A Method for Detection and Attribution of Regional Precipitation Change Using Granger Causality Application to the United States Historical Record

Mark Risser (✉ mdrisser@lbl.gov)

Lawrence Berkeley National Laboratory <https://orcid.org/0000-0003-1956-1783>

William Collins

Lawrence Berkeley Laboratory: E O Lawrence Berkeley National Laboratory

Michael Wehner

Lawrence Berkeley Laboratory: E O Lawrence Berkeley National Laboratory

Travis O'Brien

Indiana University Bloomington

Christopher Paciorek

University of California Berkeley

John P. O'Brien

National Center for Atmospheric Research

Christina Patricola

Iowa State University

Huanping Huang

Lawrence Berkeley Laboratory: E O Lawrence Berkeley National Laboratory

Paul Ullrich

University of California Davis

Burlen Loring

Lawrence Berkeley Laboratory: E O Lawrence Berkeley National Laboratory

Research Article

Keywords: Extreme precipitation, Natural variability, Local impacts, Anthropogenic aerosols, CMIP6, DAMIP

Posted Date: August 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-785460/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A method for detection and attribution of regional precipitation change using Granger causality

Application to the United States historical record

Mark D. Risser · William D. Collins · Michael F. Wehner · Travis
A. O'Brien · Christopher J. Paciorek · John P. O'Brien · Christina
M. Patricola · Huanping Huang · Paul A. Ullrich · Burlen Loring

Received: date / Accepted: date

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 propose a novel approach to regional detection and attribution (D&A) for precipitation, starting with the contiguous United States (CONUS) where observational uncertainty is minimized. In a single framework, we 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. We use output from global climate models in a perfect-data sense 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). In doing so, we turn an apparent limitation into an opportunity by using the diversity of responses to short-lived climate forcings across the CMIP6 multi-model ensemble to ensure our D&A is insensitive to structural uncertainty. Our framework is developed using a Pearl-causal perspective, but forthcoming research now underway will apply the framework to in situ measurements using a Granger-causal perspective. 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

M. Risser
Lawrence Berkeley National Laboratory, Berkeley, CA, United States
E-mail: mdrisser@lbl.gov

W. Collins
Lawrence Berkeley National Laboratory, Berkeley, CA, United States
Department of Earth and Planetary Science, University of California, Berkeley, CA, United States

M. Wehner
Lawrence Berkeley National Laboratory, Berkeley, CA, United States

T. O'Brien
Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, United States
Lawrence Berkeley National Laboratory, Berkeley, CA, United States

C. Paciorek
Department of Statistics, University of California, Berkeley, CA, United States

J. O'Brien
National Center for Atmospheric Research, Boulder, CO, United States
Lawrence Berkeley National Laboratory, Berkeley, CA, United States

C. Patricola
Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA, United States
Lawrence Berkeley National Laboratory, Berkeley, CA, United States

H. Huang
Lawrence Berkeley National Laboratory, Berkeley, CA, United States

P. Ullrich
Department of Land, Air, and Water Resources, University of California, Davis, CA, United States

B. Loring
Lawrence Berkeley National Laboratory, Berkeley, CA, United States

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 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. Two recent studies [Kirchmeier-Young and Zhang, 2020, Huang et al., 2021] detect changes in extreme precipitation at subcontinental 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. 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. 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 using the diversity of responses to SLCFs across the CMIP6 multi-model ensemble to ensure our D&A is insensitive to structural uncertainty.

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 Section 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 Section 3, we describe the climate model data sets, forcing data sets, and modes of natural climate variability used to test the hypotheses proposed in Section 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 (Figure 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

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

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. A.4 (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 Section 3.3 for more information). 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.

[†] 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 Section 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 variability). The yellow boxes correspond to the general research questions explored in the paper, and the green circles denote the methods for making conclusions about each research question. The red boxes contain the resulting conclusions.

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 Section 4 and found that we could simplify Eq. 1 (the flow chart diagram in Figure 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 Figure 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. However, 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 any 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 Section 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, namely 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). Asian and European aerosols do not affect US rainfall, and slow responses to aerosols (teleconnected from sea-surface temperatures (SSTs)) also do not affect precipitation.

Hypothesis 4d (H4d). For CONUS and over the historical record, a single CONUS-wide time series of SO_2 emissions yields the same result as using local 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:

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

The time lag τ_{WMGHG} is estimated in Appendix A. For reasons discussed in section 4.3.3, 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 Section 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). 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 E.1, 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 CO₂ (abrupt4xCO₂); and a simulation forced by a 1% yr⁻¹ CO₂ 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 CO₂ value that is used in piControl. Table E.2 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 [AerChemMIP; Collins et al., 2017] 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 SO₂), 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 E.3 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 E.4 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 Section 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 E.5).

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

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 Figure 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 , 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 Figure 2(a).

Since the response of precipitation to WMGHG forcings is dominated by the lagged response of SSTs to these forcings [Samset et al., 2016], 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 Section 4.3); the punchline is that for the CMIP6-historical, DAMIP hist-aer, and LUMIP hist-noLu experiments we find that a single CONUS-wide time series of SO_2 emissions derived from observations sufficiently characterizes the influence of anthropogenic aerosols on seasonal precipitation. Specifically, 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, shown in Figure 2(c), agree well with various other estimates of total CONUS emissions (see Supplemental Figure F.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); Section 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 Section 4.2. We use the Climate Variability Diagnostics Package [CVD; 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 Section 1. Figure 1 summarizes the motivation and testing of each hypothesis.

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 Equation 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 Equation 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 ts (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 Figure 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 1-3% 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 Figure 3, particularly given the moderate levels of warming over the historical period in the hist-GHG runs. Similar results are obtained when analyzing the CMIP6 DECK experiments piControl and 1pctCO2 introduced in Section 3.1, which involve significantly more warming to the global climate system; see Appendix B. These results support our confidence assessment in Appendix D 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,

Fig. 3 Estimated hist-GHG changes in globally averaged annual mean and extreme precipitation versus warming (panel a.), the corresponding hist-GHG scaling factors $P_{\text{mean},m}^{\text{hist-GHG}}$ and $P_{\text{GEV},m}^{\text{hist-GHG}}$ versus warming (panel b.), and statistical WMGHG coefficient estimates calculated from hist-GHG versus historical for mean and extreme precipitation (panel c.). Each point contains coordinate-wise 95% basic bootstrap confidence intervals. In panels (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 panel (c) are the 1-1 line, and the blue lines in panel (c) represent the linear fit for the historical relative to hist-GHG (95% uncertainty band in light gray).

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:

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

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

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

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 E.1 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 Figure 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 D 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 Section 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 Figure 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 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 D regarding the likelihood that individual forcings are non-negligible for secular trends in precipitation

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

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]).

4.3 Structural uncertainty in representing anthropogenic aerosols (H4)

Given the clear influence of tropospheric aerosols on seasonal precipitation identified in Figure 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 (Figure F.3).

To this end, recall that our exploration of structural uncertainty in representing anthropogenic aerosols (H4) actually involves four sub-hypotheses. The punchline of this analysis is:

- SO_2 is, in fact, the dominant aerosol species (Hypothesis 4a),
- We can use SO_2 emissions (denoted “emiso2”) just as well as any other model-specific measure of SO_2 (Hypothesis 4b),
- We can safely ignore aerosol emissions from outside the CONUS and assume that the fast (aerosol-cloud interaction mediated) response dominates the slow (SST-mediated) response of CONUS rainfall to anthropogenic SO_2 emissions (Hypothesis 4c), and
- CONUS-wide measures of SO_2 emissions work as well as localized measures for attributing changes in rainfall (Hypothesis 4d).

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

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

We now justify each of these conclusions in the next subsections using climate model output from the AerChemMIP experiments described in Section 3.1.

4.3.1 Dominant aerosol species for changes in precipitation (H4a)

First, 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 F.2. This result supports our confidence assessment in Appendix D 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 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

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).

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 D 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 Aerosol measures from outside of CONUS (H4c)

Results from the Precipitation Driver and Response Model Intercomparison Project [PDRMIP; Myhre et al., 2017] show that Asian and European sulfate aerosols have negligible effects on precipitation over CONUS, and therefore we can use aerosol measures local to CONUS [Liu et al., 2018]. 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 SO₂. Supported by the PDRMIP findings and the source apportionment by Yang et al. [2018], we hypothesize that the response of CONUS to SO₂ is dominated by fast responses to domestic emissions of this species.

4.3.4 CONUS-wide vs. local estimates of SO₂ emissions (H4d)

Finally, given the results of Hypotheses 4a, 4b, and 4c, we evaluate whether the influence of anthropogenic aerosols should be determined using local (model grid-cell) estimates of emiso2, or whether a single CONUS-wide average time series of emiso2 can be used. In principle, a local measure would be appropriate given the dominance of fast (microphysics-mediated) effects of SO₂ on rainfall (Section 4.3.3), the large geographic gradients in anthropogenic emissions and column-integrated atmospheric burdens of SO₂ 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] and Dagan et al. [2018]).

There are two reasons that a single, CONUS-wide time series would make the follow-up observational analysis analytically tractable. First, our observational analysis is limited to 1900-present due to data availability. However, over this period the correlation between the annually averaged grid cell SO₂ emissions and WMGHGs is very strong (exceeding 0.8 in absolute value for $\approx 50\%$ of CONUS) which prevents a regression-based analysis from distinguishing the effect of WMGHG on precipitation from the effect of SO₂. These strong correlations are much weaker for 1850-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). The correlation between WMGHGs and a CONUS-wide time series of SO₂ emissions is sufficiently small for both 1850-present and 1900-present. (See Appendix C for further details on the relationship between WMGHG forcing and grid-cell SO₂ emissions.) Second, 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 (see Figure C.1a.). In grid cells with very low average emissions (mostly rural areas), SO₂ emissions cannot be confounded with WMGHGs regardless of its correlation with WMGHGs for the simple reason that such low emissions would not affect precipitation. Ideally, in these grid cells, one would not include SO₂ in the analysis to avoid spurious results. On the other hand, grid cells with large emissions or grid cells downstream of and near enough to be affected by the source grid cells, correlations with WMGHGs are much weaker (again see Figure C.1). In this case, the potential causal effect of SO₂ is plausible, and hence one would include SO₂ in the analysis. Now, it would be quite difficult to figure out for what locations to include emissions and which not to when doing the analysis cell-by-cell. Using a CONUS-wide time series obviates this difficulty and furthermore provides a proxy for the emissions affecting each grid cell because emissions have generally trended up and then down similarly across the US due to broad societal reasons (i.e., industrialization and then regulations). For places unaffected by emissions, adding a variable that we know does not relate to emissions should not be problematic.

Fortunately, the AerChemMIP histSST experiments cover 1850 to 2014 and hence allow us to compare grid cell specific SO₂ with a single CONUS-wide time series. In doing so, we can explicitly test whether using an aggregate time series introduces significant biases for those grid cells with very low emissions. 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 Section 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 twice: first, using grid cell-specific SO₂ emissions time series, and second, using a single CONUS-wide average emissions time series (both of these time series are extracted from the model-specific emiso2 output). Then, conditional on the estimated SO₂ coefficient, we summarize the influence of SO₂ emissions on precipitation by calculating the change in precipitation under maximum versus minimum SO₂

Fig. 6 The CONUS-average effect of maximum versus minimum emiso2, comparing an analysis where we characterize SO_2 using model-specific grid cell emiso2 time series and a single CONUS-wide average emiso2 time series, with 95% basic bootstrap confidence intervals.

emissions (either local or CONUS-wide) at each grid cell and then take the 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 both characterizations of SO_2 in Figure 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 either method of characterizing SO_2 emissions. The spatial patterns of this effect (not shown) also agree extremely well for grid-cell versus CONUS-wide SO_2 . This result supports our confidence assessment in Appendix D that it is *virtually certain* (as precisely defined by Mastrandrea et al. [2010]) that a single CONUS-wide time series of emiso2 yields the same results as using local estimates of emiso2. Therefore, we can safely proceed with using a single CONUS-wide 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 Section 4.3.4, 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

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.). 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.

<i>(a) Mean</i>												
Driver	Bias				Relative Std. Dev.				Pattern Correlation			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
ELI	0.88	0.88	0.93	0.93	0.95	0.88	0.90	0.78	0.73	0.54	0.62	0.57
AO		0.97	0.93	0.97		0.90	0.85	0.95		0.68	0.81	0.59
NAO	0.97	0.93	0.88	0.97	0.88	0.88	0.85	0.97	0.70	0.51	0.62	0.49
PNA	0.95	0.90	0.97	0.90	0.93	0.95	0.90	0.95	0.57	0.51	0.59	0.62
AMO	0.90	0.95	0.90	0.95	0.90	0.80	0.85	0.90	0.65	0.70	0.76	0.84

<i>(b) GEV</i>												
Driver	Bias				Relative Std. Dev.				Pattern Correlation			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
ELI	0.85	0.95	0.95	0.97	0.95	0.82	0.70	0.82	0.51	0.54	0.84	0.51
AO		0.97	0.95	0.93		0.75	0.80	0.78		0.81	0.92	0.95
NAO	0.88	0.95	0.85	0.93	0.90	0.70	0.85	0.78	0.73	0.59	0.86	0.78
PNA	0.95	0.90	1.00	0.95	0.90	0.85	0.72	0.85	0.38	0.68	0.84	0.62
AMO	0.90	0.90	0.97	0.93	0.75	0.72	0.72	0.75	0.65	0.89	0.89	0.70

– historical and hist-noLu: climate drivers, WMGHGs, and SO₂ emissions

In light of the negligible secular trends identified in the hist-nat simulations (Figure 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

Fig. 7 CONUS-average bias in precipitation rate (left) and 20-year return value (right) coefficient estimates for the single-forcing experiments (hist-GHG and hist-aer, respectively) relative to the historical runs for WMGHG (top) and SO₂ (bottom), along with 95% basic bootstrap confidence intervals. The units for the WMGHG coefficient bias are mm per W m⁻²; units for the SO₂ coefficient bias are mm per 1000 kg m⁻² s⁻¹.

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 F.4-F.8 show the information summarized in Table 2 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 D 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 Section 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 Section 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 Section 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

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.

the experiment-specific estimates of the regression coefficients yield $\beta_{\text{WMGHG}}^{\text{hist-GHG}} - \beta_{\text{WMGHG}}^{\text{historical}} \sim 0$ and $\beta_{\text{SO}_2}^{\text{hist-aer}} - \beta_{\text{SO}_2}^{\text{historical}} \sim 0$ (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 SO_2 coefficients for the single-forcing experiments relative to the historical (i.e., single forcing minus historical) along with 95% basic bootstrap confidence intervals. For SO_2 , the bias is indistinguishable from zero for almost every model-season; for WMGHGs, the bias is somewhat more distinguishable from 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 β_{WMGHG} and β_{SO_2} , i.e.,

$$\begin{aligned} x(t) + y(t) &\rightarrow x(t) + y(t) + [\epsilon(t) - \epsilon(t)] = [x(t) + \epsilon(t)] + [y(t) - \epsilon(t)] \\ \beta_{\text{WMGHG}} &\rightarrow \frac{x(t) + \epsilon(t)}{F_{\text{WMGHG}}(t, \tau_{\text{WMGHG}})} \\ \beta_{\text{SO}_2} &\rightarrow \frac{y(t) - \epsilon(t)}{F_{\text{SO}_2}(t, \tau_{\text{SO}_2})} \end{aligned}$$

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 Figure 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 D 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.

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

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.

To address this hypothesis, we can utilize the C20C+ and HAPPI experiments described in Section 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 E.5) 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 ($+0.85^{\circ}\text{C}$), Plus15-Future ($+1.55^{\circ}\text{C}$), and Plus20-Future ($+2^{\circ}\text{C}$) experiments; Supplemental Figure F.9 shows corresponding plots for the total and background variability. For the models considered here, the percent background variability is roughly

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”). 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 D for more information on how each confidence statement is determined.

	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 pi-Control 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 precipitation as well as wetso2 and iso2. [†]	AerChemMIP (piClim experiments)	Yes	<i>Very likely</i>
H4d	A single CONUS-wide time series of emiso2 yields the same results as using local estimates of emiso2.	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>

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

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 F.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 Figure 9(b) 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 D 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 Figure 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. 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 D. 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].

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. 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.

Acknowledgments

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.

References

- T. P. Barnett, D. W. Pierce, K. M. AchutaRao, P. J. Gleckler, B. D. Santer, J. M. Gregory, and W. M. Washington. Penetration of human-induced warming into the world's oceans. *Science*, 309(5732):284–287, 2005. doi:10.1126/science.1112418.
- A. G. Barnston and R. E. Livezey. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review*, 115(6):1083–1126, jun 1987. ISSN 00270644. doi:10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2.
- Centre for Environmental Data Analysis (CEDA). CMIP6 data request. <http://clipc-services.ceda.ac.uk/dreq/mipVars.html>, 2016.
- J.-P. Chen, A. Hazra, C.-J. Shiu, I.-C. Tsai, and H.-H. Lee. *Interaction between Aerosols and Clouds: Current Understanding*, pages 231–281. World Scientific Publishing Co., 01 2008. ISBN 978-981-281-890-4. doi:10.13140/2.1.2481.7924.
- S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Lecture Notes in Control and Information Sciences. Springer, 2001. ISBN 978-1-85233-459-8. doi:10.1007/978-1-4471-3675-0.
- M. Collins, R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W. J. Gutowski, T. Johns, G. Krinner, et al. 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*, pages 1029–1136. Cambridge University Press, 2013. doi:10.1017/CBO9781107415324.024.
- W. J. Collins, J.-F. Lamarque, M. Schulz, O. Boucher, V. Eyring, M. I. Hegglin, A. Maycock, G. Myhre, M. Prather, D. Shindell, et al. AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6. *Geoscientific Model Development*, 10(2):585–607, 2017. doi:10.5194/gmd-10-585-2017.
- G. Dagan, I. Koren, O. Altaratz, and Y. Lehahn. Shallow convective cloud field lifetime as a key factor for evaluating aerosol effects. *iScience*, 10:192–202, 2018. ISSN 2589-0042. doi:10.1016/j.isci.2018.11.032.
- C. Deser. “Certain uncertainty: The role of internal climate variability in projections of regional climate change and risk management”. *Earth's Future*, 8(12), 2020. doi:10.1029/2020ef001854.
- S. Dong, Y. Sun, C. Li, X. Zhang, S.-K. Min, and Y.-H. Kim. Attribution of extreme precipitation with updated observations and cmip6 simulations. *Journal of Climate*, 34(3):871–881, 2021. doi:10.1175/JCLI-D-19-1017.1.
- D. Easterling, K. Kunkel, J. Arnold, T. Knutson, A. LeGrande, L. Leung, R. Vose, D. Waliser, and M. Wehner. Precipitation change in the United States. In: *Climate Science Special Report: Fourth National Climate Assessment, Volume I*, pages 207–230, 2017. doi:10.7930/J0H993CC.
- R. B. Esmaili, Y. Tian, D. A. Vila, and K.-M. Kim. A Lagrangian analysis of cold cloud clusters and their life cycles with satellite observations. *Journal of Geophysical Research: Atmospheres*, 121(19):11,723–11,738, 2016. doi:10.1002/2016JD025653.
- M. Etminan, G. Myhre, E. Highwood, and K. Shine. Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. *Geophysical Research Letters*, 43(24):12–614, 2016. doi:10.1002/2016GL071930.
- V. Eyring, S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958, 2016. doi:10.5194/gmd-9-1937-2016.
- M. Gidden, K. Riahi, S. Smith, S. Fujimori, G. Luderer, E. Kriegler, D. van Vuuren, M. van den Berg, L. Feng, D. Klein, K. Calvin, J. Doelman, S. Frank, O. Fricko, M. Harmsen, T. Hasegawa, P. Havlik, J. Hilaire, R. Hoesly, J. Horing, A. Popp, E. Stehfest, and K. Takahashi. input4mips.cmip6.scenariomip.iamc.iamc-image-ssp119-1-1, 2018. <https://doi.org/10.22033/ESGF/input4MIPs.2485>.

- N. P. Gillett, F. W. Zwiers, A. J. Weaver, and P. A. Stott. Detection of human influence on sea-level pressure. *Nature*, 422(6929):292–294, 2003. doi:10.1038/nature01487.
- N. P. Gillett, H. Shiogama, B. Funke, G. Hegerl, R. Knutti, K. Matthes, B. D. Santer, D. Stone, and C. Tebaldi. The detection and attribution model intercomparison project (DAMIP v1. 0) contribution to CMIP6. *Geoscientific Model Development*, 9(10):3685–3697, 2016. doi:10.5194/gmd-9-3685-2016.
- C. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969. doi:10.2307/1912791.
- J. M. Gregory, W. J. Ingram, M. A. Palmer, G. S. Jones, P. A. Stott, R. B. Thorpe, J. A. Lowe, T. C. Johns, and K. D. Williams. A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, 31(3), 2004. doi:10.1029/2003GL018747.
- R. Guo, C. Deser, L. Terray, and F. Lehner. Human influence on winter precipitation trends (1921–2015) over North America and Eurasia revealed by dynamical adjustment. *Geophysical Research Letters*, 46(6):3426–3434, 2019. doi:10.1029/2018GL081316.
- A. Hannart, J. Pearl, F. Otto, P. Naveau, and M. Ghil. Causal counterfactual theory for the attribution of weather and climate-related events. *Bulletin of the American Meteorological Society*, 97:99–110, 2015. doi:10.1175/BAMS-D-14-00034.1.
- D. Hartmann, A. Klein Tank, M. Rusticucci, L. Alexander, S. Brönnimann, Y. Charabi, F. Dentener, E. Dlugokencky, D. Easterling, A. Kaplan, B. Soden, P. Thorne, M. Wild, and P. Zhai. Observations: Atmosphere and surface. In T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, editors, *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, pages 159–254. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013. ISBN ISBN 978-1-107-66182-0. doi:10.1017/CBO9781107415324.008.
- G. C. Hegerl, K. Hasselmann, U. Cubasch, J. F. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz. Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. *Climate Dynamics*, 13(9):613–634, 1997. doi:10.1007/s003820050186.
- G. C. Hegerl, E. Black, R. P. Allan, W. J. Ingram, D. Polson, K. E. Trenberth, R. S. Chadwick, P. A. Arkin, B. B. Sarojini, A. Becker, et al. Challenges in quantifying changes in the global water cycle. *Bulletin of the American Meteorological Society*, 96(7):1097–1115, 2015. doi:10.1175/BAMS-D-13-00212.1.
- Ø. Hodnebrog, M. Etminan, J. Fuglestad, G. Marston, G. Myhre, C. Nielsen, K. P. Shine, and T. Wallington. Global warming potentials and radiative efficiencies of halocarbons and related compounds: A comprehensive review. *Reviews of Geophysics*, 51(2):300–378, 2013. doi:10.1002/rog.20013.
- R. M. Hoesly, S. J. Smith, L. Feng, Z. Klimont, G. Janssens-Maenhout, T. Pitkanen, J. J. Seibert, L. Vu, R. J. Andres, R. M. Bolt, T. C. Bond, L. Dawidowski, N. Kholod, J.-I. Kurokawa, M. Li, L. Liu, Z. Lu, M. C. P. Moura, P. R. O'Rourke, and Q. Zhang. Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development*, 11(1):369–408, 2018. doi:10.5194/gmd-11-369-2018.
- H. Huang, C. M. Patricola, J. M. Winter, E. C. Osterberg, and J. S. Mankin. Rise in Northeast US extreme precipitation caused by Atlantic variability and climate change. *Weather and Climate Extremes*, 33:100351, 2021. doi:10.1016/j.wace.2021.100351.
- J. W. Hurrell, Y. Kushnir, G. Ottersen, and M. Visbeck. An overview of the North Atlantic Oscillation. In *Geophysical Monograph Series*, volume 134, pages 1–35. American Geophysical Union, January 2003. ISBN 9781118669037. doi:10.1029/134GM.
- G. S. Jones, S. F. Tett, and P. A. Stott. Causes of atmospheric temperature change 1960–2000: A combined attribution analysis. *Geophysical Research Letters*, 30(5), 2003. doi:10.1029/2002GL016377.
- J. E. Kay, C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. M. Arblaster, S. C. Bates, G. Danabasoglu, J. Edwards, M. Holland, P. Kushner, J.-F. Lamarque, D. Lawrence, K. Lindsay, A. Middleton, E. Munoz, R. Neale, K. Oleson, L. Polvani, and M. Vertenstein. The Community Earth System Model (CESM) Large Ensemble Project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, 96(8):1333–1349, aug 2015. doi:10.1175/bams-d-13-00255.1.
- R. A. Kerr. A North Atlantic Climate Pacemaker for the Centuries. *Science*, 288(5473):1984–1985, 2000. doi:10.1126/science.288.5473.1984.
- V. V. Kharin, F. Zwiers, X. Zhang, and M. Wehner. Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, 119(2):345–357, 2013. doi:10.1007/s10584-013-0705-8.
- M. C. Kirchmeier-Young and X. Zhang. Human influence has intensified extreme precipitation in north america. *Proceedings of the National Academy of Sciences*, 117(24):13308–13313, 2020. doi:10.1073/pnas.1921628117.
- A. Kirkevåg, T. Iversen, J. E. Kristjansson, O. Seland, and J. B. Debernard. On the additivity of climate response to anthropogenic aerosols and CO₂, and the enhancement of future global warming by carbonaceous aerosols. *Tellus Series A-Dynamic Meteorology and Oceanography*, 60(3):513–527, 2008. doi:10.1111/j.1600-

0870.2008.00308.x.

- T. R. Knutson and F. Zeng. Model assessment of observed precipitation trends over land regions: Detectable human influences and possible low bias in model trends. *Journal of Climate*, 31(12):4617–4637, 2018. doi:10.1175/JCLI-D-17-0672.1.
- R. Knutti, D. Masson, and A. Gettelman. Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, 40(6):1194–1199, 2013. doi:10.1002/grl.50256.
- R. Knutti, J. Sedláček, B. M. Sanderson, R. Lorenz, E. M. Fischer, and V. Eyring. A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters*, 44:1908–1918, 2017. doi:10.1002/2016gl072012.
- K. E. Kunkel. North American trends in extreme precipitation. *Natural Hazards*, 29(2):291–305, 2003. doi:10.1023/A:1023694115864.
- K. E. Kunkel, D. R. Easterling, D. A. R. Kristovich, B. Gleason, L. Stoecker, and R. Smith. Meteorological Causes of the Secular Variations in Observed Extreme Precipitation Events for the Conterminous United States. *Journal of Hydrometeorology*, 13(3):1131–1141, 2012. doi:10.1175/jhm-d-11-0108.1.
- J.-F. Lamarque, F. Dentener, J. McConnell, C.-U. Ro, M. Shaw, R. Vet, D. Bergmann, P. Cameron-Smith, S. Dal-soren, R. Doherty, G. Faluvegi, S. J. Ghan, B. Josse, Y. H. Lee, I. A. MacKenzie, D. Plummer, D. T. Shindell, R. B. Skeie, D. S. Stevenson, S. Strode, G. Zeng, M. Curran, D. Dahl-Jensen, S. Das, D. Fritzsche, and M. Nolan. Multi-model mean nitrogen and sulfur deposition from the atmospheric chemistry and climate model intercom-parison project (ACCMIP): evaluation of historical and projected future changes. *Atmospheric Chemistry and Physics*, 13(16):7997–8018, 2013. doi:10.5194/acp-13-7997-2013.
- D. M. Lawrence, G. C. Hurtt, A. Arneth, V. Brovkin, K. V. Calvin, A. D. Jones, C. D. Jones, P. J. Lawrence, N. d. Noblet-Ducoudré, J. Pongratz, et al. The Land Use Model Intercomparison Project (LUMIP) contribu-tion to CMIP6: rationale and experimental design. *Geoscientific Model Development*, 9(9):2973–2998, 2016. doi:10.5194/gmd-9-2973-2016.
- C. Lee, R. V. Martin, A. van Donkelaar, H. Lee, R. R. Dickerson, J. C. Hains, N. Krotkov, A. Richter, K. Vinnikov, and J. J. Schwab. SO₂ emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations. *Journal of Geophysical Research: Atmospheres*, 116(D6), 2011. doi:10.1029/2010JD014758.
- L. Liu, D. Shawki, A. Voulgarakis, M. Kasoar, B. H. Samset, G. Myhre, P. M. Forster, Ø. Hodnebrog, J. Sillmann, S. G. Aalbergstjø, O. Boucher, G. Faluvegi, T. Iversen, A. Kirkevåg, J.-F. Lamarque, D. Olivie, T. Richardson, D. Shindell, and T. Takemura. A PDRMIP multimodel study on the impacts of regional aerosol forcings on global and regional precipitation. *Journal of Climate*, 31(11):4429 – 4447, 2018. doi:10.1175/JCLI-D-17-0439.1.
- U. Lohmann and J. Feichter. Global indirect aerosol effects: a review. *Atmospheric Chemistry and Physics*, 5(3): 715–737, 2005. doi:10.5194/acp-5-715-2005.
- B. Loring, T. A. O’Brien, A. E. Elbashandy, J. N. Johnson, H. Krishnan, N. Keen, Prabhat, and the CASCADE SFA. *Toolkit for Extreme Climate Analysis*. Lawrence Berkeley National Lab, 2016. <https://doi.org/10.20358/C8C651>.
- M. E. Mann, B. A. Steinman, and S. K. Miller. Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nature communications*, 11(1):49, 2020. doi:10.1038/s41467-019-13823-w.
- M. E. Mann, B. A. Steinman, D. J. Brouillette, and S. K. Miller. Multidecadal climate oscillations during the past millennium driven by volcanic forcing. *Science*, 371(6533):1014–1019, 2021. doi:10.1126/science.abc5810.
- K. Marvel, G. A. Schmidt, D. Shindell, C. Bonfils, A. N. LeGrande, L. Nazarenko, and K. Tsigaridis. Do responses to different anthropogenic forcings add linearly in climate models? *Environmental Research Letters*, 10(10), 2015. doi:10.1088/1748-9326/10/10/104010.
- M. D. Mastrandrea, C. B. Field, T. F. Stocker, O. Edenhofer, K. L. Ebi, D. J. Frame, H. Held, E. Kriegler, K. J. Mach, P. R. Matschoss, et al. 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.
- K. A. McKinnon and C. Deser. Internal variability and regional climate trends in an observational large ensemble. *Journal of Climate*, 31(17):6783–6802, sep 2018. doi:10.1175/jcli-d-17-0901.1.
- G. A. Meehl, C. A. Senior, V. Eyring, G. Flato, J.-F. Lamarque, R. J. Stouffer, K. E. Taylor, and M. Schlund. Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, 6(26), 2020. doi:10.1126/sciadv.aba1981.
- M. Meinshausen and Z. R. J. Nicholls. Uom-remind-magpie-ssp585-1-2-0: Remind-magpie-ssp585 GHG concen-trations, 2018. <https://doi.org/10.22033/ESGF/input4MIPs.2349>.
- M. Meinshausen and E. Vogel. input4mips.uom.ghgconcentrations.cmip.uom-cmip-1-2-0, 2016. <https://doi.org/10.22033/ESGF/input4MIPs.1118>.
- M. J. Menne, I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston. An overview of the Global Historical Climatology Network-Daily database. *Journal of Atmospheric and Oceanic Technology*, 29(7):897–910, 2012.

doi:10.1175/JTECH-D-11-00103.1.

- M. J. Mills, A. Schmidt, R. Easter, S. Solomon, D. E. Kinnison, S. J. Ghan, R. R. Neely III, D. R. Marsh, A. Conley, C. G. Bardeen, et al. Global volcanic aerosol properties derived from emissions, 1990–2014, using CESM1 (WACCM). *Journal of Geophysical Research: Atmospheres*, 121(5):2332–2348, 2016. doi:10.1002/2015JD024290.
- S.-K. Min, X. Zhang, F. W. Zwiers, and G. C. Hegerl. Human contribution to more-intense precipitation extremes. *Nature*, 470(7334):378, 2011. doi:10.1038/nature09763.
- D. Mitchell, K. AchutaRao, M. Allen, I. Bethke, U. Beyerle, A. Ciavarella, P. M. Forster, J. Fuglestad, N. Gillett, K. Haustein, W. Ingram, T. Iversen, V. Kharin, N. Klingaman, N. Massey, E. Fischer, C.-F. Schleussner, J. Scinocca, Ø. Seland, H. Shiogama, E. Shuckburgh, S. Sparrow, D. Stone, P. Uhe, D. Wallom, M. Wehner, and R. Zaaboul. Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design. *Geoscientific Model Development*, 10(2):571–583, 2017. doi:10.5194/gmd-10-571-2017.
- G. Myhre, P. M. Forster, B. H. Samset, Ø. Hodnebrog, J. Sillmann, S. G. Aalberg, T. Andrews, O. Boucher, G. Faluvegi, D. Flaschner, T. Iversen, M. Kasoar, V. Kharin, A. Kirkevåg, J.-F. Lamarque, D. Olivié, T. B. Richardson, D. Shindell, K. P. Shine, C. W. Stjern, T. Takemura, A. Voulgarakis, and F. Zwiers. PDRMIP: A precipitation driver and response model intercomparison project—protocol and preliminary results. *Bulletin of the American Meteorological Society*, 98(6):1185–1198, 2017. doi:10.1175/BAMS-D-16-0019.1.
- K. Noake, D. Polson, G. Hegerl, and X. Zhang. Changes in seasonal land precipitation during the latter twentieth-century. *Geophysical Research Letters*, 39(3), 2012. doi:10.1029/2011GL050405.
- J. P. O’Brien. *The quantification of co-occurring meteorological extremes and the anthropogenic contribution to hydrometeorological variation*. PhD thesis, University of California, Santa Cruz, 2019. URL <https://escholarship.org/uc/item/8fh4h62f>.
- C. Paciorek, D. Stone, and M. Wehner. Quantifying statistical uncertainty in the attribution of human influence on severe weather. *Weather and Climate Extremes*, 20:69–80, 2018. doi:10.1016/j.wace.2018.01.002.
- S. Paik, S.-K. Min, X. Zhang, M. G. Donat, A. D. King, and Q. Sun. Determining the anthropogenic greenhouse gas contribution to the observed intensification of extreme precipitation. *Geophysical Research Letters*, 47(12):e2019GL086875, 2020. doi:10.1029/2019GL086875.
- C. Patricola and M. Wehner. Anthropogenic influences on major tropical cyclone events. *Nature*, 563(7731):339, 2018. doi:10.1038/s41586-018-0673-2.
- A. G. Pendergrass, R. Knutti, F. Lehner, C. Deser, and B. M. Sanderson. Precipitation variability increases in a warmer climate. *Scientific Reports*, 7(1):17966, 2017. doi:10.1038/s41598-017-17966-y.
- A. S. Phillips, C. Deser, and J. Fasullo. Evaluating modes of variability in climate models. *Eos, Transactions American Geophysical Union*, 95(49):453–455, 2014. doi:10.1002/2014EO490002.
- L. Resplandy, R. F. Keeling, Y. Eddebbar, M. Brooks, R. Wang, L. Bopp, M. C. Long, J. P. Dunne, W. Koeve, and A. Oschlies. Quantification of ocean heat uptake from changes in atmospheric O₂ and CO₂ composition. *Scientific Reports*, 9(1):20244, 2019. doi:10.1038/s41598-019-56490-z.
- T. B. Richardson, P. M. Forster, T. Andrews, O. Boucher, G. Faluvegi, D. Flaschner, Ø. Hodnebrog, M. Kasoar, A. Kirkevåg, J.-F. Lamarque, G. Myhre, D. Olivié, B. H. Samset, D. Shawki, D. Shindell, T. Takemura, and A. Voulgarakis. Drivers of precipitation change: An energetic understanding. *Journal of Climate*, 31(23):9641–9657, 2018. doi:10.1175/JCLI-D-17-0240.1.
- K. L. Ricke and K. Caldeira. Maximum warming occurs about one decade after a carbon dioxide emission. *Environmental Research Letters*, 9(12):124002, 2014. doi:10.1088/1748-9326/9/12/124002.
- M. D. Risser and M. F. Wehner. Attributable human-induced changes in the likelihood and magnitude of the observed extreme precipitation during Hurricane Harvey. *Geophysical Research Letters*, 44(24):12,457–12,464, 2017. doi:10.1002/2017GL075888.
- M. D. Risser, C. J. Paciorek, T. A. O’Brien, M. F. Wehner, and W. D. Collins. Detected changes in precipitation extremes at their native scales derived from in situ measurements. *Journal of Climate*, 32(23):8087–8109, 2019a. doi:10.1175/JCLI-D-19-0077.1.
- M. D. Risser, C. J. Paciorek, M. F. Wehner, T. A. O’Brien, and W. D. Collins. A probabilistic gridded product for daily precipitation extremes over the United States. *Climate Dynamics*, 53(5):2517–2538, Sep 2019b. ISSN 1432-0894. doi:10.1007/s00382-019-04636-0.
- M. D. Risser, M. F. Wehner, J. P. O’Brien, C. M. Patricola, T. A. O’Brien, W. D. Collins, C. J. Paciorek, and H. Huang. Quantifying the influence of natural climate variability on in situ measurements of seasonal total and extreme daily precipitation. *Climate Dynamics*, 56:3205–3230, 2021. doi:10.1007/s00382-021-05638-7.
- B. H. Samset, G. Myhre, P. M. Forster, Ø. Hodnebrog, T. Andrews, G. Faluvegi, D. Flaschner, M. Kasoar, V. Kharin, A. Kirkevåg, J.-F. Lamarque, D. Olivié, T. Richardson, D. Shindell, K. P. Shine, T. Takemura, and A. Voulgarakis. Fast and slow precipitation responses to individual climate forcings: A PDRMIP multimodel study. *Geophysical Research Letters*, 43(6):2782–2791, 2016. doi:10.1002/2016GL068064.

- B. D. Santer, M. F. Wehner, T. Wigley, R. Sausen, G. Meehl, K. Taylor, C. Ammann, J. Arblaster, W. Washington, J. Boyle, et al. Contributions of anthropogenic and natural forcing to recent tropopause height changes. *science*, 301(5632):479–483, 2003. doi:10.1126/science.1084123.
- B. B. Sarojini, P. A. Stott, E. Black, and D. Polson. Fingerprints of changes in annual and seasonal precipitation from CMIP5 models over land and ocean. *Geophysical Research Letters*, 39(21), 2012. doi:10.1029/2012GL053373.
- B. B. Sarojini, P. A. Stott, and E. Black. Detection and attribution of human influence on regional precipitation. *Nature Climate Change*, 6(7):669–675, 2016. doi:10.1038/nclimate2976.
- M. Sato, J. E. Hansen, M. P. McCormick, and J. B. Pollack. Stratospheric aerosol optical depths, 1850–1990. *Journal of Geophysical Research: Atmospheres*, 98(D12):22987–22994, 1993. doi:10.1029/93JD02553.
- M. E. Schlesinger and N. Ramankutty. An oscillation in the global climate system of period 65–70 years. *Nature*, 367(6465):723–726, feb 1994. ISSN 00280836. doi:10.1038/367723a0.
- A. Schmidt, M. J. Mills, S. Ghan, J. M. Gregory, R. P. Allan, T. Andrews, C. G. Bardeen, A. Conley, P. M. Forster, A. Gettelman, et al. Volcanic radiative forcing from 1979 to 2015. *Journal of Geophysical Research: Atmospheres*, 123(22):12–491, 2018. doi:10.1029/2018JD028776.
- H. Shiogama, D. A. Stone, T. Nagashima, T. Nozawa, and S. Emori. On the linear additivity of climate forcing-response relationships at global and continental scales. *International Journal of Climatology*, 33(11):2542–2550, 2013. doi:10.1002/joc.3607.
- K. E. Taylor. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7):7183–7192, 2001. doi:10.1029/2000JD900719.
- S. F. Tett, P. A. Stott, M. R. Allen, W. J. Ingram, and J. F. Mitchell. Causes of twentieth-century temperature change near the Earth’s surface. *Nature*, 399(6736):569–572, 1999. doi:10.1038/21164.
- D. W. J. Thompson and J. M. Wallace. Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability. *Journal of Climate*, 13(5):1000–1016, mar 2000. ISSN 0894-8755. doi:10.1175/1520-0442.
- G. D. Thornhill, W. J. Collins, R. J. Kramer, D. Oliv, R. B. Skeie, F. M. O’Connor, N. L. Abraham, R. Checa-Garcia, S. E. Bauer, M. Deushi, L. K. Emmons, P. M. Forster, L. W. Horowitz, B. Johnson, J. Keeble, J.-F. Lamarque, M. Michou, M. J. Mills, J. P. Mulcahy, G. Myhre, P. Nabat, V. Naik, N. Oshima, M. Schulz, C. J. Smith, T. Takemura, S. Tilmes, T. Wu, G. Zeng, and J. Zhang. Effective radiative forcing from emissions of reactive gases and aerosols – a multi-model comparison. *Atmospheric Chemistry and Physics*, 21(2):853–874, 2021. doi:10.5194/acp-21-853-2021.
- J. M. Wallace and D. S. Gutzler. Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter. *Monthly Weather Review*, 109(4):784–812, apr 1981. ISSN 0027-0644. doi:10.1175/1520-0493.
- S. S. Wang, L. Zhao, J.-H. Yoon, P. Klotzbach, and R. R. Gillies. Quantitative attribution of climate effects on hurricane harvey’s extreme rainfall in texas. *Environmental Research Letters*, 13(5):054014, 2018. doi:10.1088/1748-9326/aabb85.
- L. J. Wilcox, E. J. Highwood, and N. J. Dunstone. The influence of anthropogenic aerosol on multi-decadal variations of historical global climate. *Environmental Research Letters*, 8(2):024033, 2013. doi:10.1088/1748-9326/8/2/024033.
- I. N. Williams and C. M. Patricola. Diversity of ENSO events unified by convective threshold sea surface temperature: a nonlinear ENSO index. *Geophysical Research Letters*, 45(17):9236–9244, 2018. doi:10.1029/2018GL079203.
- P. Wu, N. Christidis, and P. Stott. Anthropogenic impact on Earth’s hydrological cycle. *Nature Climate Change*, 3(9):807–810, 2013. doi:10.1038/nclimate1932.
- Y. Yang, H. Wang, S. J. Smith, R. Zhang, S. Lou, H. Yu, C. Li, and P. J. Rasch. Source Apportionments of Aerosols and Their Direct Radiative Forcing and Long-Term Trends Over Continental United States. *Earth’s Future*, 6(6):793–808, 2018. doi:10.1029/2018EF000859.
- F. Zhang, Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel. What is the predictability limit of midlatitude weather? *Journal of the Atmospheric Sciences*, 76(4):1077–1091, 2019. doi:10.1175/jas-d-18-0269.1.
- X. Zhang, F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa. Detection of human influence on twentieth-century precipitation trends. *Nature*, 448(7152):461–465, 2007. doi:10.1038/nature06025.
- X. Zhang, H. Wan, F. W. Zwiers, G. C. Hegerl, and S.-K. Min. Attributing intensification of precipitation extremes to human influence. *Geophysical Research Letters*, 40(19):5252–5257, 2013. doi:10.1002/grl.51010.

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

A.1 Lagged effects of forcing on precipitation

As Samset et al. [2016] showed, the response of global precipitation to forcing by CO_2 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:

$$\frac{P(t)}{P_0} - 1 = R\gamma \delta T(t) \quad (\text{A.1})$$

$$P_0 = P(t = 0)$$

$$R = \text{Unitless ratio of regional-to-global precipitation}$$

$$\gamma = \text{Global precipitation sensitivity: } 2\%/K \text{ for mean and } 6\%/K \text{ for extremes}$$

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

$$\rho c_p h \frac{d\delta T(t)}{dt} = F(t) - \alpha \delta T(t)$$

$$\rho = \text{Density of ocean water}$$

$$c_p = \text{Specific heat at constant pressure of ocean water}$$

$$h = \text{Depth of ocean}$$

$$F(t) = \text{Climate forcing}$$

$$\alpha = \text{Climate feedback parameter} \quad [W/m^2/K]$$

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

$$\frac{d\delta T(t)}{dt} + \frac{\delta T(t)}{\tau} = \frac{F(t)}{\rho c_p h} \quad (\text{A.2})$$

The solution is:

$$\delta T(t) = \alpha^{-1} F(t, \tau) \quad (\text{A.3})$$

$$F(t, \tau) = \int_0^t e^{(t'-t)/\tau} F(t') \frac{dt'}{\tau} \quad (\text{A.4})$$

In the limit $\tau = 0$, one can show that, as expected from Eq. A.2 when the system is in equilibrium,

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

This follows from Eqs. A.3 and A.4 if we define an integration variable $x = (t - t')/\tau$ so that

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

Substituting Eq. A.3 into Eq. A.1 gives an expression for the β coefficients that appear, for example, in Eq. 2 for the forced component:

$$\frac{P(t)}{P_0} - 1 = \beta(\tau) F(t, \tau) \quad (\text{A.5})$$

$$\beta(\tau) = R(\tau) \frac{\gamma}{\alpha} \quad (\text{A.6})$$

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

The lagged time series $F(t, \tau)$ calculated from Equation A.4 using the WMGHG forcing time series $F(t)$ from Figure 2 is shown in Figure A.1 for $\tau = 1, \dots, 20$.

Fig. A.1 The lagged WMGHG forcing $F(t, \tau)$ calculated from Equation A.4 for $\tau \in \{1, \dots, 20\}$ (panel 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 (panel b.). Estimates in panel (b) include 95% basic bootstrap confidence intervals. Note that panel (a) shows the decreasing lagged forcing with increasing lag as derived in Equation A.7 and panel (b) shows the increasing trend in $\beta(\tau)$ with increasing lag derived in Equation A.10.

A.2 Decreasing lagged forcing with increasing lag timescales

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

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

$$\frac{dF(t, \tau)}{d\tau} \leq 0. \quad (\text{A.7})$$

If $P(t)$ is prescribed and the forcing obeys the assumptions used to obtain Eq. A.7, then Eqs. A.5 and A.7 imply that

$$\frac{d}{d\tau} \left[\frac{P(t)}{P_0} - 1 \right] = 0 = \frac{d\beta(\tau)}{d\tau} F(t, \tau) + \beta(\tau) \frac{dF(t, \tau)}{d\tau} \quad (\text{A.8})$$

$$\frac{1}{\beta(\tau)} \frac{d\beta(\tau)}{d\tau} = -\frac{1}{F(t, \tau)} \frac{dF(t, \tau)}{d\tau} \quad (\text{A.9})$$

$$\therefore \frac{d \ln \beta(\tau)}{d\tau} \geq 0 \quad (\text{A.10})$$

The positive trend in $\beta(\tau)$ with respect to τ is observed the DAMIP WMGHG-only runs (Figure A.1).

A.3 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

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

Fig. A.2 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 Equation A.15 with a 95% uncertainty band and the blue line denotes the breakpoint $\tau_{b,i}$.

1048 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
1049 1st law of thermodynamics, is

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

1050 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
1051 et al., 2004, Meehl et al., 2020]. Then the models typically shift to a much slower rate of change.

1052 The time-series of $\ln N_i(t)$ from the abrupt4xCO2 MME indexed by i for the various models and ensemble members is computed
1053 via

$$N_i(t) = \text{rsdt}_i(t) - \text{rsut}_i(t) - \text{rlut}_i(t) \quad (\text{A.11})$$

$$\text{rsdt}_i(t) = \text{TOA incident shortwave radiation} \quad [W/m^2] \quad (\text{A.12})$$

$$\text{rsut}_i(t) = \text{TOA outgoing shortwave radiation} \quad [W/m^2] \quad (\text{A.13})$$

$$\text{rlut}_i(t) = \text{TOA outgoing longwave radiation} \quad [W/m^2] \quad (\text{A.14})$$

1054 where the radiative fluxes on the RHS of Eq. A.11 have been globally and annually averaged and are denoted by the names of CMIP6
1055 output variables [Centre for Environmental Data Analysis (CEDA), 2016]. We have determined three timescales $\tau_{f,i}$, $\tau_{s,i}$, and $\tau_{b,i}$
1056 characterizing the “fast” and “slow” evolution of $N_i(t)$ and the “break” between these two regimes as follows:

$$\ln N_i(t) = \ln F_0 - \frac{t}{\tau_{f,i}} \Theta(\tau_{b,i} - t) - \frac{t}{\tau_{s,i}} \Theta(t - \tau_{b,i}) \quad (\text{A.15})$$

1057 where Θ is the Heaviside function and $F_0 = 8.053$ is determined from the formulae in Etminan et al. [2016].

1058 We use a Bayesian least squares fit to estimate the timescales $\tau_{f,i}$, $\tau_{s,i}$, and $\tau_{b,i}$ in Equation A.15 using the $N_i(t)$ time series
1059 from each GCM with abrupt4xCO2 ensembles in Table E.2, aggregating over multiple ensemble members when available. A Bayesian
1060 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
1061 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
1062 from the two-member CanESM5 ensemble and six-member MRI-ESM2-0 ensemble; see Figure A.2. The natural log of the net radiative
1063 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
1064 lagged WMGHG forcing time series $F(t, \tau_i)$ is set equal to $\tau_{f,i}$ if $\tau_{f,i} < 33$ (the 90th percentile of the multi-model ensemble) and $\tau_{b,i}$
1065 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
1066 CMIP6 models with abrupt4xCO2 data are shown in Table A.1. For models without abrupt4xCO2 data (e.g., FGOALS-g3), we use the
1067 CMIP6 multi-model ensemble average of $\bar{\tau} = 14$.

1068 B Test of H1 under significantly increased warming

1069 To assess our test of H1 under much greater levels of warming than is imposed on the CMIP6 historical experiment, we can utilize
1070 climate model runs from the CMIP6 DECK experiments piControl and 1pctCO2 introduced in Section 3.1. For each CMIP6 model
1071 considered, we first extract the global average of mean daily precipitation and Rx1Day as well as monthly surface temperature for each
1072 year in both the piControl and 1pctCO2 runs. Furthermore, we derive the time series of WMGHG forcing for the 1pctCO2 runs using the
1073 specified boundary conditions for the 1pctCO2 experiment and various formulae described in Section 3.2. We then estimate the changes

Table A.1 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 $F(t, \tau)$. 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 E.2), we use the CMIP6 multi-model ensemble average of $\bar{\tau} = 14$.

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

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 Equations 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 Figure B.1, 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 Figure B.1(a) and (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.

C Correlations between WMGHGs and SO₂ emissions for 1900-present

As was mentioned in Section 4.3.4, 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 counterintuitive 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 Section 4.3.4.

The average annual SO₂ emissions at each grid cell are shown in Figure C.1(a) and (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 Figure C.1e.), 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 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

Fig. B.1 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).

with the WMGHG time series (Figure C.1b.), 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 C.1(c) and (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 Figure C.1(e) tend to correspond to grid cells with small average emissions. This fact explains the counterintuitive result wherein the grid cell correlations between SO₂ emissions and WMGHGs can be so strong while the CONUS-wide SO₂ 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 Confidence statements for each hypothesis

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

Fig. C.1 Average annual SO_2 emissions for (panels a. and d.), with grid cell correlations for annual SO_2 emissions versus the time series of WMGHGs for (panels b. and e.) and a scatter plot of average annual emissions and WMGHGs (panels c. and f.). The top row corresponds to 1850-2015 and the bottom row corresponds to 1900-2015. The red lines in panels (c) and (f) show the correlation between the CONUS-wide SO_2 emissions and WMGHGs.

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 Figure 3(b). 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 Figure 3(b): $1.7\% K^{-1}$ 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 Figure 3(c). 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 Figure 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 Figure 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 4d. The quantities of interest here are the CONUS-average effect of maximum vs. minimum emiso2 on precipitation using grid-cell vs. a CONUS-wide time series of SO₂ emissions shown in Figure 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 0.82, 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 Figure 7; and second, the Pearson correlations in the WMGHG vs. SO₂ coefficient biases shown in Figure 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	65	0.74	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 Figure 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

1182

1183 E Supplemental tables

Table E.1 Number of ensemble members (“Ens”) and years (“Yrs”) for each of the CMIP6 models that provide daily precipitation rate output for historical runs as well as the various DAMIP (hist-GHG, hist-aer, hist-nat, and hist-stratO3) and LUMIP (hist-noLu) experiments. Generally, the historical and hist-noLu runs cover 1851-2014, while the DAMIP experiments cover 1851-2020.

Model	historical		hist-GHG		hist-aer		hist-nat		hist-stratO3		hist-noLu	
	Ens	Yrs	Ens	Yrs	Ens	Yrs	Ens	Yrs	Ens	Yrs	Ens	Yrs
ACCESS-ESM1-5	10	1540	3	510	3	410	3	510	–	–	–	–
BCC-CSM2-MR	3	492	3	510	3	510	3	510	–	–	1	164
CanESM5	50	8200	10	1700	10	1700	10	1700	10	1700	–	–
CESM2	9	1485	2	330	2	329	3	495	–	–	3	495
FGOALS-g3	5	830	3	510	3	510	3	510	–	–	–	–
GFDL-ESM4	1	164	1	170	1	170	1	170	–	–	–	–
HadGEM3-GC31-LL	4	656	4	680	4	680	4	680	–	–	–	–
IPSL-CM6A-LR	32	5248	9	1530	9	1530	9	1530	7	1190	4	656
MIROC6	10	1640	3	510	3	510	3	510	3	510	–	–
MRI-ESM2-0	5	820	5	850	5	850	5	850	3	510	–	–
NorESM2-LM	3	492	3	510	3	510	3	510	–	–	–	–

Table E.2 Number of ensemble members for the abrupt4xCO₂ (N_{abrupt4x}) and 1pctCO₂ (N_{1pct}) for each of the CMIP6 models that provide the requisite data. The CMIP6 DECK criteria requires a minimum of 150 years for each experiment.

Model	N_{abrupt4x}	N_{1pct}	Model	N_{abrupt4x}	N_{1pct}
ACCESS-CM2	1	1	GISS-E2-1-H	3	–
ACCESS-ESM1-5	2	1	GISS-E2-2-G	1	1
AWI-CM-1-1-MR	1	–	HadGEM3-GC31-LL	1	4
BCC-CSM2-MR	1	–	HadGEM3-GC31-MM	1	1
BCC-ESM1	1	–	IITM-ESM	1	1
CAMS-CSM1-0	2	–	INM-CM4-8	1	1
CanESM5	2	6	INM-CM5-0	1	1
CESM2	1	1	IPSL-CM5A2-INCA	1	–
CESM2-FV2	1	–	IPSL-CM6A-LR	1	1
CESM2-WACCM	1	1	IPSL-CM6A-LR-INCA	1	–
CESM2-WACCM-FV2	1	–	KACE-1-0-G	1	–
CIESM	1	–	KIOST-ESM	1	–
CMCC-CM2-HR4	1	–	MIROC-ES2L	1	1
CMCC-CM2-SR5	1	1	MIROC6	1	1
CMCC-ESM2	1	–	MPI-ESM-1-2-HAM	1	1
CNRM-CM6-1	1	1	MPI-ESM1-2-HR	1	1
CNRM-CM6-1-HR	1	1	MPI-ESM1-2-LR	1	1
CNRM-ESM2-1	3	10	MRI-ESM2-0	6	2
E3SM-1-0	1	–	NorCPM1	1	–
EC-Earth3	2	1	NorESM2-LM	1	1
EC-Earth3-AerChem	1	–	NorESM2-MM	1	1
EC-Earth3-CC	1	–	TaiESM1	1	–
EC-Earth3-Veg	1	1	UKESM1-0-LL	1	4
FIO-ESM-2-0	3	–	FGOALS-f3-L	–	3
GFDL-CM4	1	–	FGOALS-g3	–	3
GFDL-ESM4	1	1	NESM3	–	1
GISS-E2-1-G	4	–	SAM0-UNICON	–	1

Table E.3 Number of years and ensemble members for each AerChemMIP piClim experiment that have daily precipitation rate and monthly aerosol data.

Model	# Years	piClim-control				piClim-aer			
		<i>pr</i>	<i>emiso2</i>	<i>iso2</i>	<i>wetso2</i>	<i>pr</i>	<i>emiso2</i>	<i>iso2</i>	<i>wetso2</i>
ACCESS-CM2	30	1	0	0	0	1	0	0	0
ACCESS-ESM1-5	30	1	0	0	0	1	0	0	0
BCC-ESM1	31	0	0	1	0	0	0	1	0
CanESM5	50	1	1	0	1	1	1	0	1
CESM2	31	1	0	0	0	1	1	0	1
CESM2-WACCM	31	1	1	0	1	0	0	0	0
CNRM-CM6-1	30	1	0	0	0	1	0	0	0
CNRM-ESM2-1	30	1	1	1	0	1	1	0	0
EC-Earth3	30	1	0	0	0	1	0	0	0
EC-Earth3-AerChem	31	0	1	1	1	0	0	0	0
GISS-E2-1-G	41	0	1	1	1	0	1	1	1
HadGEM3-GC31-LL	30	1	0	0	0	1	0	0	0
IPSL-CM6A-LR	150	5	0	0	0	4	0	0	0
IPSL-CM6A-LR-INCA	30	1	1	0	0	1	1	0	0
MIROC6	30	0	1	0	1	0	1	0	1
MPI-ESM-1-2-HAM	30	1	1	1	1	1	1	1	1
MPI-ESM1-2-LR	91	3	0	0	0	0	0	0	0
MRI-ESM2-0	31	1	1	1	1	1	1	1	1
NorESM2-LM	60	2	2	2	2	2	2	2	2
NorESM2-MM	30	1	1	1	1	1	1	1	1
UKESM1-0-LL	45	1	1	0	1	1	1	0	1

Model	piClim-SO2				piClim-BC	piClim-OC	piClim-NH3
	<i>pr</i>	<i>emiso2</i>	<i>iso2</i>	<i>wetso2</i>			
ACCESS-CM2	0	0	0	0	0	0	0
ACCESS-ESM1-5	0	0	0	0	0	0	0
BCC-ESM1	0	0	1	0	0	0	0
CanESM5	0	0	0	0	0	0	0
CESM2	0	0	0	0	0	0	0
CESM2-WACCM	0	0	0	0	0	0	0
CNRM-CM6-1	0	0	0	0	0	0	0
CNRM-ESM2-1	0	1	1	0	0	0	0
EC-Earth3	0	0	0	0	0	0	0
EC-Earth3-AerChem	0	0	0	0	0	0	0
GISS-E2-1-G	0	1	1	1	0	0	0
HadGEM3-GC31-LL	0	0	0	0	0	0	0
IPSL-CM6A-LR	0	0	0	0	0	0	0
IPSL-CM6A-LR-INCA	1	1	0	1	1	1	1
MIROC6	1	1	0	1	1	1	0
MPI-ESM-1-2-HAM	1	1	1	1	1	1	0
MPI-ESM1-2-LR	0	0	0	0	0	0	0
MRI-ESM2-0	1	1	1	1	1	1	0
NorESM2-LM	0	2	2	2	0	0	0
NorESM2-MM	0	0	0	0	0	0	0
UKESM1-0-LL	1	2	0	2	1	1	0

Table E.4 Number of ensemble members (“Ens”) and years (“Yrs”) for AerChemMIP histSST models that provide daily precipitation rate and monthly aerosol data. Generally, these runs cover 1851-2014. Note: GISS-E2-1-G does not have daily precipitation, and the daily precipitation for MRI-ESM2-0 is on a different grid than the emiso2 data; hence, these two models are excluded.

Model	histSST	
	<i>Ens</i>	<i>Yrs</i>
CESM2-WACCM	1	164
CNRM-ESM2-1	1	164
MIROC6	1	164
MPI-ESM-1-2-HAM	1	158
NorESM2-LM	1	164
UKESM1-0-LL	1	164

Table E.5 Number of ensemble members (“Ens”) and years (“Yrs”) for models that have daily precipitation for at least 50 ensemble members of at least 10 years for each of C20C+ All-Hist and HAPPI Plus15-Future and Plus20-Future.

Model	Experiment	Ens	Yrs
MIROC5	All-Hist	100	11
CanAM4	All-Hist	100	10
CAM4-2degree	All-Hist	501	10
ECHAM6-3-LR	All-Hist	100	10
NorESM1-HAPPI	All-Hist	125	10
MIROC5	Plus15-Future	100	11
CanAM4	Plus15-Future	100	10
CAM4-2degree	Plus15-Future	501	10
ECHAM6-3-LR	Plus15-Future	100	10
NorESM1-HAPPI	Plus15-Future	125	10
MIROC5	Plus20-Future	100	11
CanAM4	Plus20-Future	100	10
CAM4-2degree	Plus20-Future	501	10
ECHAM6-3-LR	Plus20-Future	100	10
NorESM1-HAPPI	Plus20-Future	81	10

1184 **F Supplemental figures**

Fig. F.1 A comparison of the CONUS-wide SO₂ emissions (1000 Gg per year) from the input4MIPs data set (black), Organization for Economic Co-operation and Development (https://stats.oecd.org/Index.aspx?DataSetCode=AIR_EMISSIONS), and the Environmental Protection Agency (<https://www.statista.com/statistics/501303/volume-of-sulfur-dioxide-emissions-us/>).

Fig. F.2 Change in precipitation as a function of each aerosol species (piClim-* minus piClim-control) with 95% basic bootstrap confidence interval for all model grid cells over Northern Hemisphere land regions. Across most combinations of model/season/precipitation type, the sum of BC, OC, and NH3 is indistinguishable from zero. On the other hand, it is possible that SO₂ has a larger influence than the sum of the other three species across all combinations.

Fig. F.3 DJF area-weighted CONUS averages of SO₂ fields from the AerChemMIP histSST runs for surface emissions (emiso2), atmospheric-column-integrated abundances (iso2), dry deposition rate to the Earth's surface (dryso2), wet deposition rate to the Earth's surface (wetso2), and the sum of the deposition rates. Notice that the emiso2 time series are nearly identical across models, whereas the other fields (e.g., wetso2) show significant variability across models.

Fig. F.4 CONUS-wide bias (panel a.) and Taylor diagram statistics (relative standard deviation, panel b., and pattern correlation, panel c.) for the ENSO Longitude Index coefficient relative to the model-specific hist-nat ensemble across season, precipitation type, experiment, and model, along with 95% bootstrap confidence intervals.

Fig. F.5 CONUS-wide bias (panel a.) and Taylor diagram statistics (relative standard deviation, panel b., and pattern correlation, panel c.) for the Arctic Oscillation coefficient relative to the model-specific hist-nat ensemble across season (excluding DJF), precipitation type, experiment, and model, along with 95% bootstrap confidence intervals.

Fig. F.6 CONUS-wide bias (panel a.) and Taylor diagram statistics (relative standard deviation, panel b., and pattern correlation, panel c.) for the North Atlantic Oscillation coefficient relative to the model-specific hist-nat ensemble across season, precipitation type, experiment, and model, along with 95% bootstrap confidence intervals.

Fig. F.7 CONUS-wide bias (panel a.) and Taylor diagram statistics (relative standard deviation, panel b., and pattern correlation, panel c.) for the Pacific North American pattern coefficient relative to the model-specific hist-nat ensemble across season, precipitation type, experiment, and model, along with 95% bootstrap confidence intervals.

Fig. F8 CONUS-wide bias (panel a.) and Taylor diagram statistics (relative standard deviation, panel b, and pattern correlation, panel c.) for the Atlantic Multidecadal Oscillation coefficient relative to the model-specific hist-nat ensemble across season, precipitation type, experiment, and model, along with 95% bootstrap confidence intervals.

Fig. F.9 Total (panel a.) and background (panel b.) variability across models and seasons for the All-Hist experiment and trends in these quantities (panels c. and d.) along with 95% bootstrap confidence intervals. Trends are calculated using the All-Hist and HAPPI experiments.