceeding 0.8 in absolute value for $\approx 50\%$ of CONUS; in most cases emis02 and WMGHGs are anti-correlated). This means that it will be essentially impossible for a regression-based analysis to distinguish the effect of WMGHG on precipitation from the effect of SO₂. Note, however, that the grid cell emis02 versus WMGHG correlations are much weaker for 1850 to present and become sufficiently small for a regression analysis. As such, we can confidently separate the influence of WMGHG from the influence of grid cell SO₂, but only when the time series extends back to 1850 (and not when we are limited to a starting year of 1900). See Appendix C for further details on the relationship between WMGHG forcing and grid-cell SO₂ emissions. At the CONUS scale, we see that emissions have generally trended up and then down similarly across the US due to broad societal reasons (i.e., industrialization and then regulations; see Fig. 2) such that the correlation between the CONUS-average emissions and WMGHG forcing is negligible.

A secondary problem arises due to the fact that the $0.5^\circ \times 0.5^\circ$ gridded observational data set of SO₂ emissions (Hoesly et al. 2018; Gidden et al. 2018) is extremely heterogeneous across CONUS, such that the average annual emissions in nearby grid cells can differ by a factor of nearly 1000 (see Fig. 13a). This extreme spatial variability occurs at very small scales: in some cases, for neighboring grid cells (e.g., portions of the Great Plains; again see Fig. 13a). In this case one must be concerned with transport of SO₂ emissions, since it is unlikely that the grid cells in the Great Plains with essentially zero emissions adjacent to grid cells with very large emissions have negligible atmospheric SO₂ concentrations. Unfortunately, there are neither long-term nor continental-scale observations of SO₂ transport, and simulations of this phenomenon are subject to large structural and parametric uncertainties in convective transport and subsequent scavenging (Feng et al. 2011; Tost et al. 2010) that act to reduce confidence in model estimates of aerosol-cloud interactions (Storelvmo 2012).

To address both of these concerns, we propose an approach that yields a compromise between using grid cell emissions (which are strongly correlated with WMGHG forcing over 1900-present and do not realistically represent SO₂ transport) and a CONUS-wide average emissions time series (which has the correct trajectory but ignores the important local effects of SO₂). Specifically, we use “stochastically regionalized” time series of SO₂ emissions, wherein the grid cell emissions are averaged within ensembles of a set of regions across the CONUS (Appendix D).

Since we want the local emissions to follow the same trajectory as the CONUS-wide time series (i.e., generally trending up through 1970 and then decaying), we determine these regions as follows:

1. In each season, identify pairs of years from $\{1900, \dots, 1966\}$ and $\{1967, \dots, 2020\}$ such that the absolute difference of the SO₂ emissions in the CONUS-wide time series in the year-pair is less than a small fraction of peak emissions. The midpoint year of 1966 is chosen since this is the year in which SO₂ emissions reached their maximum; see Fig. 2.
2. The stochastic regionalizer generates a partitioning of CONUS into connected regions that minimize a cost function chosen to decorrelate regional SO₂ emissions from WMGHGs (Appendix D).
3. In order to account for uncertainty in the regionalization, we repeat Step 2 100 times to obtain an “ensemble” of regions.
4. For each ensemble member and using the averaged emissions for the regions identified in Step 2, we reconstruct the grid cell map of seasonal SO₂ emissions in each year.

By construction, the stochastically regionalized (SR) SO₂ emissions time series are only moderately correlated with WMGHGs (on average, roughly -0.2) such that they can be used to simultaneously assess the influence of both emissions and WMGHGs on rainfall.

In order to explicitly test whether using stochastically regionalized emissions introduces significant biases relative to using either grid cell specific emissions or CONUS-wide emissions, we use the AerChemMIP histSST experiments which cover 1850 to 2014. For the histSST simulations, we conduct a regression analysis at each model grid cell where, similar to the station-specific analysis in Risser et al. (2021), we simultaneously assess the influence of WMGHGs, SO₂, and the set of five climate drivers discussed in Sect. 3.3 as follows: for the mean (extreme) precipitation analysis, we allow the mean (location parameter) to change linearly according to the two forcing time series and five climate drivers with a time-varying error term as in Eq. 5. For each grid cell, we conduct the analysis three times: first, using grid cell-specific SO₂ emissions time series; second, using the stochastically-regionalized emissions (here, the analysis is done for each ensemble member, after which we average coefficient estimates across the ensemble); and third, using a single CONUS-wide average emissions time series (all of these time series are extracted from the model-specific emis02 output). Conditional on the estimated SO₂ coefficients, we calculate the effect of emis02 on precipitation by multiplying coefficient estimates by the range of emissions experienced at that grid cell and then calculate an area-weighted CONUS average. All other drivers (as well as the

time-varying error) are fixed at their average values. These estimates are shown for all three characterizations of SO_2 in Fig. 6 for the six models that have all required input data and across seasons and mean/extreme precipitation along with 95% basic bootstrap confidence intervals. Clearly, across all seasons, the estimated effect on seasonal precipitation (for both mean and extreme) is essentially identical for all three methods of characterizing SO_2 emissions. The spatial patterns of this effect (not shown) also agree extremely well for each approach. This result supports our confidence assessment in Appendix E that it is *virtually certain* (as precisely defined by Mastrandrea et al. (2010)) that regionalized time series of emiso_2 yield the same results as using either local or CONUS-wide estimates of emiso_2 . Therefore, we can safely proceed with using regionalized time series of SO_2 emissions to characterize the influence of anthropogenic aerosols in an analysis of the historical record.

4.4 Influence of forcing agents on driver vs. precipitation relationships (H5)

As is documented in Risser et al. (2021), the five climate drivers ELI, AO, NAO, PNA, and AMO are sufficient for describing year-to-year variability in seasonal mean and extreme precipitation that is driven by large-scale modes of

climate variability, i.e., not due to background weather variability. However, a follow-up question that was not directly addressed in Risser et al. (2021) has to do with whether the relationship of these climate drivers and seasonal precipitation is altered by external forcing.

We can directly assess the influence of external forcings on the driver versus precipitation relationships using the DAMIP, LUMIP, and CMIP6 historical experiments. Similar to the station-specific analysis in Risser et al. (2021) and the analysis of AerChemMIP histSST described in Sect. 4.3.3, we conduct a grid-cell specific time series analysis of seasonal mean (using OLS) and extreme (using GEV) precipitation for each model/experiment, again aggregating over any available ensemble members to obtain a single estimate for each grid cell across the ensemble. Both mean and extreme analyses use the time-varying error variance in Eq. 5, and specify that the mean/location parameter vary linearly with a set of experiment-specific terms:

- hist-nat and hist-stratO3: climate drivers only
- hist-GHG: climate drivers and WMGHGs
- hist-aer: climate drivers and SO_2 emissions
- historical and hist-noLu: climate drivers, WMGHGs, and SO_2 emissions

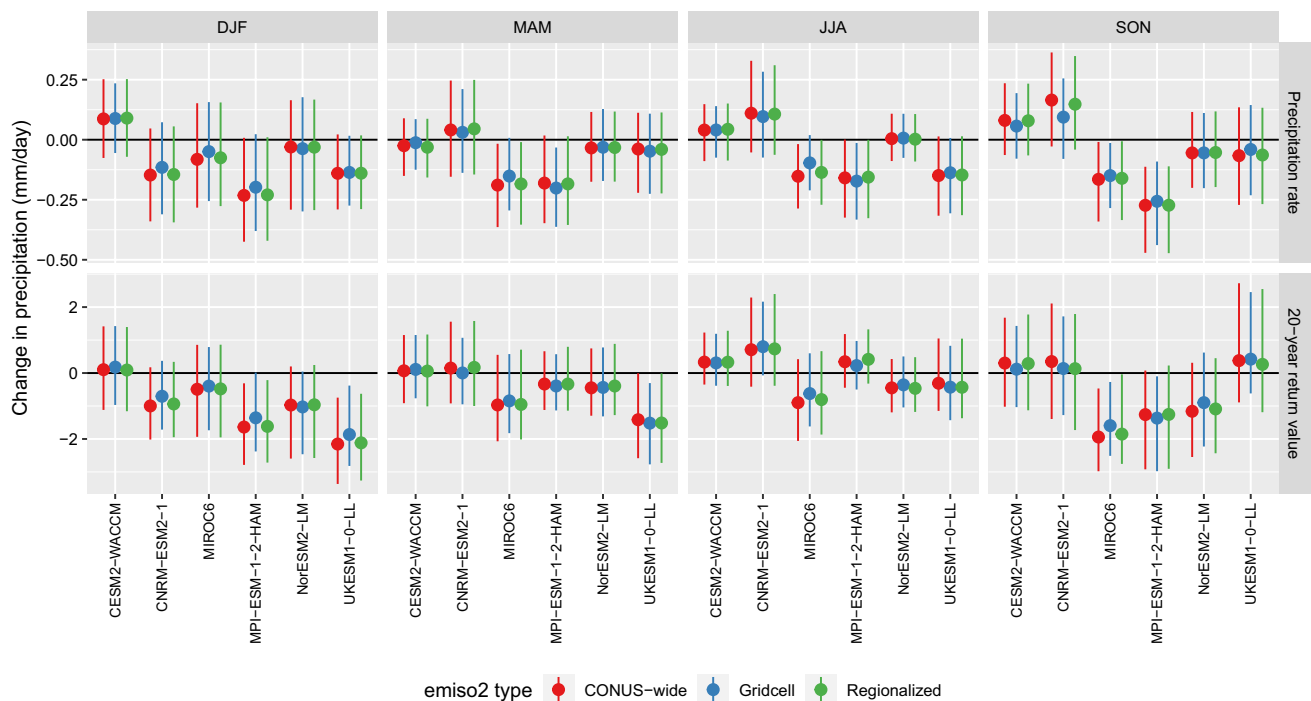


Fig. 6 The CONUS-average effect of maximum versus minimum emis_2 , comparing an analysis where we characterize SO_2 in one of three ways: (1) using model-specific grid cell emis_2 time series

(blue), (2) a single CONUS-wide average emis_2 time series (red), and (3) stochastically regionalized emis_2 (green). Each estimate includes a 95% basic bootstrap confidence interval

For the historical, hist-aer, and hist-noLu simulations, we use stochastically regionalized SO₂ emissions as described in Sect. 4.3.3 and Appendix D. In light of the negligible secular trends identified in the hist-nat simulations (Fig. 4), we can use the hist-nat estimates as reference values to compare versus the estimated driver-precipitation relationships in the other experiments. For example, if the hist-GHG analysis (which includes only “main effects” of the drivers/forcing in statistical terminology) recovers the same driver effects as hist-nat, this indicates that there is not a driver-forcing interaction that modifies the driver effect. Similarly, in this case there would be no modification to the forcing effect, either.

To compare the various regression coefficients for each climate driver across the different experiments versus what we obtain from the hist-nat experiment, we use the statistics of a Taylor diagram (Taylor 2001), a traditional diagnostic plot for comparing climate model output with a reference data set, usually an observationally based product. In this case, we use the spatial fields of hist-nat coefficients from each model as the “reference” data set, and then calculate the spatial pattern correlation (Pearson correlation across pairs of grid cells) and relative standard deviation (standard deviation of the comparison data divided by the standard deviation of the reference data) for all other experiments from each model. We also calculate the overall CONUS-average bias of the driver coefficients for the various experiments relative to hist-nat. After accounting for uncertainty in the bias, relative standard deviation, and pattern correlation via 95% bootstrap confidence intervals, we tally the number of models/experiments for which the 95% confidence intervals do not differ significantly from the hist-nat (all comparisons

are done within-model). For the bias and relative standard deviation, there are straightforward null values for these statistics: a bias of zero and a relative standard deviation of one. For the pattern correlation, we instead calculate the null value as the intra-experiment pattern correlation calculated across the hist-nat ensemble. For the bias and relative standard deviation, the total number of model/experiments is 40 (there are 51 total, 11 of which are the reference hist-nat runs); for the pattern correlation, the total number of models/experiments is 37 since the GFDL-ESM4 hist-nat runs only have a single ensemble member and hence we cannot calculate a within-experiment pattern correlation.

The tallied results are shown in Table 2. For the bias and relative standard deviation, in most cases 80% (or more) of the models/experiments have estimated driver coefficients that are indistinguishable from the hist-nat experiment when aggregating across CONUS. Given that these comparisons are based on 95% confidence intervals, we would expect approximately 95% agreement if the experiments were truly indistinguishable: this is almost always the case for the bias metric, while the relative standard deviation tends to show slightly less agreement with hist-nat. The agreement for pattern correlation is somewhat reduced, relative to the bias and relative standard deviation, although in almost all cases the agreement is better than 50% (and in many cases much higher). This could be due in part to the relatively small hist-nat ensembles, i.e., the intra-experiment pattern correlation is calculated based on an ensemble of as few as three (in most cases), with only CanESM5 and IPSL-CM6A-LR having more than nine ensemble members. Supplemental Figures G.4–G.8 show the information summarized in Table 2

Table 2 Fraction of model/experiment pairs that do not differ significantly from the model’s hist-nat experiment (40 total for a. and b.; 37 for c.)

Driver	Bias				Relative std. dev.				Pattern correlation			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
(a) Mean												
ELI	0.88	0.85	0.93	0.93	0.93	0.85	0.88	0.80	0.68	0.51	0.62	0.57
AO		0.95	0.90	0.95		0.88	0.82	0.95		0.65	0.78	0.57
NAO	0.95	0.90	0.88	0.97	0.85	0.90	0.85	0.95	0.68	0.46	0.62	0.43
PNA	0.93	0.88	0.95	0.90	0.90	0.90	0.88	0.93	0.57	0.54	0.57	0.57
AMO	0.90	0.95	0.93	0.97	0.90	0.80	0.82	0.88	0.65	0.70	0.76	0.81
(b) GEV												
ELI	0.81	0.78	0.86	0.76	0.86	0.73	0.76	0.73	0.41	0.53	0.88	0.53
AO		0.86	0.89	0.95		0.78	0.78	0.81		0.79	0.91	0.97
NAO	0.89	0.89	0.86	0.95	0.92	0.73	0.84	0.73	0.68	0.65	0.82	0.82
PNA	0.92	0.76	0.76	0.95	0.86	0.73	0.70	0.78	0.32	0.68	0.68	0.71
AMO	0.84	0.89	0.97	0.78	0.78	0.76	0.78	0.78	0.56	0.91	0.88	0.74

Bias not different from zero; relative standard deviation not different from 1; inter-experiment pattern correlation does not differ from the intra-experiment pattern correlation. “Does not differ” is determined when the 95% confidence interval includes the null value. Darker colors indicate a higher proportion of pairs that do not differ. Mean/GEV refers to the OLS regression on precipitation rate and the GEV regression on Rx1Day, respectively

in a more verbose fashion, from which we can see that biases are generally centered on zero, the relative standard deviations are mostly centered on one, and the intra-experiment pattern correlations are generally as large or larger than the inter-experiment pattern correlations. This result supports our confidence assessment in Appendix E that it is *likely* (as precisely defined by Mastrandrea et al. (2010)) that the individual forcing agents do not significantly alter the driver versus seasonal precipitation relationships. Hence, we can conclude that the driven component $P_D(\cdot)$ of seasonal precipitation in Eq. 6 can be well-approximated by a linear combination of the five modes of climate variability considered here.

4.5 Isolating WMGHG and SO₂ for single-forcing vs. historical (H6)

In Sect. 4.2, we found that the externally forced component of seasonal precipitation over CONUS in the historical record can be well-approximated by two forcings, namely well-mixed greenhouse gas emissions (which generally causes an increase in precipitation) and anthropogenic aerosols (which generally cause a decrease in precipitation). The conclusion of much of the literature summarized in Sect. 1 (e.g., Min et al. 2011; Knutson and Zeng 2018) is that trends over CONUS are indistinguishable from zero. If the roughly zero signal we want to characterize, say $z(t)$, is the sum of two components $x(t)$ (WMGHGs) and $y(t)$ (anthropogenic aerosols), then this further implies $x(t) \sim -y(t)$ such

that any influence of WMGHGs on seasonal precipitation over CONUS is offset by an equal and opposite influence from anthropogenic aerosols. While a regression analysis should be able to estimate the effect of each forcing so long as WMGHGs and aerosols are sufficiently uncorrelated, it is nonetheless important to ensure that we can accurately distinguish the influence of WMGHGs from the influence of anthropogenic aerosols.

Once again, the single-forcing DAMIP experiments (hist-GHG and hist-aer) together with the CMIP6 historical runs provide a test bed for answering this question. Again using the gridcell-specific analyses for these three experiments outlined in Sect. 4.4, note that across experiments we are modeling both $x(t)$ and $y(t)$ as a constant multiplied by a time series of forcing (in the terminology of Eq. 6, this is $x(t) = \beta \text{WMGHG}^F \text{WMGHG}(t, \tau \text{WMGHG}_m)$ and $y(t) = \beta \text{SO}_2^F \text{SO}_2(t, \tau \text{SO}_2)$). Since the forcing time series are the same for the single forcing and historical runs, then if the experiment-specific estimates of the regression coefficients yield $\beta^{\text{hist-GHG}}_{\text{WMGHG}} - \beta^{\text{historical}}_{\text{WMGHG}} \sim 0$ and $\beta^{\text{hist-aer}}_{\text{SO}_2} - \beta^{\text{historical}}_{\text{SO}_2} \sim 0$ (on average) we can conclude that our estimates of the influence of WMGHGs and anthropogenic aerosols are unbiased. Figure 7 shows the CONUS-average bias for the WMGHG and SO₂ coefficients for the single-forcing experiments relative to the historical (i.e., single forcing minus historical) along with 95% basic bootstrap confidence intervals. For SO₂, the bias is indistinguishable from zero for almost every model-season; for WMGHGs, the bias is somewhat more distinguishable from

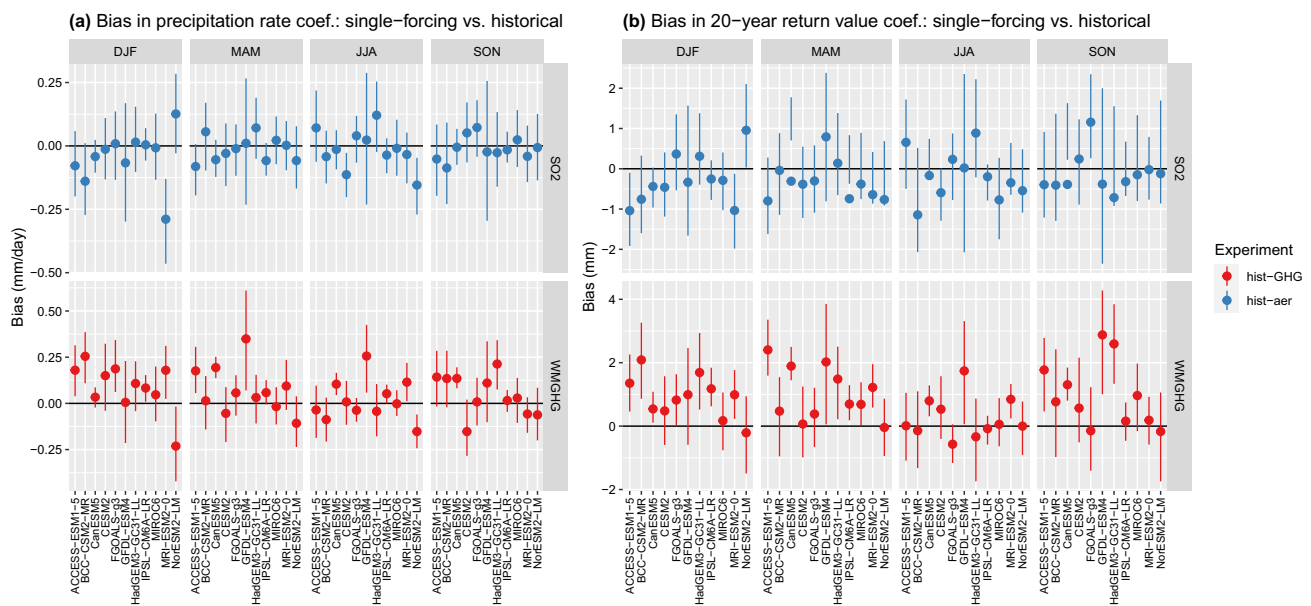


Fig. 7 CONUS-average bias in precipitation rate (left) and 20-year return value (right) coefficient estimates, both in mm/day, for the single-forcing experiments (hist-GHG and hist-aer, respectively) rela-

tive to the historical runs for WMGHG (top) and SO₂ (bottom), along with 95% basic bootstrap confidence intervals

zero (in that two to five of the models have confidence intervals that do not overlap zero), but in most cases the confidence intervals include zero. For WMGHGs, when the bias is nonzero, it is generally positive, meaning that we (on average) underestimate the influence of WMGHGs in the historical record. Nonetheless, given this general agreement between the single-forcing and historical experiments, we can indeed distinguish the influence of WMGHGs from the influence of anthropogenic aerosols.

The same coefficient estimates can be further interrogated to assess whether the single-forcing vs. historical biases for WMGHGs are correlated with the SO_2 biases. Anti-correlation in the grid-cell WMGHG and SO_2 bias would imply that there is a compensating trade off between our estimates of the effect of these two forcings. We seek to minimize equal and opposite errors $\epsilon(t)$ in $x(t) = \beta \text{WMGHG}^F \text{WMGHG}(t, \tau \text{WMGHG})$ and $y(t) = \beta \text{SO}_2^F \text{SO}_2(t, \tau \text{SO}_2)$ such that the fit to the observed precipitation time series is unperturbed, yet yields biased estimates of βWMGHG and βSO_2 , i.e.,

$$\begin{aligned} x(t) + y(t) &\rightarrow x(t) + y(t) + [\epsilon(t) - \epsilon(t)] \\ &= [x(t) + \epsilon(t)] + [y(t) - \epsilon(t)] \\ \beta \text{WMGHG} &\rightarrow \frac{x(t) + \epsilon(t)}{F \text{WMGHG}(t, \tau \text{WMGHG})} \quad \text{and} \\ \beta \text{SO}_2 &\rightarrow \frac{y(t) - \epsilon(t)}{F \text{SO}_2(t, \tau \text{SO}_2)} \end{aligned}$$

To answer this question, we calculate the Pearson correlation of the grid-cell SO_2 and WMGHG biases as well as a 95% bootstrap confidence interval; see Fig. 8. Across a large majority of the model/season/precipitation type, the confidence intervals for the correlation in biases includes zero. This result supports our confidence assessment in Appendix E that it is *likely* (as precisely defined by Mastrandrea et al. (2010)) that the compensating errors in the effects of WMGHG and anthropogenic aerosol forcing are negligible.

4.6 State-dependence of signal-to-noise ratio (H7)

Our last test concerns the signal-to-noise ratio of seasonal precipitation, i.e., $1 - \frac{\text{Var } P_w(t)}{\text{Var } P(t)}$, which we have assumed does not depend on the background state of the climate system. Note that we make no specific claims about the variability of the weather term $P_w(t)$, nor do we make any claims about the total year-to-year variability of the entire time series $P(t)$. Rather, this concerns the ratio of these variances, specifically over CONUS. This question is of particular importance because the uncertainty due to weather and fast internal variability, i.e., $\text{Var } P_w(t)$, is the largest term (by far) in the time series: Risser et al. (2021) find that the noise of this term accounts for 90–95% (or more) of the variance in seasonal mean and extreme precipitation over CONUS. So, if the proportion of variance due to fast internal variability increases even slightly with warming, this reduces the magnitude of a signal that is minimal to begin with.

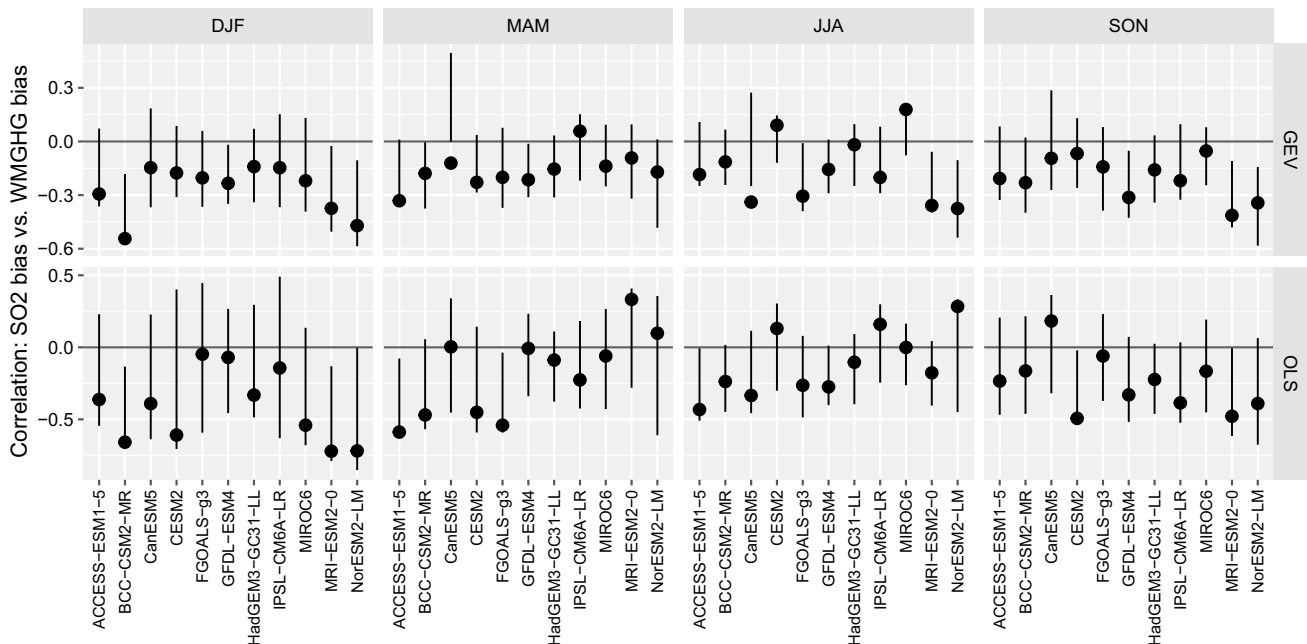


Fig. 8 Pearson correlations of the single-forcing versus historical bias in the SO_2 coefficient against the same quantity for the WMGHG coefficient across model grid cells with 95% basic bootstrap confidence intervals

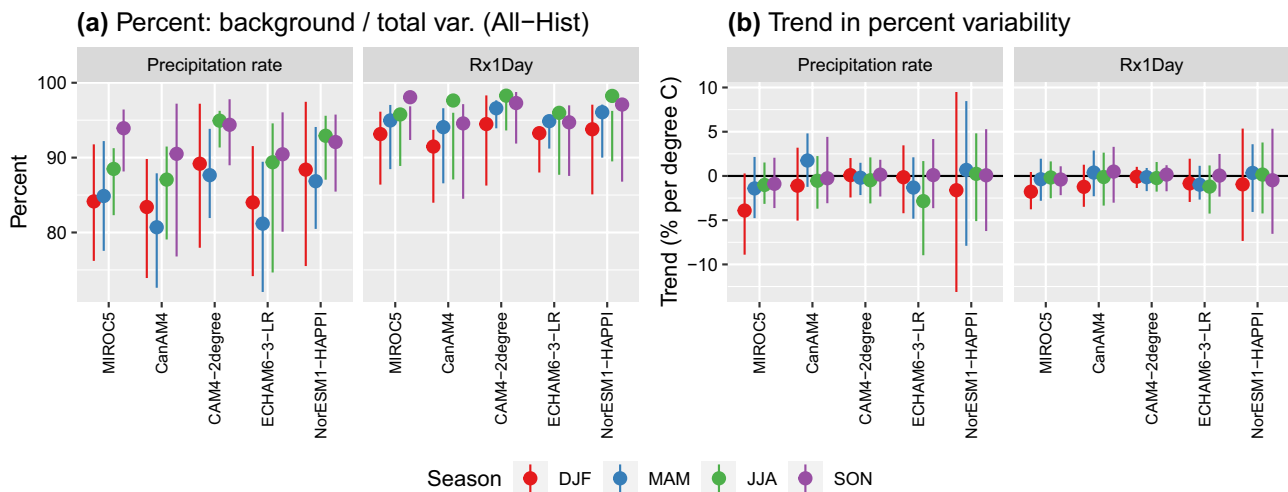


Fig. 9 Percent of background to total variability for each model and season in the All-Hist experiment (a) along with trends in the percent background variability per °C (b) calculated from the All-Hist and HAPPI simulations, with 95% bootstrap confidence intervals

Table 3 A summary of the various hypotheses examined in this paper, along with the model data sets used, our conclusion (“Conclusion”), and a confidence statement (“Confidence”)

	Hypothesis	Model data sets used	Conclusion	Confidence
H1	The WMGHG forcing is consistent with $\approx 2\%/K$ (for means) and $\approx 6\%/K$ (for extremes).	DAMIP (hist-GHG); CMIP6 piControl and 1pctCO2	Yes	<i>Likely</i>
H2	All cross-correlation terms in the forced component of seasonal precipitation are negligible.	DAMIP (hist-GHG, hist-aer); CMIP6 historical	Yes*	<i>Very likely</i>
H3	Individual forcings are non-negligible for secular trends in precipitation.	DAMIP (hist-GHG, hist-aer, hist-nat, hist-stratO3); LUMIP (hist-noLu)	Yes* and no	<i>Likely</i> (GHG, aer); <i>About as likely as not</i> (stratO3, LULCC); <i>Unlikely</i> (nat)
H4a	SO ₂ is the dominant aerosol species for changes in precipitation.	AerChemMIP (piClim experiments)	Yes	<i>Likely</i>
H4b	emiso2 correlates with changes in precip. as well as wetso2 and iso2. [†]	AerChemMIP (piClim experiments)	Yes	<i>Very likely</i>
H4c	Regionalized time series of emiso2 yield the same results as using either local or CONUS-wide estimates.	AerChemMIP (histSST)	Yes	<i>Virtually certain</i>
H5	All interaction terms between climate drivers and external forcing agents are negligible.	DAMIP (hist-GHG, hist-aer, hist-nat, hist-stratO3); LUMIP (hist-noLu); CMIP6 historical	Yes*	<i>Likely</i>
H6	Compensating errors in WMGHG and aerosol forcing are negligible.	DAMIP (hist-GHG, hist-aer); CMIP6 historical	Yes*	<i>Likely</i>
H7	The signal-to-noise ratio of seasonal precipitation is a constant with respect to warming.	C20C+ All-Hist; HAPPI Plus15- and Plus20-Future	Yes	<i>Virtually certain</i>

All statements and conclusions are strictly limited to the continental United States for the pre-industrial period to present day; the confidence labels reflect the categories defined by (Mastrandrea et al. 2010). See Appendix E for more information on how each confidence statement is determined

[†] Note: “emiso2” refers to SO₂ emissions; “wetso2” refers to deposition of SO₂ by rainfall; “iso2” refers to column integrated SO₂ mass

* Indicates cases where we fail to reject the null hypothesis versus cases where results are conclusive

To address this hypothesis, we can utilize the C20C+ and HAPPI experiments described in Sect. 3.1. The aforementioned analysis in Risser et al. (2021) utilized a large ensemble from the C20C+ All-Hist experiment; we repeat the analysis here with both the C20C+ All-Hist and HAPPI Plus15- and Plus20-Future. Specifically, for the five models (Table F.5 in Supplementary material) that provide at least 50 members over at least 10 years for each experiment, we calculate three quantities: the background variability, calculated as the across-year average of the within-year ensemble variance (denoted V_B in Risser et al. 2021); the total variability, calculated as the ensemble average of the across-year variance (denoted V_T in Risser et al. 2021); and the proportion of background vs. total variance. All quantities are first calculated for each CONUS grid cell and then averaged; uncertainty in the background (total) variability is quantified by bootstrapping years (ensemble members). The primary justification for comparing the intra-ensemble variability to weather noise is that the ensembles in each of the C20C+ and HAPPI experiments have identical upper (incoming solar) and lower (SST) boundary conditions as well as identical radiative forcings. The identical lower boundary conditions imply that long-term modes of internal variability (e.g., ENSO) that manifest in SST are also identical. Since the forcings, upper and lower boundary conditions, and long-term SST-driven modes of variability are the same across each ensemble, we can confidently postulate that the intra-ensemble variability is driven by the $\mathcal{O}(\epsilon) \ll 1$ (i.e., limit of machine precision) perturbations added to the SST at the time $t = 0$ for each ensemble member. These perturbations result in rapid solution separation (Zhang et al. 2019) and variance across the ensemble in instantaneous meteorological states, which we treat as weather noise (Risser et al. 2021).

Figure 9 shows the percent of background variability across models and in each season for the All-Hist experiment as well as trends in the percent variability with warming, calculated from the All-Hist (+0.85°C), Plus15-Future (+1.55°C), and Plus20-Future (+2°C) experiments; Supplemental Figure G.9 shows corresponding plots for the total and background variability. For the models considered here, the percent background variability is roughly 80–95% for mean and 90–98% for extreme precipitation, which is consistent with what Risser et al. (2021) found for a corresponding CAM5.1-1degree ensemble. Across mean and extreme precipitation, there are indeed increases in both the background and total variability with warming (see Supplemental Figure G.9) which are in almost all cases significant. The general increases in both background and total variability, however, translate to roughly consistent *proportion* of background variability: in Fig. 9b the trends in the percent of background variability are indistinguishable from zero in all models, seasons, and for both mean/extreme precipitation. This result

supports our confidence assessment in Appendix E that it is *virtually certain* (as precisely defined by Mastrandrea et al. (2010)) that the signal-to-noise ratio $1 - \text{Var } P_w(t)/\text{Var } P(t)$ is unaffected by warming.

5 Discussion

In this paper, we have described novel methodology to address the difficult problem of conducting regional detection and attribution for mean and extreme precipitation. Our framework can simultaneously account for anthropogenic forcing (the “signal”) while explicitly modeling the natural variability (the “noise”). Furthermore, we use output from global climate models in a perfect data sense to rigorously test the application of a general formula for modeling a time series of precipitation. Modeling uncertainty is well-known to be one of the main obstacles to conclusive attribution statements for regional precipitation, but here we turn this limitation into an opportunity by making sure the D&A methodology developed in this paper is insensitive to model uncertainty. Our focus in this paper has been on the CONUS due to the relatively low observational uncertainty, with a forthcoming paper applying our framework to century-length records of in situ measurements of seasonal precipitation. However, in principle, the series of hypotheses enumerated in this paper could be similarly applied to other global land regions to develop an appropriate formula for regional D&A of precipitation.

The analyses described in this paper are extensive and involve a large set of global climate model output. We have already attempted to summarize the scientific process that led to the development and testing of each hypothesis with the flow chart shown in Fig. 1. To further summarize our conclusions, we developed Table 3, which restates the hypotheses, the model data sets used for each hypothesis, our conclusions, and a measure of our confidence in each conclusion. Table 3 also denotes the hypotheses for which we failed to reject the null versus cases where we have definitively verified a conclusion. The confidence statement attempts to summarize the evidence for each hypothesis, albeit in a highly aggregated manner; nonetheless, we intend that a high-level statement for each of the lengthy analyses in the paper will provide a helpful and concise summary of our conclusions. Specific information on how these labels are determined is provided in Appendix E. The specific confidence labels (e.g., *likely*, *very likely*, etc.) reflect the categories defined by the Intergovernmental Panel on Climate Change (Mastrandrea et al. 2010). We reiterate that in this work the categories do not summarize a multi-expert assessment but instead provide a qualification of evidence obtained from the model analysis.

Traditional D&A techniques are based on the concept of Pearl causality (Hannart et al. 2015), which is applicable when using model-based counterfactual methods where an “intervention” can be applied and assessed. For example, the C20C+ D&A experiment explores pairs of experiments: one set includes anthropogenic and natural forcings, and the other includes natural forcings only. In this case, the “intervention” is anthropogenic forcing. When observations can be successfully projected onto the anthropogenically-forced experiments, the reliance on global climate models is a strength because a clear causal statement can be made regarding the presence of human-induced climate change. However, when the goal is regional D&A for precipitation, the large internal variability combined with large model uncertainty (to say nothing of observational uncertainty) essentially prevents conclusive attribution statements regarding any anthropogenic influence. Of course, this may change in the future as models improve or the anthropogenic signal intensifies. Furthermore, there could be value in being selective with respect to which GCMs are used: first, one might validate a set of models for their ability to simulate salient features of the precipitation regime governing a regional hydroclimate (e.g., Gleckler et al. 2008) and rely on the most realistic models for D&A. Meanwhile, the methods identified in this paper incorporate the guidance of Sarojini et al. (2016) and are designed to maximize our usage of existing models and data sets. Our philosophy in this work uses climate models to derive and test a formula for regional D&A, but we plan to then set models aside and construct maps of attributable changes in precipitation using observations only. Such attribution statements use the concept of Granger causality (Granger 1969), which is complementary to more traditional techniques that use Pearl causality. Papers such as Risser and Wehner (2017) successfully made such Granger-causal statements for localized extreme precipitation over the Houston, Texas region in the aftermath of Hurricane Harvey. Granger-causal attribution statements are not as controlled as Pearl-causal statements since they do not involve a set of interventions, but such analyses can both motivate dynamical studies (e.g., the Hurricane Harvey study provided motivation for Patricola and Wehner 2018; Wang et al. 2018) and enhance confidence in attribution statements by using multiple analysis techniques to explore the causes of climate change. And, when climate models prove to be an insufficient tool for Pearl-causal attribution (as in the case of regional precipitation), Granger-causal statements are an important tool for both quantifying the local impacts of climate change on the water cycle and informing methodological developments that utilize Pearl causality. We note that Granger-causal statements are subject to the usual caveats regarding the effect of hidden covariates (i.e., “correlation does not imply causation”). However, using both Pearl and Granger causality together as we have done here removes

some of the usual Granger limitations: the correlation identified in an observational analysis using the framework developed in this paper may indeed be causation.

A. Accounting for time lags in the response of precipitation to WMGHG forcing

A1. Lagged effects of forcing on precipitation

As Samset et al. (2016) showed, the response of global precipitation to forcing by CO₂ is dominated by a slow response associated with the effects of warming SSTs. SSTs have a lagged response to instantaneous WMGHG forcing due to thermal inertia of the oceans, which have absorbed over 90% of the excess energy gained by the climate system (Resplandy et al. 2019) and have a $\mathcal{O}(10^3)$ higher heat capacity than the atmosphere. To include the lagged response of precipitation to instantaneous WMGHG forcing in our general D&A formula, let us hypothesize that the regional precipitation $P(t)$ tracks changes in global mean SST $\delta T(t)$ as a function of time t and that lags in the precipitation response are due to time lags in the response of the global temperature:

$$\begin{aligned} \frac{P(t)}{P_0} - 1 &= R \gamma \delta T(t) \\ P_0 &= P(t=0) \\ R &= \text{Unitless ratio of regional-to-global precipitation} \\ \gamma &= \text{Global precipitation sensitivity:} \\ &2\%/\text{K for mean and } 6\%/\text{K for extremes} \end{aligned} \quad (11)$$

The time evolution of $\delta T(t)$ is dictated by the first law of thermodynamics applied to the global ocean:

$$\begin{aligned} \rho c_p h \frac{d \delta T(t)}{dt} &= F(t) - \alpha \delta T(t) \\ \rho &= \text{Density of ocean water} \\ c_p &= \text{Specific heat at constant pressure of ocean water} \\ h &= \text{Depth of ocean} \\ F(t) &= \text{Climate forcing} \\ \alpha &= \text{Climate feedback parameter } [W/m^2/K] \end{aligned}$$

Define a timescale $\tau = \rho c_p h / \alpha$. Then the first law can be written as:

$$\frac{d \delta T(t)}{dt} + \frac{\delta T(t)}{\tau} = \frac{F(t)}{\rho c_p h} \quad (12)$$

The solution is:

$$\delta T(t) = \alpha^{-1} F(t, \tau) \quad (13)$$

$$F(t, \tau) = \int_0^t e^{(t'-t)/\tau} F(t') \frac{dt'}{\tau} \quad (14)$$

In the limit $\tau = 0$, one can show that, as expected from Eq. 12 when the system is in equilibrium,

$$\delta T(t) = F(t)/\alpha$$

This follows from Eqs. 13 and 14 if we define an integration variable $x = (t - t')/\tau$ so that

$$\begin{aligned} \delta T(t) &= \alpha^{-1} \lim_{\tau \rightarrow 0} \int_0^{t/\tau} e^{-x} F(t - \tau x) dx \\ &= \alpha^{-1} \lim_{\tau \rightarrow 0} \int_0^{t/\tau} e^{-x} \left[F(t) + \sum_{n=1}^{\infty} \frac{1}{n!} \frac{d^n F(t)}{dt^n} (-\tau x)^n \right] dx \\ &= \frac{F(t)}{\alpha} \lim_{\tau \rightarrow 0} \int_0^{t/\tau} e^{-x} dx \\ &= \frac{F(t)}{\alpha} \lim_{\tau \rightarrow 0} e^{-x} \Big|_0^{t/\tau} \\ &= \frac{F(t)}{\alpha} \end{aligned}$$

Substituting Eq. 13 into Eq. 11 gives an expression for the β coefficients that appear, for example, in Eq. 2 for the forced component:

$$\frac{P(t)}{P_0} - 1 = \beta(\tau) F(t, \tau) \quad (15)$$

$$\beta(\tau) = R(\tau) \frac{\gamma}{\alpha} \quad (16)$$

where we have introduced the possibility that β might exhibit dependence on τ .

The lagged time series $F(t, \tau)$ calculated from Eq. 14 using the WMGHG forcing time series $F(t)$ from Fig. 2 is shown in Fig. 10 for $\tau = 1, \dots, 20$.

A2. Decreasing lagged forcing with increasing lag timescales

The variation of the lagged forcing with lag timescale can be established by computing

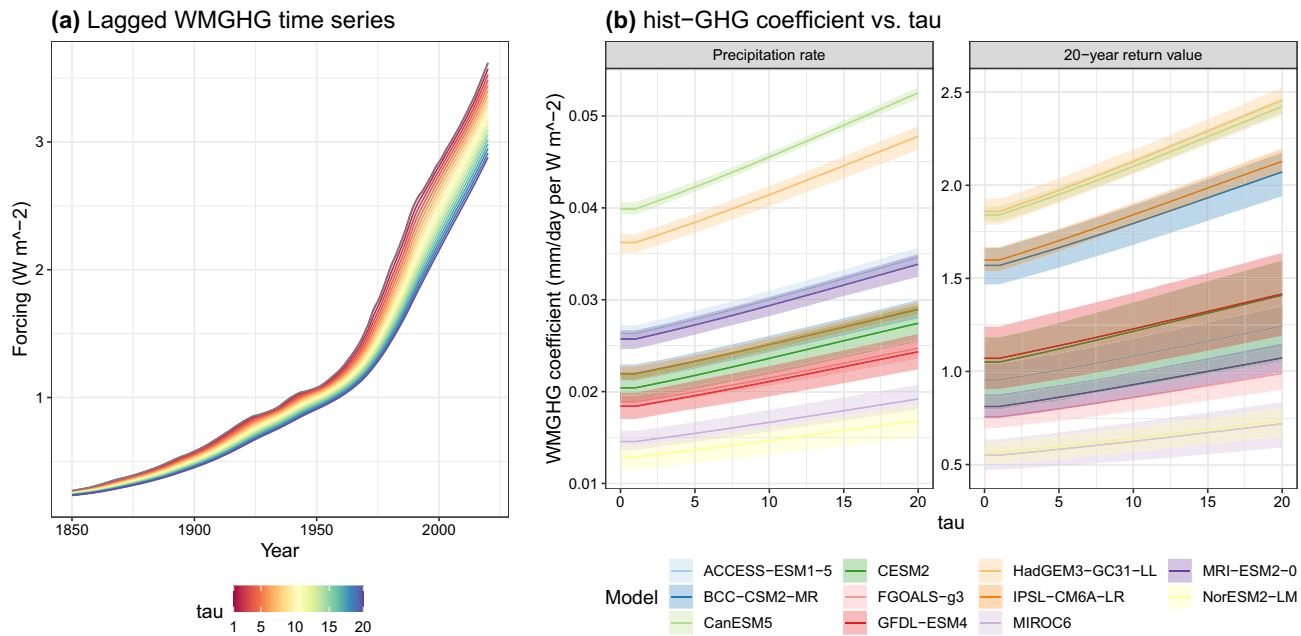


Fig. 10 The lagged WMGHG forcing $F(t, \tau)$ calculated from Eq. 14 for $\tau \in \{1, \dots, 20\}$ (a), as well as the estimated WMGHG regression coefficient $\beta(\tau)$ for global average annual precipitation rate as a function of the time lag τ (b). Estimates in b include 95% basic bootstrap

confidence intervals. Note that a shows the decreasing lagged forcing with increasing lag as derived in Eq. 17 and b shows the increasing trend in $\beta(\tau)$ with increasing lag derived in Eq. 20.

$$\begin{aligned}
\frac{dF(t, \tau)}{d\tau} &= \alpha^{-1} \underbrace{\frac{d}{d\tau} \int_0^{t/\tau} e^{-x} F(t - \tau x) dx}_A \\
&= \alpha^{-1} \underbrace{\left\{ \left(-\frac{t}{\tau^2} \right) \left[e^{-x} F(t - \tau x) \right] \Big|_{x=t/\tau} \right\}}_A \\
&\quad + \alpha^{-1} \underbrace{\left\{ \int_0^{t/\tau} e^{-x} \left[\frac{dF(y)}{dy} \frac{dy}{d\tau} \right] \Big|_{y=t-\tau x} dx \right\}}_B \\
&= \alpha^{-1} \underbrace{\left\{ \left(-\frac{t}{\tau^2} \right) \left[e^{-t/\tau} F(0) \right] \right\}}_A + \alpha^{-1} \\
&\quad \underbrace{\left\{ \int_0^{t/\tau} e^{-x} \left[\frac{dF(y)}{dy} \Big|_{y=t-\tau x} (-x) \right] dx \right\}}_B
\end{aligned}$$

Note that term A is negative indefinite assuming $F(0) \geq 0$. Assuming $F(t)$ is monotonically increasing, i.e. $dF(t)/dt \geq 0 \forall t \geq 0$, then term B is also negative indefinite. Therefore

$$\frac{dF(t, \tau)}{d\tau} \leq 0. \quad (17)$$

If $P(t)$ is prescribed and the forcing obeys the assumptions used to obtain Eq. 17, then Eqs. 15 and 17 imply that

$$\frac{d}{d\tau} \left[\frac{P(t)}{P_0} - 1 \right] = 0 = \frac{d\beta(\tau)}{d\tau} F(t, \tau) + \beta(\tau) \frac{dF(t, \tau)}{d\tau} \quad (18)$$

$$\frac{1}{\beta(\tau)} \frac{d\beta(\tau)}{d\tau} = - \frac{1}{F(t, \tau)} \frac{dF(t, \tau)}{d\tau} \quad (19)$$

$$\therefore \frac{d \ln \beta(\tau)}{d\tau} \geq 0 \quad (20)$$

The positive trend in $\beta(\tau)$ with respect to τ is observed the DAMIP WMGHG-only runs (Fig. 10).

A3. Estimation of lag timescale using Gregory regression plots

Suppose $F(t) = F_0$, i.e., a constant. This is the case in the abrupt4xCO2 runs, which we analyze to estimate τ . Then the solution for $\delta T(t)$ is

$$\delta T(t) = \frac{F_0}{\alpha} (1 - e^{-t/\tau})$$

This implies that the net energy input to the climate system, evaluated at the top of the atmosphere (TOA), which is just the RHS of the 1st law of thermodynamics, is

$$\begin{aligned}
N(t) &= F(t) - \alpha \delta T(t) \\
&= F_0 e^{-t/\tau}
\end{aligned}$$

$$\therefore \ln N(t) = \ln F_0 - \frac{t}{\tau}$$

This is indeed the behavior that's observed in climate models for $t \lesssim 20$ years in Gregory regression plots of $N(t)$ versus $\delta T(t)$ (Gregory et al. 2004; Meehl et al. 2020). Then the models typically shift to a much slower rate of change.

The time-series of $\ln N_i(t)$ from the abrupt4xCO2 MME indexed by i for the various models and ensemble members is computed via

$$N_i(t) = \text{rsdt}_i(t) - \text{rsut}_i(t) - \text{rlut}_i(t) \quad (21)$$

$$\text{rsdt}_i(t) = \text{TOA incident shortwave radiation} \quad [W/m^2] \quad (22)$$

$$\text{rsut}_i(t) = \text{TOA outgoing shortwave radiation} \quad [W/m^2] \quad (23)$$

$$\text{rlut}_i(t) = \text{TOA outgoing longwave radiation} \quad [W/m^2] \quad (24)$$

where the radiative fluxes on the RHS of Eq. 21 have been globally and annually averaged and are denoted by the names of CMIP6 output variables (Centre for Environmental Data Analysis (CEDA) 2016). We have determined three timescales $\tau_{f,i}$, $\tau_{s,i}$, and $\tau_{b,i}$ characterizing the “fast” and “slow” evolution of $N_i(t)$ and the “break” between these two regimes as follows:

$$\ln N_i(t) = \ln F_0 - \frac{t}{\tau_{f,i}} \Theta(\tau_{b,i} - t) - \frac{t}{\tau_{s,i}} \Theta(t - \tau_{b,i}) \quad (25)$$

where Θ is the Heaviside function and $F_0 = 8.053$ is determined from the formulae in Etminan et al. (2016).

We use a Bayesian least squares fit to estimate the timescales $\tau_{f,i}$, $\tau_{s,i}$, and $\tau_{b,i}$ in Eq. 25 using the $N_i(t)$ time series from each GCM with abrupt4xCO2 ensembles in Table F.2 in Supplementary material, aggregating over multiple ensemble members when available. A Bayesian framework is useful for ensuring that the timescales are consistent, i.e., $\tau_{b,i} < \tau_{f,i}$ and $\tau_{s,i} < \tau_{f,i}$, as well as quantifying uncertainty in the break point $\tau_{b,i}$. Also note that the intercept $\ln F_0$ is fixed at its theoretical value of 8.053. To illustrate the analysis, we show results from the two-member CanESM5 ensemble and six-member MRI-ESM2-0 ensemble; see Fig. 11. The natural log of the net radiative input in each year is shown along with the estimated timescales $\tau_{f,i}$, $\tau_{s,i}$, and $\tau_{b,i}$. In general, the timescale τ_i used to calculate the lagged WMGHG forcing time series $F(t, \tau_i)$ is set equal to $\tau_{f,i}$ if $\tau_{f,i} < 33$ (the

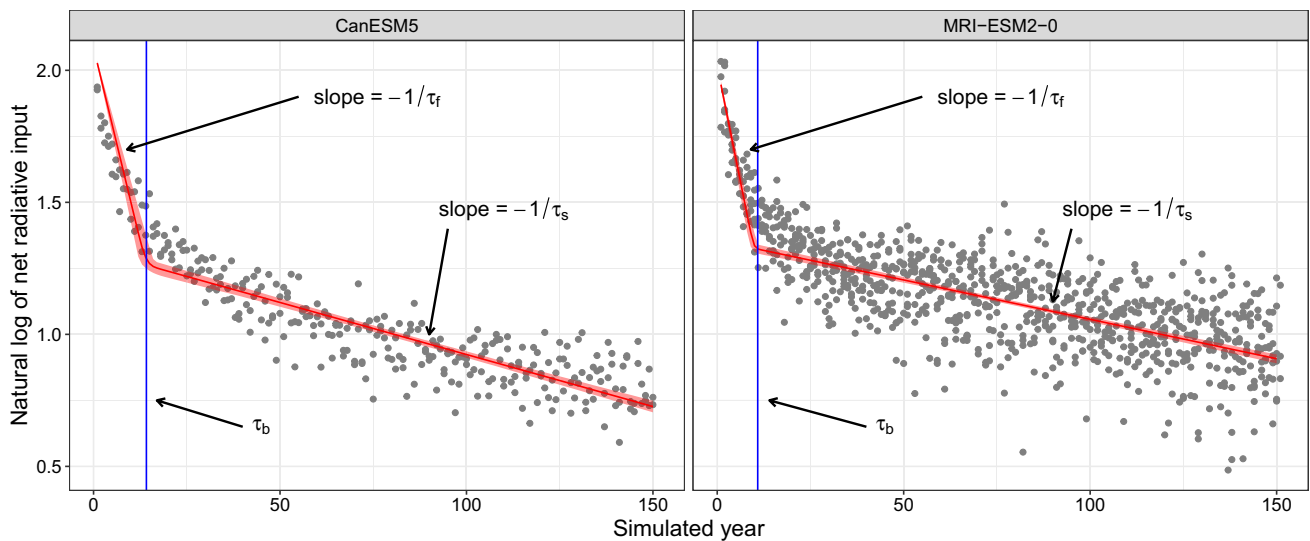


Fig. 11 Estimated timescales $\tau_{f,i}$, $\tau_{s,i}$, and $\tau_{b,i}$ for the two-member CanESM5 ensemble (left) and six-member MRI-ESM2-0 ensemble (right). The red line shows the least-squares fit described by Eq. 25 with a 95% uncertainty band and the blue line denotes the breakpoint $\tau_{b,i}$

90th percentile of the multi-model ensemble) and $\tau_{b,i}$ otherwise. For these two models, we have $\tau_i = 14$ for CanESM5 and $\tau_i = 11$ for MRI-ESM2-0. The estimated timescales for all other CMIP6 models with abrupt4xCO2 data are shown in Table 4. For models without abrupt4xCO2 data (e.g., FGOALS-g3), we use the CMIP6 multi-model ensemble average of $\bar{\tau} = 14$.

B. Test of H1 under significantly increased warming

To assess our test of H1 under much greater levels of warming than is imposed on the CMIP6 historical experiment, we can utilize climate model runs from the CMIP6 DECK experiments piControl and 1pctCO2 introduced in Sect. 3.1. For each CMIP6 model considered, we first extract the global average of mean daily precipitation and Rx1Day as well as monthly surface temperature for each year in both the piControl and 1pctCO2 runs. Furthermore, we derive the time series of WMGHG forcing for the 1pctCO2 runs using the specified boundary conditions for the 1pctCO2 experiment and various formulae described in Sect. 3.2. We then estimate the changes in precipitation (mean and extreme) using both the raw values of surface temperature and a regression versus WMGHG forcing, for each of two formulations. In formulation 1, the baseline temperature and precipitation come from the last 50 years of the piControl run; then, for year $y = 50, 51, \dots, N_{yr}$ of each 1pctCO2 run (where N_{yr} is the number of years in the simulation), we

compare the precipitation and temperature measurement in year y with the baseline (here, piControl) precipitation and temperature values to estimate the change in precipitation ΔP (%), change in temperature δT (K), and scaling factor $\Delta P/\delta T$. In formulation 2, we instead use a moving window approach, wherein year $y - 49$ is the baseline year; in other words, we look at successive 50-year subsets of the 1pctCO2 run. For both of these formulations, we further calculate the precipitation totals in two ways: first, using the “raw” values, i.e., the annual precipitation rate and Rx1Day in the baseline and comparison years; second, we regress the temperature, precipitation rate, and Rx1Day versus the WMGHG forcing (temperature and precipitation rate via ordinary least square; Rx1Day using a GEV distribution as in Eqs. 7 and 8). In the second approach, the starting and ending precipitation/temperature value come from the fitted regression analysis. The results are shown in Fig. 12, where each point represents a global climate model / ending year and the points from an individual GCM are plotted with a line of best fit. Panels a and b show change in precipitation rate versus warming and scaling versus warming, respectively; panels c and d show the corresponding quantities for 20-year return values. From Fig. 12a, c, it is clear that changes in both mean and extreme precipitation globally are linear in warming, as expected, whether the precipitation/temperature changes are calculated using raw values or a regression versus WMGHG. Interestingly, the scaling relationships are quite noisy with respect to the level of warming, and in many cases do not show large increases with warming.

Table 4 Central estimates of model-specific timescales τ_f , τ_b , and τ_s , estimated from the CMIP6 abrupt4xCO2 experiment, as well as the timescale τ used to calculate the lagged WMGHG forcing time series $F(t, \tau)$

Model	τ_f	τ_b	τ_s	τ	Model	τ_f	τ_b	τ_s	τ
ACCESS-CM2	16	14	338	16	GFDL-ESM4	9	11	321	9
ACCESS-ESM1-5	8	10	268	8	GISS-E2-1-G	8	9	363	8
AWI-CM-1-1-MR	12	15	202	12	GISS-E2-1-H	9	10	199	9
BCC-CSM2-MR	7	10	113	7	GISS-E2-2-G	6	8	344	6
BCC-ESM1	7	10	179	7	HadGEM3-GC31-LL	15	12	256	15
CAMS-CSM1-0	12	12	250	12	HadGEM3-GC31-MM	13	9	227	13
CanESM5	17	14	255	17	IITM-ESM	14	12	447	14
CESM2	17	12	373	17	INM-CM4-8	39	17	744	17
CESM2-FV2	11	9	494	11	INM-CM5-0	11	10	408	11
CESM2-WACCM	11	10	300	11	IPSL-CM5A2-INCA	8	8	248	8
CESM2-WACCM-FV2	11	11	489	11	IPSL-CM6A-LR	18	17	306	18
CIESM	82	29	496	29	IPSL-CM6A-LR-INCA	14	13	283	14
CMCC-CM2-HR4	24	15	422	24	KACE-1-0-G	25	14	431	25
CMCC-CM2-SR5	32	20	554	32	KIOST-ESM	8	9	290	8
CMCC-ESM2	40	20	554	20	MIROC-ES2L	16	14	557	16
CNRM-CM6-1	24	13	455	24	MIROC6	12	11	673	12
CNRM-CM6-1-HR	40	21	390	21	MPI-ESM-1-2-HAM	14	14	324	14
CNRM-ESM2-1	19	11	472	19	MPI-ESM1-2-HR	11	12	255	11
E3SM-1-0	17	21	237	17	MPI-ESM1-2-LR	13	12	287	13
EC-Earth3	11	12	131	11	MRI-ESM2-0	14	11	336	14
EC-Earth3-AerChem	12	14	127	12	NorCPM1	11	11	270	11
EC-Earth3-CC	10	11	127	10	NorESM2-LM	6	9	537	6
EC-Earth3-Veg	11	13	128	11	NorESM2-MM	8	13	561	8
FIO-ESM-2-0	40	16	438	16	TaiESM1	14	12	278	14
GFDL-CM4	9	9	356	9	UKESM1-0-LL	15	13	245	15

The timescale τ is set equal to τ_f if $\tau_f < 33$ (the 90th percentile of the MME) and τ_b otherwise. For models without abrupt4xCO2 data (e.g., FGOALS-g3—see Table F.2 in Supplementary material), we use the CMIP6 multi-model ensemble average of $\bar{\tau} = 14$

C. Correlations between WMGHGs and SO₂ emissions for 1900-present

As was mentioned in Sect. 4.3.3, there are very strong correlations between the SO₂ emissions time series and the annual WMGHG time series across CONUS for 1900-present when considering the individual $0.5^\circ \times 0.5^\circ$ grid cell measurements of SO₂ emissions (Hoesly et al. 2018; Gidden et al. 2018). On the other hand, the CONUS-wide sum-total emissions time series over the same time period has a much weaker correlation with the WMGHG time series. We now provide additional details to explain this somewhat counter-intuitive result. In the next few paragraphs, we focus on the annual SO₂ emissions time series, although similar results hold for the seasonal time series used for the analysis in Sect. 4.3.3.

The average annual SO₂ emissions at each grid cell are shown in Fig. 13a, d for 1850–2015 and 1900–2015, respectively. Even with the color bar shown on the logarithmic scale, it is clear that the typical SO₂ emissions show significant heterogeneity across CONUS: in most of the eastern

U.S. (east of 95°W), there are moderate emissions, with particularly strong emissions at locations that presumably include an industrial center. Alternatively, in the rural areas of the central U.S., the emissions are much smaller (by a factor of > 100), particularly across large swaths of the Northern Great Plains. Indeed, this heterogeneity is what led us to initially consider using spatially resolved local estimates of SO₂ emissions in the analyses throughout this paper.

However, recall that observational limitations prevent us from having reliable measurements of daily precipitation before 1900, and so our subsequent analysis of the weather station data is constrained to 1900-present. Unfortunately, the grid cell correlations between SO₂ emissions and WMGHGs for 1900-present are generally quite strong (most often anticorrelated; see Fig. 13e), with the correlation being larger than 0.8 in absolute value for approximately 50% of the CONUS grid cells. Such strong (anti) correlation renders a regression-based analysis essentially useless for distinguishing the effect of WMGHG on precipitation from the effect of SO₂ when using local estimates of SO₂ emissions. On the other hand, the length of the AerChemMIP

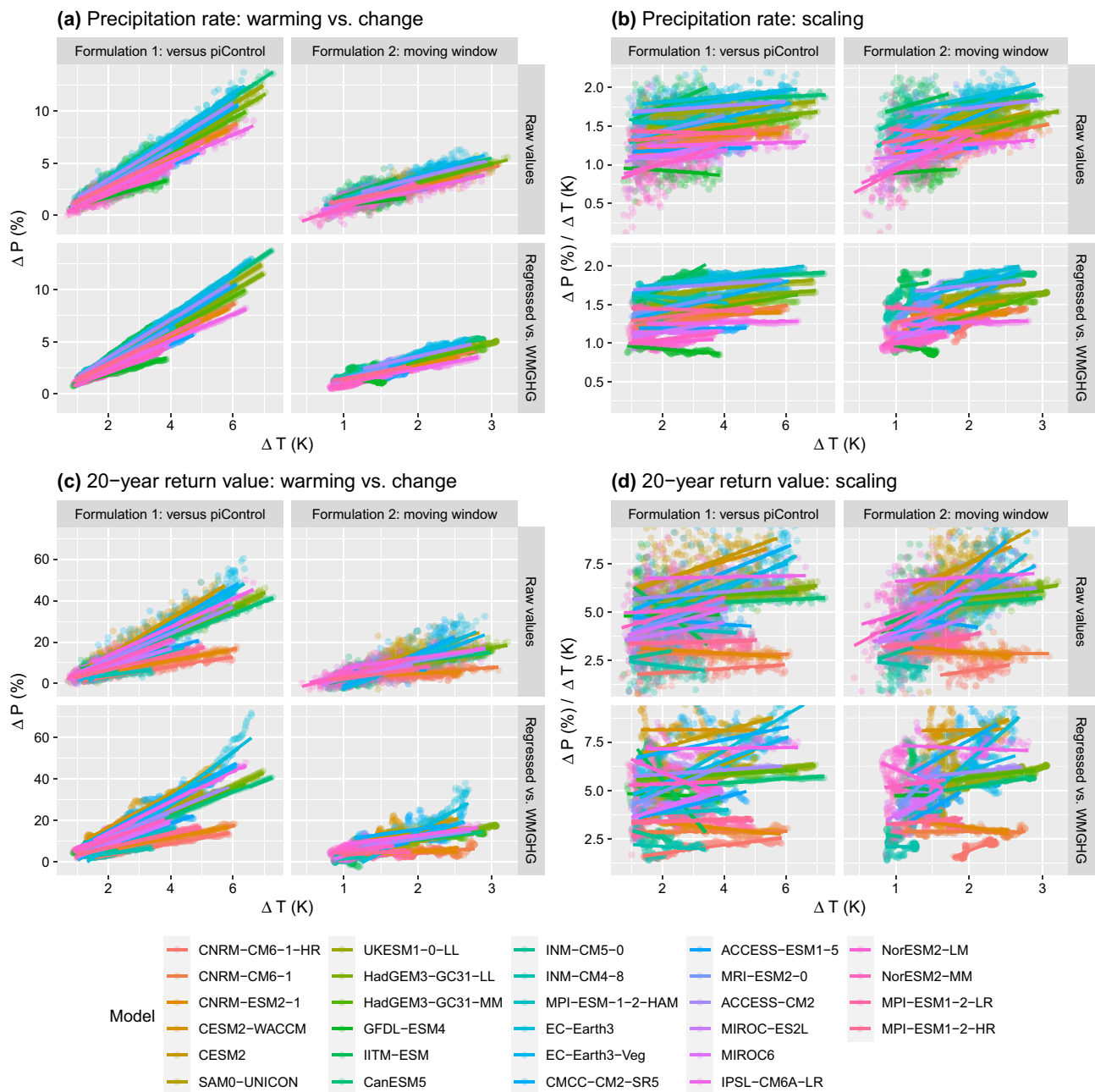


Fig. 12 Estimated changes in annual precipitation (left) and corresponding scaling (right) versus warming (precipitation rate, top; Rx1Day, bottom) for the 1pctCO₂ experiment, where the changes in precipitation and temperature are calculated directly from the start/

end year (“raw values”) and from a regression versus WMGHG forcing. All panels show two formulations, where the starting values come from the piControl runs (formulation 1) and also using a 50 year moving-window (formulation 2)

histSST runs allow us to analyze the entire 1850–2014 time period. For this longer period (1850–present), note that the grid cell estimates of SO₂ emissions have much weaker correlations with the WMGHG time series (Fig. 13b), generally less than ± 0.2 for roughly 75% of the CONUS grid cells. These diminished correlations at the grid cell level present no problems for a regression-based analysis. Figure 13c, f show scatter plots of the quantities mapped in the left and

middle columns, along with an indication of the correlation between the WMGHG time series and the CONUS-wide SO₂ emissions (approximately 0.4 for 1850–2015 and -0.2 for 1900–2015), where we can see that the strong anticorrelations in Fig. 13e tend to correspond to grid cells with small average emissions. This fact explains the counter-intuitive result wherein the grid cell correlations between SO₂ emissions and WMGHGs can be so strong while the

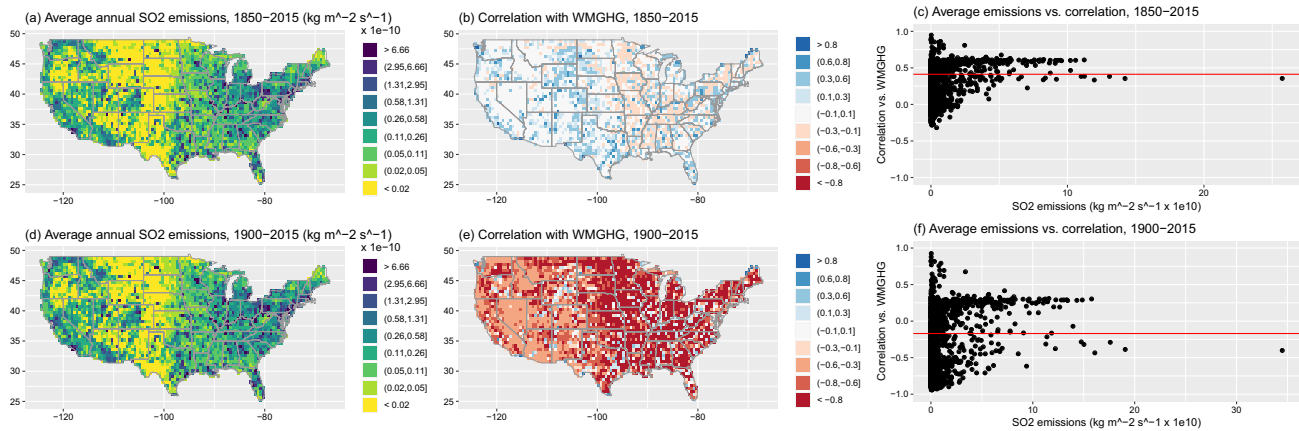


Fig. 13 Average annual SO_2 emissions for (a, d), with grid cell correlations for annual SO_2 emissions versus the time series of WMGHGs for (b, e) and a scatter plot of average annual emissions and

WMGHGs (c, f). The top row corresponds to 1850–2015 and the bottom row corresponds to 1900–2015. The red lines in c and f show the correlation between the CONUS-wide SO_2 emissions and WMGHGs

CONUS-wide SO_2 emissions are only weakly correlated with WMGHGs: the CONUS-wide time series is dominated by the grid cells that have huge emissions, while the individual grid cells with strong anticorrelation tend to have much smaller emissions (by a factor of > 100).

D. The algorithm for the Stochastic Regionalizer

The Stochastic Regionalizer (SR hereafter) is designed to divide a connected region of the Earth's surface, in this case the continental United States, into non-overlapping domains with specific properties to facilitate separating the effects of WMGHG forcing $F_{\text{WMGHG}}(t)$ and SO_2 emissions $F_{\text{aer}}(t)$ on precipitation. The specific properties are set by a user-selectable cost function $C(\mathbf{s}, t)$ that depends on location $\mathbf{s}$ and time t . In general, $C(\mathbf{s}, t)$ is a function of the time series of these forcings:

$$C(\mathbf{s}, t) = C(F_{\text{WMGHG}}(t), F_{\text{aer}}(\mathbf{s}, t), t) \quad (26)$$

The SR operates on data on rectilinear grids from the family of cylindrical projections. The locations $\mathbf{s}$ correspond to points in the center of each grid cell. Let $R_n = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n]$ denote a region of n adjacent grid points in contact with each

other along either meridians or circles of latitude, i.e., not at just at the corners of the grid boxes. Let

$$\langle F_{\text{aer}}(t) \rangle_{R_n} = \sum_{i=1}^n a_i F_{\text{aer}}(\mathbf{s}_i, t) / \sum_{i=1}^n a_i \quad (27)$$

denote the area-weighted regional average of SO_2 emissions where a_i is the area of the grid cell centered on $\mathbf{s}_i$. The cost function can also be computed on this region as:

$$C(R_n, t) = C(F_{\text{WMGHG}}(t), \langle F_{\text{aer}}(t) \rangle_{R_n}, t) \quad (28)$$

The SR constructs regions R_n such that the absolute value of the cost function falls below a user-specified threshold C_{max} :

$$|C(R_n, t)| \leq C_{\text{max}} \quad (29)$$

The algorithm is stochastic to ensure that our results are insensitive to the precise geographical details of any single regional partitioning. In our application of the SR to test hypotheses H4, H5, and H6, for each season we generated 100 realizations of regional partitioning differentiated by random numbers utilized throughout the construction of each region.

The SR algorithm $m > 0$ is as follows:

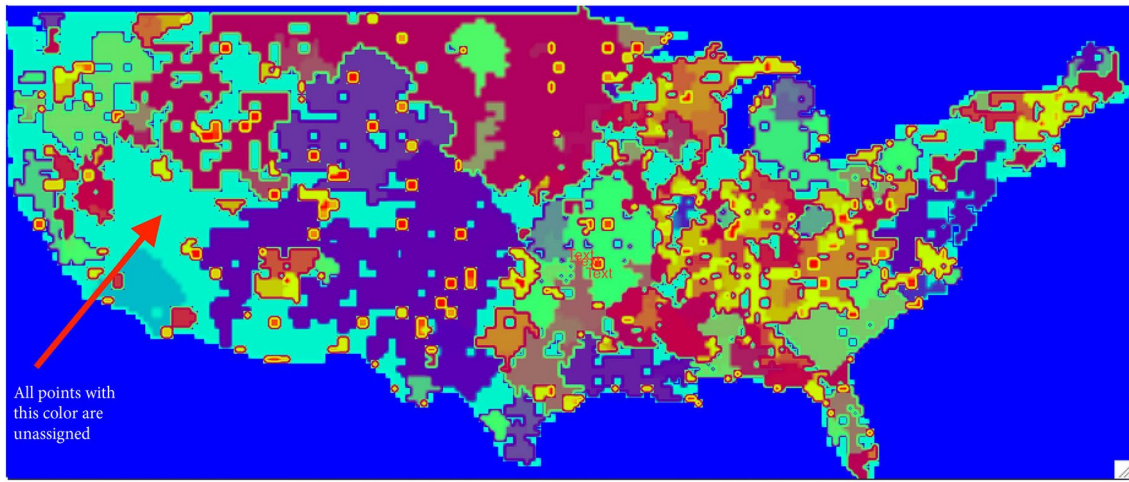


Fig. 14 Sample map of regions output by SR using the Pearson correlation cost function Eq. 32 – each region is assigned a different color

While $m > 0$,

1. Let $U_m = [s_1, s_2, \dots, s_m]$ denote the m CONUS grid points that have yet to be assigned, and let $F_l = [s_1, s_2, \dots, s_l]$ be the set of “failed” grid points around which it is not possible to construct a region satisfying Eq. D.4.
2. If there are one or more j such that $s_j \in U_m$ has $|C(s_j, t)| > C_{\max}$, then set phase $p = 1$, otherwise set phase $p = 2$.
3. Randomly draw either $s_{j'} \in U_m$ with

$$C(s_{j'}, t) = \max_{s' \in U_m} C(s', t) \quad (D.5)$$

or $s_{j''} \in U_m$ with

$$C(s_{j''}, t) = \min_{s'' \in U_m} C(s'', t) \quad (D.6)$$

as the first “seed” grid point s_1 in R_n , and remove that grid point from U_m .

4. Grow the cluster iteratively by:
 - (a) Randomly choose to traverse the region’s outer perimeter in a clockwise or counter-clockwise direction.
 - (b) For each grid point on the perimeter (in order of the traversal),
 - i. Randomly select one of the immediately adjacent grid points s_* that are in U_m . Only unassigned grid points with the same longitude or latitude (i.e., directly above, below, left, or right) as the perimeter grid point are considered.
 - ii. Form a test region $R_{n+1} = [s_1, s_2, \dots, s_n, s_*]$.
 - iii. Compute the cost function $C(R_{n+1}, t)$.
 - iv. If $|C(R_{n+1}, t)| \leq |C(R_n, t)|$, remove s_* from U_m , replace U_m with U_{m-1} , add s_* to R_n , and replace R_n with R_{n+1} . This new point will be on the region’s perimeter in the next iteration.
 - (c) If the region grew during this iteration, move points on the perimeter surrounded on all sides by other points in the cluster from the perimeter to the interior of cluster. Likewise, if points on the perimeter have not been loci of growth in several iterations, these points are also relabeled as interior points.
 - (d) Conclude this iteration, and start the next iteration at step 4a above.
5. If the cluster R_n did not grow during the last iteration:
 - (a) If phase $p = 1$, and
 - i. If $|C(R_n, t)| \leq C_{\max}$, then output the cluster, and decrement m by n .
 - ii. Otherwise, return points $[s_2, \dots, s_n]$ in R_n to U_m , replace U_m by U_{m+n-1} , and increment m by $n - 1$. Add the “seed” s_1 to F_l , replace F_l by F_{l+1} , and increment l by 1.
 - (b) If phase $p = 2$, then by construction $|C(R_n, t)| \leq C_{\max}$. Output the cluster, and decrement m by n .
6. Begin construction of the next cluster at step 2 above.

Once $m = 0$,

1. Output F_l as a “region” which need not be spatially connected. Experience has shown that generally $|C(F_l, t)| \sim C_{\max}$.

We have utilized two cost functions in the SR. Originally we chose to minimize the Pearson correlation $\rho(x, y)$ between two vectors x and y implemented as

$$C(R_n, t) = \rho\left(F \text{ WMGHG}(t), \langle F_{\text{aer}}(t) \rangle_{R_n}\right) \quad (32)$$

We set $C_{\max}$ using the student t-test. Let α be the significance level for the null student t-test hypothesis H_0 that the two time series x and y are uncorrelated. We use $\alpha = 0.5$. We solve for the maximum Pearson correlation coefficient ρ consistent with this significance level by solving for the t value from the cumulative student t-test distribution function, then inverting the relationship between t and ρ to obtain $\rho_{\max} = C_{\max}$.

Let n be the length of the time series x and y , and define $\nu = n - 2$. Denoting the cumulative distribution for the student t-test by $\mathcal{T}(t, \nu)$, we solve for the test statistic t from

$$\alpha = 2[1 - \mathcal{T}(t, \nu)] \quad (33)$$

$$\text{where } \mathcal{T}(t, \nu) = 1 - \frac{1}{2} I_{\nu/2}\left(\frac{t}{\sqrt{2}}, \frac{1}{2}\right), \quad (34)$$

$$x(t) = \frac{\nu}{t^2 + \nu}, \quad (35)$$

and I is an incomplete Bessel function. Using

$$t = \frac{\rho\sqrt{\nu}}{\sqrt{1 - \rho^2}}, \quad (36)$$

we obtain

$$|\rho_{\max}| = \sqrt{\frac{t^2/\nu}{1 + t^2/\nu}} \quad (37)$$

For the tests of hypotheses H4, H5, and H6, we used a different cost function designed to minimize the baseline emissions and change in emissions of SO_2 in each region. After fitting $\langle F_{\text{aer}}(t) \rangle_{R_n}$ with $a + bt$ using linear least squares, we set the cost function to

$$C(R_n, t) = \begin{cases} a & \text{if } |a| > |b \max(t)| \\ b \max(t) & \text{if } |a| \leq |b \max(t)| \end{cases} \quad (38)$$

We set $C_{\max}$ to a user-settable fraction $f \ll 1$ of the 1970 (near maximum) CONUS-mean anthropogenic flux of SO_2 expressed in units of $\text{kg}/\text{m}^2/\text{s}$, the same as units those used in input4MIPs (Feng et al. 2020). Adopting an area $A = 7663941.7 \text{ km}^2$ for the 48 contiguous states and 1970 SO_2 emissions of $E = 31218000$ metric tons (EPA 2020), we obtain a CONUS-mean annually-averaged flux of SO_2 given by

$$F_{\text{SO}_2}(1970) = \frac{E \times 10^3 (\text{kg}/\text{ton})}{A \times [10^3 (\text{m}/\text{km})]^2 \times 365 (\text{days}/\text{year}) \times 86400 (\text{s}/\text{day})} \quad (39)$$

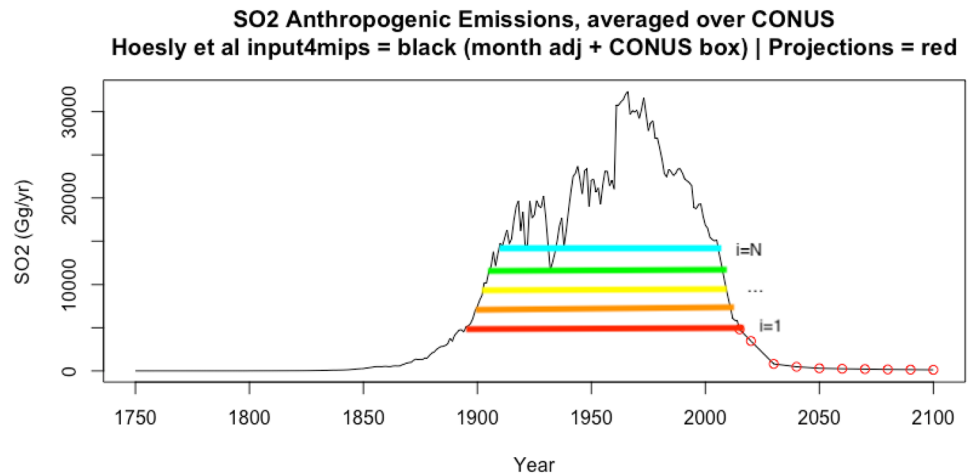
$$= 1.29 \times 10^{-10} \text{ kg}/\text{m}^2/\text{s} \quad (40)$$

In our tests of hypotheses H4, H5, and H6, we set $f = 0.03$ and set

$$C_{\max} = f F_{\text{SO}_2}(1970) \quad (41)$$

To better separate the effects of WMGHGs and SO_2 emissions on precipitation (Fig. 14), we exploit the fact that SO_2 emissions for CONUS rose until the early 1970s, after which the emissions have recently fallen to levels typical of the 19th century (Figs. 2 and 15) due to imposition of mitigation measures for acid rain (NAS 1981). Let $t_{\max}$ be the year when CONUS-mean SO_2 emissions $F_{\text{SO}_2}(t_{\max}) = \max_t [F_{\text{SO}_2}(t)]$, and denote years $t < t_{\max}$ by $t_{<}$ and times $t > t_{\max}$ by $t_{>}$. For each season, we have constructed series of pairs of years $(t_{<,i}, t_{>,i})$ such that the seasonally-averaged CONUS-mean

Fig. 15 Schematic of the δ_i emissions method (Eq. 42). Each colored line extends between a pair of times $(t_{<,i}, t_{>,i})$



$F_{\text{SO}_2}(t_{<,i}) \simeq F_{\text{SO}_2}(t_{>,i})$ (Fig. 15). We denote differences in a quantity $x(t)$ between times $t_{<,i}$ and $t_{>,i}$ by

$$\delta_i x(t) = x(t_{>,i}) - x(t_{<,i}) \quad (42)$$

We then define a fractional difference in emissions by

$$e_i = \frac{\delta_i F_{\text{SO}_2}(t)}{F_{\text{SO}_2}(1970)} * 100\% \quad (43)$$

The series of year pairs and values of e_i are listed for each season in Table 5. The average absolute values of e_i are less than 0.5% throughout the annual cycle. This implies that at the CONUS scale, time series of $\delta_i F_{\text{SO}_2}(t)$ exhibit minimal trends with time. However since the lagged WMGHG forcing is monotonically increasing with time, time series of $\delta_i F_{\text{WMGHG}}(t, \tau \text{ WMGHG})$ increase rapidly with increasing $\delta_i t = t_{>,i} - t_{<,i}$. We use this difference in the temporal trends of increments in forcing by SO_2 and WMGHGs to facilitate separating the effects of these agents

on precipitation. The cost function we employ in the SR algorithm is a form of Eq. 28 modified as follows:

$$C(R_n, t) = C\left(\left[\delta_i F_{\text{WMGHG}}(t)\right], \left[\delta_i \langle F_{\text{aer}}(t) \rangle|_{R_n}\right], \left[\delta_i t\right]\right) \quad (44)$$

E. Confidence statements for each hypothesis

As described in Sect 5, we provided a summary of the extensive analysis in this paper in Table 3, which includes a measure of our confidence in the conclusion of each hypothesis. Any attempt to distill the results of each analysis might feel oversimplified since the various hypothesis analyze numerous models, four seasons, and two types of precipitation (mean and extreme). Nonetheless, such statements provide an important summary that can quickly communicate the results of our explorations and the strength of evidence for

Table 5 Lists of $t_{<,i}$ (years CE), $t_{>,i}$ (years CE), and e_i (Eq. 43, in percent) for each season

DJF			MAM			JJA			SON		
$t_{<,i}$	$t_{>,i}$	e_i	$t_{<,i}$	$t_{>,i}$	e_i	$t_{<,i}$	$t_{>,i}$	e_i	$t_{<,i}$	$t_{>,i}$	e_i
1900	2010	-5.3	1900	2010	-4.9	1900	2011	-3.5	1900	2010	-2.3
1901	2009	-0.6	1901	2009	-0.8	1902	2010	-2	1901	2009	2
1902	2008	4.7	1902	2008	4	1903	2009	-1.7	1903	2008	-4.1
1904	2007	1.1	1904	2007	2.2	1905	2008	-0.7	1908	2006	0.6
1908	2003	-0.9	1905	2006	2.4	1915	2004	1.4	1909	2003	0.4
1909	2002	-0.7	1909	2002	-2.3	1916	2000	0.1	1910	2001	0.7
1913	1996	0	1914	2000	0	1918	1996	-1	1913	1999	0.6
1922	2000	0.1	1917	1994	2.2	1926	1997	1.9	1915	2000	-0.6
1928	1994	-1.3	1918	1993	0.4	1929	1998	0.4	1916	1995	-0.6
1929	1991	0.2	1931	2001	0.8	1933	2007	-0.2	1924	1997	-0.3
1932	2005	2.5	1934	2003	-0.2	1936	2001	0.7	1925	1998	-2.3
1933	2006	-1.7	1936	1996	2	1938	2006	-0.5	1932	2007	-3.8
1935	2001	2.8	1937	1998	-1	1939	2002	-1.3	1934	2004	1.7
1936	1999	0.5	1939	1999	-0.1	1943	1993	1.7	1937	1996	-0.6
1940	1995	-0.3	1941	1991	1.1	1944	1986	0.2	1938	2002	0.3
1941	1992	-0.1	1944	1980	3.2	1947	1992	0.3	1941	1993	1.3
1944	1980	3.7	1945	1984	-0.4	1948	1983	0.7	1942	1989	0.1
1945	1981	0	1947	1981	-0.3	1949	1999	-0.4	1944	1981	-1.2
1946	1985	-0.1	1950	1988	0.3	1956	1994	-0.8	1945	1990	-0.5
1953	1990	0.3	1951	1989	0.1	1961	1972	-5.9	1951	1987	-0.2
1954	1993	-0.3	1952	1992	0.2	1964	1973	0.1	1955	1991	1.2
1955	1987	0.4	1955	1986	0.2				1956	1984	-1.3
1959	1982	-1.7	1958	1983	0				1958	1992	0
1961	1977	-4.9	1959	1990	-0.6				1959	1983	-0.2
1962	1973	-10.1	1961	1973	-0.3				1960	1994	-0.2
									1962	1973	0.7
Average		-0.5	Average		0.3	Average		-0.5	Average		-0.3

each test. We now describe in greater detail how these confidence statements are derived.

The general principle used to derive the confidence statements is to consider the proportion of models, seasons, precipitation types, etc. for which the hypothesis is verified. While this approach implicitly assumes model democracy (i.e., the proportion is not weighted based on any assessment of model quality; see e.g., Knutti et al. (2017)) and furthermore does not account for model dependence (i.e., that the behavior of multiple models may be related due to similar model components; see e.g., Knutti et al. (2013)), more refined statements are beyond the scope of the current work. The conclusions for an individual model/season/etc. is made based on the error bars of a 95% confidence interval and hence robustly accounts for uncertainty throughout. Using a fixed confidence level (95%), the proportion of individual comparisons that satisfy each hypothesis is then compared with the categories in Table 1 of Mastrandrea et al. (2010): ≥ 0.99 implies “virtually certain,” ≥ 0.9 and < 0.99 implies “very likely,” ≥ 0.66 and < 0.9 implies “likely,” ≥ 0.33 and < 0.66 implies “about as likely as not,” ≥ 0.1 and < 0.33 implies “unlikely,” ≥ 0.01 and < 0.1 implies “very unlikely,” and < 0.01 implies “exceptionally unlikely.” We now describe the specific quantities used to determine the confidence label for each hypothesis.

Hypothesis 1. The quantities of interest here are the DAMIP hist-GHG scaling estimates shown in Fig. 3b. Our standard of comparison for this hypothesis is the interquartile range of scaling rates estimated for CMIP5 models in Kharin et al. (2013) and shown with blue bars on Fig. 3b: $1.7\%K^{-1}$ to $2.2\%K^{-1}$ for mean precipitation and $4.3\%K^{-1}$ to $8.0\%K^{-1}$ for extreme precipitation. Across the 11 models and precipitation type (mean/extreme), we count the number of scaling estimates for which the 95% confidence interval is indistinguishable from the Kharin et al. (2013) (denoted K13) ranges:

Number of comp.	# CI indist. from K13	Confidence	Label
22	18	0.82	Likely

Hypothesis 2. The quantities of interest here are the DAMIP hist-GHG and CMIP6 historical global WMGHG coefficients shown in Fig. 3c. Our standard of comparison for this hypothesis is the gray 95% uncertainty band for each precipitation type from a linear fit to the model estimates: since this uncertainty band covers the 1-1 line for the range of estimates in the historical and hist-GHG experiments, we tally the number of models that are consistent with this uncertainty band (i.e., the individual model CIs overlap with the uncertainty band in at least one axis):

Number of comp.	# CI indist. from unc. band	Confidence	Label
22	20	0.91	Very likely

Hypothesis 3. The quantities of interest here are the single-forcing trend estimates summarized in Fig. 4, given for each experiment, model, season, and precipitation type. For each single-forcing experiment (hist-Lu, hist-GHG, hist-aer, hist-nat, and hist-stratO3), we tally the proportion of models/seasons/precipitation type for which the trend error bars do not include zero. Here, “Number of comp.” indicates the number of models/seasons/precipitation type for which we have trend estimates, “# CI $\neq 0$ ” indicates the number of trend confidence intervals that do not include zero, “Confidence” is the third column divided by the second column, and “Label” converts the proportion to the IPCC categories:

Experiment	Number of comp.	# CI $\neq 0$	Confidence	Label
hist-Lu	24	12	0.50	About as likely as not
hist-GHG	88	73	0.83	Likely
hist-aer	88	71	0.81	Likely
hist-nat	88	22	0.25	Unlikely
hist-stratO3	32	13	0.41	About as likely as not

The individual forcings that we deem non-negligible are those that are at least *likely*, i.e., WMGHGs and anthropogenic aerosols, while the others (LULCC, natural, and stratospheric ozone) are deemed negligible.

Hypothesis 4a. For this hypothesis, we consider the number of model-season pairs for which the sum of black carbonaceous aerosols, organic carbonaceous aerosols, and ammonia cannot be reliably distinguished from the control run for Northern Hemisphere land regions. Here, “does not differ” is determined when the 95% confidence interval includes 0. As shown in Table 1, 9/16 (for mean) and 15/16 (for extreme) pairs are indistinguishable from zero:

Number of comp.	# CI = 0	Confidence	Label
32	24	0.75	Likely

Hypothesis 4b. The quantities of interest here are the correlations between changes in the various SO₂ types and precipitation summarized in Fig. 5 across models, seasons, and precipitation type. For the models that have output data for at least two of the three SO₂ types (emiso2, iso2, wetso2), we tally the pairs of error bars that overlap. Here, “Number of comp.” indicates the number of pairs of SO₂ type across models/seasons/precipitation types, “# overlap CI” indicates

the number of pairs of SO₂ correlations that have overlapping CIs, “Confidence” is the second column divided by the first column, and “Label” converts the proportion to the IPCC categories:

Number of comp.	# overlap CI	Confidence	Label
88	86	0.98	Very likely

Hypothesis 4c. The quantities of interest here are the CONUS-average effect of maximum vs. minimum emis02 on precipitation using grid-cell, regionalized, and CONUS-wide time series of SO₂ emissions shown in Fig. 6 across models, seasons, and precipitation type. We simply tally the number of models/seasons/precipitation types for which the grid-cell and CONUS-wide confidence intervals overlap. Here, “Number of comp.” indicates the number of models/seasons/precipitation types, “# overlap CI” indicates the number of comparisons that have overlapping CIs, “Confidence” is the second column divided by the first column, and “Label” converts the proportion to the IPCC categories:

Number of comp.	# overlap CI	Confidence	Label
48	48	1.00	Virtually certain

Hypothesis 5. Given the large number of experiments, models, seasons, precipitation types, and statistics (bias, relative standard deviation, and pattern correlation), to summarize this analysis we simply take the average of all proportions shown in Table 2. The average is 0.79, which translates to a confidence label of “Likely.”

Hypothesis 6. There are two quantities of interest here: first, the CONUS-average bias in WMGHG and SO₂ coefficients for the single-forcing experiments vs. the historical shown in Fig. 7; and second, the Pearson correlations in the WMGHG vs. SO₂ coefficient biases shown in Fig. 8. For the first assessment, we tally the number of models/seasons/precipitation types for which the bias confidence interval includes zero, for both WMGHG and SO₂. For the second, we tally the number of models/seasons/precipitation types for which the correlation confidence interval includes zero. Here, “Number of comp.” indicates the number of models/seasons/precipitation types across WMGHGs and SO₂, “# CI = 0” indicates the number of comparisons for which the CI includes zero, “Confidence” is the third column divided by the second column, and “Label” converts the proportion to the IPCC categories:

Quantity of interest	Number of comp.	# CI = 0	Confidence	Label
Bias	176	131	0.74	Likely
Pearson correlation	88	68	0.77	Likely

Given that both quantities result in a label of “likely,” this is the category used in Table 3 for H6.

Hypothesis 7. The quantities of interest here are the trends in percent of background/total variability shown in Fig. 9 across models, seasons, and precipitation type. We tally the number of trend estimates for which the confidence interval does not include zero. Here, “Number of comp.” indicates the number of models/seasons/precipitation types, “# CI = 0” indicates the number of comparisons for which the trend CI includes zero, “Confidence” is the second column divided by the first column, and “Label” converts the proportion to the IPCC categories:

Number of comp.	# overlap CI	Confidence	Label
40	40	1.00	Virtually certain

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00382-022-06321-1>.

Acknowledgements This research was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 and used resources of the National Energy Research Scientific Computing Center (NERSC), also supported by the Office of Science of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231. This project was supported by the Environmental Resilience Institute, funded by Indiana University’s Prepared for Environmental Change Grand Challenge initiative. This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Author contributions All authors contributed to the study conception, design, and development. Material preparation, data collection, and analysis were performed by MR, WC, PU, TO, and BL. The first draft of the manuscript was written by MR and WC and all authors

commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 and used resources of the National Energy Research Scientific Computing Center (NERSC), also supported by the Office of Science of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231. This project was supported by the Environmental Resilience Institute, funded by Indiana University's Prepared for Environmental Change Grand Challenge initiative.

Availability of data and materials All climate model simulations used (CMIP6-historical, DAMIP, LUMIP, CMIP6 DECK, AerChemMIP, C20C+, HAPPI) are publicly available through the Earth System Grid Federation (<https://esgf-node.llnl.gov/projects/cmip6/>).

Code availability Not applicable.

Declarations

Conflict of interest The authors declare no competing interests.

Ethics approval This report did not study humans and/or animals subjects.

Consent to participate This report did not study humans and/or animals subjects.

Consent for publication This report did not study humans and/or animals subjects.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barnett TP, Pierce DW, AchutaRao KM, Gleckler PJ, Santer BD, Gregory JM, Washington WM (2005) Penetration of human-induced warming into the world's oceans. *Science* 309(5732):284–287. <https://doi.org/10.1126/science.1112418>
- Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon Weather Rev* 115(6):1083–1126. Jun 1987. ISSN 00270644
- Centre for Environmental Data Analysis (CEDA) (2016) CMIP6 data request. <http://clipc-services.ceda.ac.uk/dreq/mipVars.html>
- Chen JP, Hazra A, Shiu C-J, Tsai I-C, Lee H-H (2008) Interaction between Aerosols and Clouds: Current Understanding. World Scientific Publishing Co., pp 231–281, 01 2008. ISBN 978-981-281-890-4. <https://doi.org/10.13140/2.1.2481.7924>
- Coles S (2001) An Introduction to Statistical Modeling of Extreme Values. Lecture Notes in Control and Information Sciences. Springer, 2001. ISBN 978-1-85233-459-8. 10.1007/978-1-4471-3675-0
- Collins WJ, Lamarque J-F, Schulz M, Boucher O, Eyring V, Hegglin MI, Maycock A, Myhre G, Prather M, Shindell D et al (2017) AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6. *Geosci Model Dev* 10(2):585–607. <https://doi.org/10.5194/gmd-10-585-2017>
- Collins M, Knutti R, Arblaster J, Dufresne J-L, Fichefet T, Friedlingstein P, Gao X, Gutowski WJ, Johns T, Krinner G et al (2013) Long-term climate change: projections, commitments and irreversibility. In: *Climate Change 2013-The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pp 1029–1136. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9781107415324.024>
- Dagan G, Koren I, Altaratz O, Lehahn Y (2018) Shallow convective cloud field lifetime as a key factor for evaluating aerosol effects. *iScience* 10:192–202. <https://doi.org/10.1016/j.isci.2018.11.032> (ISSN 2589-0042)
- DelSole T, Trenary L, Yan X, Tippett MK (2019) Confidence intervals in optimal fingerprinting. *Clim Dyn* 52(7):4111–4126
- Deser C (2020) Certain uncertainty: the role of internal climate variability in projections of regional climate change and risk management. *Earth's Future* 8(12). <https://doi.org/10.1029/2020ef001854>
- Dong S, Sun Y, Li C, Zhang X, Min S-K, Kim Y-H (2021) Attribution of extreme precipitation with updated observations and cmip6 simulations. *J Clim* 34(3):871–881. <https://doi.org/10.1175/JCLI-D-19-1017.1>
- Douville H, John A (2021) Fast adjustment versus slow sst-mediated response of daily precipitation statistics to abrupt 4xco2. *Clim Dyn* 56(3):1083–1104
- EPA. National emissions inventory (nei) (2020). URL <https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei>. URL validated 3/5/2022
- Esmaili RB, Tian Y, Vila DA, Kim K-M (2016) A Lagrangian analysis of cold cloud clusters and their life cycles with satellite observations. *J Geophys Res Atmos* 121(19):11723–11738. <https://doi.org/10.1002/2016JD025653>
- Etminan M, Myhre G, Highwood E, Shine K (2016) Radiative forcing of carbon dioxide, methane, and nitrous oxide: a significant revision of the methane radiative forcing. *Geophys Res Lett* 43(24):12–614. <https://doi.org/10.1002/2016GL071930>
- Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci Model Dev* 9(5):1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Feng W, Chipperfield MP, Dhomse S, Monge-Sanz BM, Yang X, Zhang K, Ramonet M (2011) Evaluation of cloud convection and tracer transport in a three-dimensional chemical transport model. *Atmos Chem Phys* 11(12):5783–5803. <https://doi.org/10.5194/acp-11-5783-2011>
- Feng L, Smith SJ, Braun C, Crippa M, Gidden MJ, Hoesly R, Klimont Z, van Marle M, van den Berg M, van der Werf GR (2020) The generation of gridded emissions data for CMIP6. *Geosci Model Dev* 13(2):461–482. <https://doi.org/10.5194/gmd-13-461-2020>
- Gershunov A, Schneider N, Barnett T (2001) Low-frequency modulation of the Enso-Indian monsoon rainfall relationship: signal or noise? *J Clim* 14(11):2486–2492
- Gershunov A, Shulgina T, Clemesha RE, Guirguis K, Pierce DW, Dettinger MD, Lavers DA, Cayan DR, Polade SD, Kalansky J et al (2019) Precipitation regime change in western north America: the role of atmospheric rivers. *Sci Rep* 9(1):1–11

- Gidden M, Riahi K, Smith S, Fujimori S, Luderer G, Kriegler E, van Vuuren D, van den Berg M, Feng L, Klein D, Calvin K, Doelman J, Frank S, Fricko O, Harmsen M, Hasegawa T, Havlik P, Hilaire J, Hoesly R, Horing J, Popp A, Stehfest E, Takahashi K (2018) input4mips.cmp6.scenariomip.iamc.iamc-image-ssp119-1-1. <https://doi.org/10.22033/ESGF/input4MIPs.2485>
- Gillett NP, Zwiers FW, Weaver AJ, Stott PA (2003) Detection of human influence on sea-level pressure. *Nature* 422(6929):292–294. <https://doi.org/10.1038/nature01487>
- Gillett NP, Shiogama H, Funke B, Hegerl G, Knutti R, Matthes K, Santer BD, Stone D, Tebaldi C (2016) The detection and attribution model intercomparison project (DAMIP v1.0) contribution to CMIP6. *Geosci Model Dev* 9(10):3685–3697. <https://doi.org/10.5194/gmd-9-3685-2016>
- Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models. *J Geophys Res Atmos* 113(D6)
- Granger C (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econom J Econom Soc* 424–438. <https://doi.org/10.2307/1912791>
- Gregory JM, Ingram WJ, Palmer MA, Jones GS, Stott PA, Thorpe RB, Lowe JA, Johns TC, Williams KD (2004) A new method for diagnosing radiative forcing and climate sensitivity. *Geophys Res Lett* 31(3). <https://doi.org/10.1029/2003GL018747>
- Groisman PY, Knight RW, Easterling DR, Karl TR, Hegerl GC, Razuvaev VN (2005) Trends in intense precipitation in the climate record. *J Clim* 18(9):1326–1350
- Guo R, Deser C, Terray L, Lehner F (2019) Human influence on winter precipitation trends (1921–2015) over North America and Eurasia revealed by dynamical adjustment. *Geophys Res Lett* 46(6):3426–3434. <https://doi.org/10.1029/2018GL081316>
- Hannart A, Pearl J, Otto F, Naveau P, Ghil M (2015) Causal counterfactual theory for the attribution of weather and climate-related events. *Bull Am Meteorol Soc* 97:99–110. <https://doi.org/10.1175/BAMS-D-14-00034.1>
- Hartmann D, Klein Tank A, Rusticucci M, Alexander L, Brönnimann S, Charabi Y, Dentener F, Dlugokencky E, Easterling D, Kaplan A, Soden B, Thorne P, Wild M, Zhai P (2013) Observations: Atmosphere and surface. In: Stocker T, Qin D, Plattner G-K, Tignor M, Allen S, Boschung J, Nauels A, Xia Y, Bex V, Midgley P (eds) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 2, pp 159–254. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. ISBN ISBN 978-1-107-66182-0. <https://doi.org/10.1017/CBO9781107415324.008>
- Hegerl GC, Hasselmann K, Cubasch U, Mitchell JF, Roeckner E, Voss R, Waszkewitz J (1997) Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. *Clim Dyn* 13(9):613–634. <https://doi.org/10.1007/s003820050186>
- Hegerl GC, Black E, Allan RP, Ingram WJ, Polson D, Trenberth KE, Chadwick RS, Arkin PA, Sarojini BB, Becker A et al (2015) Challenges in quantifying changes in the global water cycle. *Bull Am Meteorol Soc* 96(7):1097–1115. <https://doi.org/10.1175/BAMS-D-13-00212.1>
- Hodnebrog Ø, Etmann M, Fuglestedt J, Marston G, Myhre G, Nielsen C, Shine KP, Wallington T (2013) Global warming potentials and radiative efficiencies of halocarbons and related compounds: a comprehensive review. *Rev Geophys* 51(2):300–378. <https://doi.org/10.1002/rog.20013>
- Hoesly RM, Smith SJ, Feng L, Klimont Z, Janssens-Maenhout G, Pitkanen T, Seibert JJ, Vu L, Andres RJ, Bolt RM, Bond TC, Dawidowski L, Kholod N, Kurokawa J-I, Li M, Liu L, Lu Z, Moura MCP, O'Rourke PR, Zhang Q (2018) Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geosci Model Dev* 11(1):369–408. <https://doi.org/10.5194/gmd-11-369-2018>
- Huang H, Patricola CM, Winter JM, Osterberg EC, Mankin JS (2021) Rise in Northeast US extreme precipitation caused by Atlantic variability and climate change. *Weather Clim Extrem* 33:100351. <https://doi.org/10.1016/j.wace.2021.100351>
- Hurrell JW, Kushnir Y, Ottensen G, Visbeck M (2003) An overview of the North Atlantic Oscillation. In: *Geophysical Monograph Series*, vol 134, pp 1–35. American Geophysical Union. ISBN 9781118669037. 10.1029/134GM
- Jones GS, Tett SF, Stott PA (2003) Causes of atmospheric temperature change 1960–2000: a combined attribution analysis. *Geophys Res Lett* 30(5). 10.1029/2002GL016377
- Kay JE, Deser C, Phillips A, Mai A, Hannay C, Strand G, Arblaster JM, Bates SC, Danabasoglu G, Edwards J, Holland M, Kushner P, Lamarque J-F, Lawrence D, Lindsay K, Middleton A, Munoz E, Neale R, Oleson K, Polvani L, Vertenstein M (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 Am Meteorol Soc* 96(8):1333–1349. <https://doi.org/10.1175/bams-d-13-00255.1>
- Kerr RA (2000) A North Atlantic Climate Pacemaker for the Centuries. *Science* 288(5473):1984–1985. <https://doi.org/10.1126/science.288.5473.1984>
- Kharin VV, Zwiers F, Zhang X, Wehner M (2013) Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Clim Change* 119(2):345–357. <https://doi.org/10.1007/s10584-013-0705-8>
- Kirchmeier-Young MC, Zhang X (2020) Human influence has intensified extreme precipitation in North America. *Proc Natl Acad Sci* 117(24):13308–13313. <https://doi.org/10.1073/pnas.1921628117>
- Kirkevåg A, Iversen T, Kristjansson JE, Seland O, Debernard JB (2008) On the additivity of climate response to anthropogenic aerosols and CO₂, and the enhancement of future global warming by carbonaceous aerosols. *Tellus Ser A-Dyn Meteorol Oceanogr* 60(3):513–527. <https://doi.org/10.1111/j.1600-0870.2008.00308.x>
- Knutson TR, Zeng F (2018) Model assessment of observed precipitation trends over land regions: detectable human influences and possible low bias in model trends. *J Clim* 31(12):4617–4637. <https://doi.org/10.1175/JCLI-D-17-0672.1>
- Knutti R, Masson D, Gettelman A (2013) Climate model genealogy: generation CMIP5 and how we got there. *Geophys Res Lett* 40(6):1194–1199. <https://doi.org/10.1002/grl.50256>
- Knutti R, Sedláček J, Sanderson BM, Lorenz R, Fischer EM, Eyring V (2017) A climate model projection weighting scheme accounting for performance and interdependence. *Geophys Res Lett* 44:1908–1918. <https://doi.org/10.1002/2016gl072012>
- Kunkel KE (2003) North American trends in extreme precipitation. *Nat Hazards* 29(2):291–305. <https://doi.org/10.1023/A:1023694115864>
- Kunkel KE, Easterling DR, Kristovich DAR, Gleason B, Stoecker L, Smith R (2012) Meteorological causes of the secular variations in observed extreme precipitation events for the conterminous United States. *J Hydrometeorol* 13(3):1131–1141. <https://doi.org/10.1175/jhm-d-11-0108.1>
- Lamarque J-F, Dentener F, McConnell J, Ro C-U, Shaw M, Vet R, Bergmann D, Cameron-Smith P, Dalsoren S, Doherty R, Faluvegi G, Ghan SJ, Josse B, Lee YH, MacKenzie IA, Plummer D, Shindell DT, Skeie RB, Stevenson DS, Strode S, Zeng G, Curran M, Dahl-Jensen D, Das S, Fritzsche D, Nolan M (2013) Multi-model mean nitrogen and sulfur deposition from the atmospheric chemistry and climate model intercomparison project (ACC-MIP): evaluation of historical and projected future changes. *Atmos Chem Phys* 13(16):7997–8018. <https://doi.org/10.5194/acp-13-7997-2013>

- Lawrence DM, Hurtt GC, Arneth A, Brovkin V, Calvin KV, Jones AD, Jones CD, Lawrence PJ, Noblet-Ducoudré ND, Pongratz J et al (2016) The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: rationale and experimental design. *Geosci Model Dev* 9(9):2973–2998. <https://doi.org/10.5194/gmd-9-2973-2016>
- Lee C, Martin RV, van Donkelaar A, Lee H, Dickerson RR, Hains JC, Krotkov N, Richter A, Vinnikov K, Schwab JJ (2011) SO₂ emissions and lifetimes: estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations. *J Geophys Res Atmos* 116(D6). 10.1029/2010JD014758
- Liu L, Shawki D, Voulgarakis A, Kasoar M, Samset BH, Myhre G, Forster PM, Hodnebrog Ø, Sillmann J, S. G. Aalbergssjø O, Boucher, G, Faluvegi, T, Iversen, A, Kirkevåg, J.-F. Lamarque, D. Olivie, T. Richardson, D. Shindell, and T. Takemura, (2018) A PDRMIP multimodel study on the impacts of regional aerosol forcings on global and regional precipitation. *J Clim* 31(11):4429–4447. <https://doi.org/10.1175/JCLI-D-17-0439.1>
- Lohmann U, Feichter J (2005) Global indirect aerosol effects: a review. *Atmos Chem Phys* 5(3):715–737. <https://doi.org/10.5194/acp-5-715-2005>
- Loring B, O'Brien TA, Elbashandy AE, Johnson JN, Krishnan H, Keen N (2016) Prabhat, and the CASCADE SFA. Toolkit for Extreme Climate Analysis. Lawrence Berkeley National Lab. <https://doi.org/10.20358/C8C651>
- Mann ME, Steinman BA, Miller SK (2020) Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nat Commun* 11(1):49. <https://doi.org/10.1038/s41467-019-13823-w>
- Mann ME, Steinman BA, Brouillette DJ, Miller SK (2021) Multidecadal climate oscillations during the past millennium driven by volcanic forcing. *Science* 371(6533):1014–1019. <https://doi.org/10.1126/science.abc5810>
- Marvel K, Schmidt GA, Shindell D, Bonfils C, LeGrande AN, Nazarenko L, Tsigaridis K (2015) Do responses to different anthropogenic forcings add linearly in climate models? *Environ Res Lett* 10(10). 10.1088/1748-9326/10/10/104010
- Mastrandrea MD, Field CB, Stocker TF, Edenhofer O, Ebi KL, Frame DJ, Held H, Kriegler E, Mach KJ, Matschoss PR et al (2021) Guidance note for lead authors of the IPCC Fifth Assessment Report on consistent treatment of uncertainties, 2010. https://www.ipcc.ch/site/assets/uploads/2017/08/AR5_Uncertainty_Guidance_Note.pdf, verified 21. Jun. 2021
- McKinnon KA, Deser C (2018) Internal variability and regional climate trends in an observational large ensemble. *J Clim* 31(17):6783–6802. <https://doi.org/10.1175/jcli-d-17-0901.1>
- Meehl GA, Senior CA, Eyring V, Flato G, Lamarque J-F, Stouffer RJ, Taylor KE, Schlund M (2020) Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Sci Adv* 6(26). 10.1126/sciadv.aba1981
- Meinshausen M, Nicholls ZRJ (2018) Uom-remind-magpie-ssp585-1-2-0: Remind-magpie-ssp585 GHG concentrations. <https://doi.org/10.22033/ESGF/input4MIPs.2349>
- Meinshausen M, Vogel E (2016) input4mips.uom.ghgconcentrations. cmip.uom-cmip-1-2-0. <https://doi.org/10.22033/ESGF/input4MIPs.1118>
- Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG (2012) An overview of the Global Historical Climatology Network-Daily database. *J Atmos Oceanic Tech* 29(7):897–910. <https://doi.org/10.1175/JTECH-D-11-00103.1>
- Mills MJ, Schmidt A, Easter R, Solomon S, Kinnison DE, Ghan SJ, Neely RR III, Marsh DR, Conley A, Bardeen CG et al (2016) Global volcanic aerosol properties derived from emissions, 1990–2014, using CESM1 (WACCM). *J Geophys Res Atmos* 121(5):2332–2348. <https://doi.org/10.1002/2015JD024290>
- Min S-K, Zhang X, Zwiers FW, Hegerl GC (2011) Human contribution to more-intense precipitation extremes. *Nature* 470(7334):378. <https://doi.org/10.1038/nature09763>
- Mitchell D, AchutaRao K, Allen M, Bethke I, Beyerle U, Ciavarella A, Forster PM, Fuglestedt J, Gillett N, Haustein K, Ingram W, Iversen T, Kharin V, Klingaman N, Massey N, Fischer E, Schleussner C-F, Scinocca J, Seland Ø, Shiogama H, Shuckburgh E, Sparrow S, Stone D, Uhe P, Wallom D, Wehner M, Zaaboul R (2017) Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design. *Geosci Model Dev* 10(2):571–583. <https://doi.org/10.5194/gmd-10-571-2017>
- Myhre G, Forster PM, Samset BH, Hodnebrog Ø, Sillmann J, S. G. Aalbergssjø T, Andrews, O, Boucher, G, Faluvegi, D, Flaschner, T, Iversen, M, Kasoar, V, Kharin, A, Kirkevåg, J.-F. Lamarque, D. Olivie, T. B. Richardson, D. Shindell, K. P. Shine, C. W. Stjern, T. Takemura, A. Voulgarakis, and F. Zwiers, (2017) PDRMIP: a precipitation driver and response model intercomparison project-protocol and preliminary results. *Bull Am Meteorol Soc* 98(6):1185–1198. <https://doi.org/10.1175/BAMS-D-16-0019.1>
- NAS. *Atmosphere-Biosphere Interactions*. National Academies Press, 1981. ISBN 978-0-309-03196-7. 10.17226/135
- Noake K, Polson D, Hegerl G, Zhang X (2012) Changes in seasonal land precipitation during the latter twentieth-century. *Geophys Res Lett* 39(3). 10.1029/2011GL050405
- O'Brien JP (2019) The quantification of co-occurring meteorological extremes and the anthropogenic contribution to hydrometeorological variation. PhD thesis, University of California, Santa Cruz. URL <https://escholarship.org/uc/item/8fh4h62f>
- Paciorek C, Stone D, Wehner M (2018) Quantifying statistical uncertainty in the attribution of human influence on severe weather. *Weather Clim Extrem* 20:69–80. <https://doi.org/10.1016/j.wace.2018.01.002>
- Paik S, Min S-K, Zhang X, Donat MG, King AD, Sun Q (2020) Determining the anthropogenic greenhouse gas contribution to the observed intensification of extreme precipitation. *Geophys Res Lett* 47(12):e2019GL086875. <https://doi.org/10.1029/2019GL086875>
- Patricola C, Wehner M (2018) Anthropogenic influences on major tropical cyclone events. *Nature* 563(7731):339. <https://doi.org/10.1038/s41586-018-0673-2>
- Pendergrass AG, Knutti R, Lehner F, Deser C, Sanderson BM (2017) Precipitation variability increases in a warmer climate. *Sci Rep* 7(1):17966. <https://doi.org/10.1038/s41598-017-17966-y>
- Phillips AS, Deser C, Fasullo J (2014) Evaluating modes of variability in climate models. *EOS Trans Am Geophys Union* 95(49):453–455. <https://doi.org/10.1002/2014EO490002>
- Polade SD, Pierce DW, Cayan DR, Gershunov A, Dettinger MD (2014) The key role of dry days in changing regional climate and precipitation regimes. *Sci Rep* 4(1):1–8
- Polade SD, Gershunov A, Cayan DR, Dettinger MD, Pierce DW (2017) Precipitation in a warming world: assessing projected hydroclimate changes in California and other Mediterranean climate regions. *Sci Rep* 7(1):1–10
- Report CSS, Easterling D, Kunkel K, Arnold J, Knutson T, LeGrande A, Leung L, Vose R, Waliser D, Wehner M (2017) Precipitation change in the United States. In: Fourth National Climate Assessment I 207–230. <https://doi.org/10.7930/J0H993CC>
- Resplandy L, Keeling RF, Eddebbar Y, Brooks M, Wang R, Bopp L, Long MC, Dunne JP, Koeve W, Oschlies A (2019) Quantification of ocean heat uptake from changes in atmospheric O₂ and CO₂ composition. *Sci Rep* 9(1):20244. <https://doi.org/10.1038/s41598-019-56490-z>
- Richardson TB, Forster PM, Andrews T, Boucher O, Faluvegi G, Flaschner D, Hodnebrog Ø, Kasoar M, Kirkevåg A, Lamarque J-F, Myhre G, Olivie D, Samset BH, Shawki D, Shindell D,

- Takemura T, Voulgarakis A (2018) Drivers of precipitation change: an energetic understanding. *J Clim* 31(23):9641–9657. <https://doi.org/10.1175/JCLI-D-17-0240.1>
- Ricke KL, Caldeira K (2014) Maximum warming occurs about one decade after a carbon dioxide emission. *Environ Res Lett* 9(12):124002. <https://doi.org/10.1088/1748-9326/9/12/124002>
- Risser MD, Wehner MF (2017) Attributable human-induced changes in the likelihood and magnitude of the observed extreme precipitation during Hurricane Harvey. *Geophys Res Lett* 44(24):12457–12464. <https://doi.org/10.1002/2017GL075888>
- Risser MD, Paciorek CJ, O'Brien TA, Wehner MF, Collins WD (2019) Detected changes in precipitation extremes at their native scales derived from in situ measurements. *J Clim* 32(23):8087–8109. <https://doi.org/10.1175/JCLI-D-19-0077.1>
- Risser MD, Paciorek CJ, Wehner MF, O'Brien TA, Collins WD (2019) A probabilistic gridded product for daily precipitation extremes over the United States. *Clim Dyn* 53(5):2517–2538. <https://doi.org/10.1007/s00382-019-04636-0> (ISSN 1432-0894)
- Risser MD, Wehner MF, O'Brien JP, Patricola CM, O'Brien TA, Collins WD, Paciorek CJ, Huang H (2021) Quantifying the influence of natural climate variability on in situ measurements of seasonal total and extreme daily precipitation. *Clim Dyn* 56:3205–3230. <https://doi.org/10.1007/s00382-021-05638-7>
- Samset BH, Myhre G, Forster PM, Hodnebrog Ø, Andrews T, Faluvegi G, Fläschner D, Kasoar M, Kharin V, Kirkevåg A, Lamarque J-F, Olivé D, Richardson T, Shindell D, Shine KP, Takemura T, Voulgarakis A (2016) Fast and slow precipitation responses to individual climate forcings: a PDRMIP multimodel study. *Geophys Res Lett* 43(6):2782–2791. <https://doi.org/10.1002/2016GL068064>
- Santer BD, Wehner MF, Wigley T, Sausen R, Meehl G, Taylor K, Ammann C, Arblaster J, Washington W, Boyle J et al (2003) Contributions of anthropogenic and natural forcing to recent tropopause height changes. *Science* 301(5632):479–483. <https://doi.org/10.1126/science.1084123>
- Sarojini BB, Stott PA, Black E, Polson D (2012) Fingerprints of changes in annual and seasonal precipitation from CMIP5 models over land and ocean. *Geophys Res Lett* 39(21). 10.1029/2012GL053373
- Sarojini BB, Stott PA, Black E (2016) Detection and attribution of human influence on regional precipitation. *Nat Clim Change* 6(7):669–675. <https://doi.org/10.1038/nclimate2976>
- Sato M, Hansen JE, McCormick MP, Pollack JB (1993) Stratospheric aerosol optical depths, 1850–1990. *J Geophys Res Atmos* 98(D12):22987–22994. <https://doi.org/10.1029/93JD02553>
- Schlesinger ME, Ramankutty N (1994) An oscillation in the global climate system of period 65–70 years. *Nature* 367(6465):723–726. <https://doi.org/10.1038/367723a0> (ISSN 00280836)
- Schmidt A, Mills MJ, Ghan S, Gregory JM, Allan RP, Andrews T, Bardeen CG, Conley A, Forster PM, Gettelman A et al (2018) Volcanic radiative forcing from 1979 to 2015. *J Geophys Res Atmos* 123(22):12–491. <https://doi.org/10.1029/2018JD028776>
- Shiogama H, Stone DA, Nagashima T, Nozawa T, Emori S (2013) On the linear additivity of climate forcing-response relationships at global and continental scales. *Int J Climatol* 33(11):2542–2550. <https://doi.org/10.1002/joc.3607>
- Storelvmo T (2012) Uncertainties in aerosol direct and indirect effects attributed to uncertainties in convective transport parameterizations. *Atmos Res* 118:357–369. <https://doi.org/10.1016/j.atmosres.2012.06.022>
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res Atmos* 106(D7):7183–7192. <https://doi.org/10.1029/2000JD900719>
- Tett SF, Stott PA, Allen MR, Ingram WJ, Mitchell JF (1999) Causes of twentieth-century temperature change near the Earth's surface. *Nature* 399(6736):569–572. <https://doi.org/10.1038/21164>
- Thompson DWJ, Wallace JM (2000) Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability. *J Clim* 13(5):1000–1016. [https://doi.org/10.1175/1520-0442\(1999\)013<1000:AMIE>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)013<1000:AMIE>2.0.CO;2) (ISSN 0894-8755)
- Thornhill GD, Collins WJ, Kramer RJ, Olivé D, Skeie RB, O'Connor FM, Abraham NL, Checa-Garcia R, Bauer SE, Deushi M, Emmons LK, Forster PM, Horowitz LW, Johnson B, Keeble J, Lamarque J-F, Michou M, Mills MJ, Mulcahy JP, Myhre G, Nabat P, Naik V, Oshima N, Schulz M, Smith CJ, Takemura T, Tilmes S, Wu T, Zeng G, Zhang J (2021) Effective radiative forcing from emissions of reactive gases and aerosols—a multi-model comparison. *Atmos Chem Phys* 21(2):853–874. <https://doi.org/10.5194/acp-21-853-2021>
- Tost H, Lawrence MG, Brühl C, Jöckel P (2010) Uncertainties in atmospheric chemistry modelling due to convection parameterisations and subsequent scavenging. *Atmos Chem Phys* 10(4):1931–1951. <https://doi.org/10.5194/acp-10-1931-2010>
- Wallace JM, Gutzler DS (1981) Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter. *Mon Weather Rev* 109(4):784–812. [https://doi.org/10.1175/1520-0493\(1981\)109<0784:TE>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<0784:TE>2.0.CO;2) (ISSN 0027-0644)
- Wang SS, Zhao L, Yoon J-H, Klotzbach P, Gillies RR (2018) Quantitative attribution of climate effects on Hurricane Harvey's extreme rainfall in Texas. *Environ Res Lett* 13(5):054014. <https://doi.org/10.1088/1748-9326/aabb85>
- Wilcox LJ, Highwood EJ, Dunstone NJ (2013) The influence of anthropogenic aerosol on multi-decadal variations of historical global climate. *Environ Res Lett* 8(2):024033. <https://doi.org/10.1088/1748-9326/8/2/024033>
- Williams IN, Patricola CM (2018) Diversity of ENSO events unified by convective threshold sea surface temperature: a nonlinear ENSO index. *Geophys Res Lett* 45(17):9236–9244. <https://doi.org/10.1029/2018GL079203>
- Wu P, Christidis N, Stott P (2013) Anthropogenic impact on Earth's hydrological cycle. *Nat Clim Change* 3(9):807–810. <https://doi.org/10.1038/nclimate1932>
- Yang Y, Wang H, Smith SJ, Zhang R, Lou S, Yu H, Li C, Rasch PJ (2018) Source Apportionments of Aerosols and Their Direct Radiative Forcing and Long-Term Trends Over Continental United States. *Earth's Future* 6(6):793–808. <https://doi.org/10.1029/2018EF000859>
- Zhang X, Zwiers FW, Hegerl GC, Lambert FH, Gillett NP, Solomon S, Stott PA, Nozawa T (2007) Detection of human influence on twentieth-century precipitation trends. *Nature* 448(7152):461–465. <https://doi.org/10.1038/nature06025>
- Zhang X, Wan H, Zwiers FW, Hegerl GC, Min S-K (2013) Attributing intensification of precipitation extremes to human influence. *Geophys Res Lett* 40(19):5252–5257. <https://doi.org/10.1002/grl.51010>
- Zhang F, Sun YQ, Magnusson L, Buizza R, Lin S-J, Chen J-H, Emanuel K (2019) What is the predictability limit of midlatitude weather? *J Atmos Sci* 76(4):1077–1091. <https://doi.org/10.1175/jas-d-18-0269.1>

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