AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2013JD020648

Special Section:
The Geoengineering Model Intercomparison Project (GeoMIP)

Key Points:
- The G1 experiment features significant changes in climate extremes
- Rapid climate responses in G1 lead to significant regional warming over land
- Extreme temperatures decrease and cold spells increase over oceans in G1

Supporting Information:
- Readme
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6
- Figure S7
- Figure S8
- Figure S9
- Figure S10
- Tables S1 and S2

Correspondence to:
C. L. Curry,
cc@uvic.ca

Citation:
Curry, C. L., et al. (2014), A multimodel examination of climate extremes in an idealized geoengineering experiment, *J. Geophys. Res. Atmos.*, *119*, 3900–3923, doi:10.1002/2013JD020648.

Received 29 JUL 2013
Accepted 27 FEB 2014
Accepted article online 3 MAR 2014
Published online 14 APR 2014

A multimodel examination of climate extremes in an idealized geoengineering experiment

Charles L. Curry[1], Jana Sillmann[1,2], David Bronaugh[3], Kari Alterskjaer[2], Jason N. S. Cole[4], Duoying Ji[5], Ben Kravitz[6], Jón Egill Kristjánsson[2], John C. Moore[5], Helene Muri[2], Ulrike Niemeier[7], Alan Robock[8], Simone Tilmes[9], and Shuting Yang[10]

[1]School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada, [2]Department of Geosciences, University of Oslo, Oslo, Norway, [3]Pacific Climate Impacts Consortium, Victoria, British Columbia, Canada, [4]Canadian Centre for Climate Modelling and Analysis, Environment Canada, Toronto, Ontario, Canada, [5]State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [6]Pacific Northwest National Laboratory, Richland, Washington, USA, [7]Max Planck Institute for Meteorology, Hamburg, Germany, [8]Department of Environmental Sciences, Rutgers, State University of New Jersey, New Brunswick, New Jersey, USA, [9]National Center for Atmospheric Research, Boulder, Colorado, USA, [10]Danish Meteorological Institute, Copenhagen, Denmark

Abstract Temperature and precipitation extremes are examined in the Geoengineering Model Intercomparison Project experiment *G1*, wherein an instantaneous quadrupling of CO_2 from its preindustrial control value is offset by a commensurate reduction in solar irradiance. Compared to the preindustrial climate, changes in climate extremes under *G1* are generally much smaller than under $4 \times CO_2$ alone. However, it is also the case that extremes of temperature and precipitation in *G1* differ significantly from those under preindustrial conditions. Probability density functions of standardized anomalies of monthly surface temperature $\mathcal{T}$ and precipitation $\mathcal{P}$ in *G1* exhibit an extension of the high-$\mathcal{T}$ tail over land, of the low-$\mathcal{T}$ tail over ocean, and a shift of $\mathcal{P}$ to drier conditions. Using daily model output, we analyzed the frequency of extreme events, such as the coldest night (*TNn*), warmest day (*TXx*), and maximum 5 day precipitation amount, and also duration indicators such as cold and warm spells and consecutive dry days. The strong heating at northern high latitudes simulated under $4 \times CO_2$ is much alleviated in *G1*, but significant warming remains, particularly for *TNn* compared to *TXx*. Internal feedbacks lead to regional increases in absorbed solar radiation at the surface, increasing temperatures over Northern Hemisphere land in summer. Conversely, significant cooling occurs over the tropical oceans, increasing cold spell duration there. Globally, *G1* is more effective in reducing changes in temperature extremes compared to precipitation extremes and for reducing changes in precipitation extremes versus means but somewhat less effective at reducing changes in temperature extremes compared to means.

1. Introduction

The persistent rise in atmospheric greenhouse gas concentrations over the last century, and the climate change it has wrought, has prompted discussions beyond those focusing solely on efforts to curtail greenhouse gas emissions. The notion of a more purposeful alteration of the Earth's radiation balance on the global scale, known as climate engineering or geoengineering, has received increased attention in recent years. In principle, there are many means by which this might be achieved [e.g., *Royal Society*, 2009]. Perhaps the simplest conceptual scheme is one in which incoming solar energy at the top of the planet's atmosphere is reduced by a certain fraction. Whether such a straightforward form of solar radiation management (SRM) could be achieved in practice is open to question [*Angel*, 2006; *Royal Society*, 2009], and the ethical dilemmas of intentional alteration of the Earth's climate continue to be widely discussed [e.g., *Robock et al.*, 2009; *Gardiner*, 2010].

The scientific investigation of the effect of SRM on the climate system has been undertaken using a variety of modeling approaches, ranging from simple zero-dimensional energy balance models [*Lenton and Vaughan*, 2009] to more complex simulations that have exposed marked differences in regional climate response to these equal but opposite global forcings [*Irvine et al.*, 2010]. In an effort to compare these modeling results on a more equal footing, the Geoengineering Model Intercomparison Project (GeoMIP) was

conceived [*Kravitz et al.*, 2011]. The overall aim of the project is to provide an experimental protocol for assessing the impact of offsetting increased longwave radiation due to an increase in CO_2 concentration by a reduction of incoming shortwave radiation through modification of the solar irradiance or stratospheric sulfate aerosol concentrations. The GeoMIP protocol comprises both idealized experiments designed to gauge the effect of simplified solar irradiance changes (known as *G1* and *G2*) and more realistic simulations including stratospheric aerosols (*G3* and *G4*). Further details of the GeoMIP motivation and methodology are provided in *Kravitz et al.* [2011].

The object of this work is to examine model-simulated changes in extreme temperature and precipitation events as a consequence of reducing solar irradiance in one of the GeoMIP experiments, *G1*. Although a highly idealized experiment, the behavior of models under the *G1* forcings can illustrate some of the potential consequences of a deliberate modification of the Earth's climate. Moreover, while the experimental design is simplified, the expected response is not: the use of state-of-the-art coupled Earth system models (identical to those used in the Coupled Model Intercomparison Project, CMIP5) ensures that the applied forcing triggers a unique response in each of the individual climate system components (i.e., atmosphere, land, ocean, sea ice, biogeochemical systems on land and ocean, etc.). For the purpose of comparison, we also examine results for the same climate variables under the CMIP5 *abrupt4* $\times$ CO_2 scenario, in which only the positive forcing, leading to widespread warming, is applied. Model-projected changes in climate extremes under several different warming scenarios were examined by *Tebaldi et al.* [2006], *Orlowsky and Seneviratne* [2012], *Caesar and Lowe* [2012], and *Sillmann et al.* [2013a].

The few studies completed to date on the multimodel ensemble of GeoMIP output have focused primarily on characteristics of the temporal means of diverse climate variables (e.g., temperature, sea level pressure, precipitation and other elements of the hydrological cycle, and terrestrial net primary productivity [*Schmidt et al.*, 2012; *Kravitz et al.*, 2013; *Tilmes et al.*, 2013; *Jones et al.*, 2013]). As part of their study on the hydrologic impact of geoengineering, *Tilmes et al.* [2013] drew attention to shifts in the upper percentiles of annual and seasonal mean precipitation in *G1* and *abrupt4* $\times$ CO_2 experiments relative to the preindustrial control state. Aside from this work, changes in climate extremes under geoengineering have not received special attention, even though climate extremes are more readily perceived by society and can have more immediate detrimental consequences than changes in mean quantities [*IPCC*, 2012].

In order to capture extreme temperature and precipitation events in the model simulations, we make use of a set of widely used indices for climate extremes as defined by the Expert Team of Climate Change Detection and Indices (ETCCDI) [*Peterson and Manton*, 2008; *Zhang et al.*, 2011]. These indices, computed from model output at daily resolution, describe various aspects of extreme climate events based upon either absolute or relative (percentile) thresholds (see http://www.climdex.org/indices.html). The ETCCDI indices have proven useful in the analysis of observations [*Donat et al.*, 2013], the evaluation of global climate models [*Sillmann et al.*, 2013b], and the projection of changes in climate extremes [e.g., *Tebaldi et al.*, 2006; *Orlowsky and Seneviratne*, 2012; *Seneviratne et al.*, 2012; *Sillmann et al.*, 2013a].

The GeoMIP experiment considered in this study, the models analyzed, and the indices of extreme temperature and precipitation events are described in section 2. The probability density functions of monthly mean temperature and precipitation are examined in section 3, as a means of gauging expected changes in the moments of their frequency distributions under both the *G1* and *abrupt4* $\times$ CO_2 experiments. In section 4, we shift the focus exclusively to extremes and examine both global mean and geographic differences between various ETCCDI indices in the three experiments. Certain prominent features seen in the difference maps of *G1* and *piControl* are studied in more detail using one of the models, which suggests that climate feedbacks lead to strong regional responses. In this section we also assess the efficacy of *G1* as a means of offsetting climate change due to a CO_2 quadrupling. Finally, in section 5, we discuss various implications of our results and conclude with an assessment of further research needs.

2. GeoMIP Experiment, Models, and Extreme Indices

2.1. The *G1* Experiment

We focus our analysis on the *G1* experiment as proposed within the GeoMIP framework [*Kravitz et al.*, 2011]. In *G1*, the global mean radiative forcing from a quadrupling of the atmospheric CO_2 concentration $F_{4\times CO2}$ is balanced by a specified decrease of the solar irradiance δS_0, with the latter estimated for each model individually as $\delta S_0 \simeq -4F_{4\times CO2}/(1-\alpha)$, where α is the planetary albedo [*Schmidt et al.*, 2012]. The experiment

Table 1. GeoMIP Models Analyzed in This Study[a]

	Model	Institution	Resolution Longitude × Latitude Levels	Ensemble Members/Years		
				piControl	*G1*	*abrupt4* × CO_2
1	BNU-ESM	Beijing Normal University, China	2.8° × 2.8°(T42) L26	1/523	1/51	1/100
2	CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	2.8° × 2.8°(T63) L35	1/796	3/100	1/100
3	CCSM4	National Center for Atmospheric Research, USA	1.25° × 0.94° L26	1/156	1/75	1/75
4	EC-EARTH	Danish Meteorological Institute	1.125° × 1.125°(T159) L62	1/100	1/75	1/100
5	GISS-E2-R	Goddard Institute for Space Studies, USA	2.5° × 2.0° L40	3/70	3/70	3/70
6	HadGEM2-ES	Met Office Hadley Centre, UK	1.88° × 1.24° L40	1/238	1/70	1/151
7	IPSL-CM5A-LR	Institut Pierre-Simon Laplace, France	3.75° × 1.88° L39	1/550	1/50	1/100
8	MPI-ESM-LR	Max Planck Institute for Meteorology, Germany	1.88° × 1.88° (T63) L47	1/500	1/70	1/100
9	NorESM1-M	Norwegian Climate Centre, Norway	2.5° × 1.88° L26	1/500	1/50	1/100

[a]The rightmost column gives information on members and simulation lengths used in this study; a larger body of results may now be available for some models. For additional model details, see *Kravitz et al.* [2013].

is initiated from each model's preindustrial control run (hereafter *piControl*), and the forcings are applied instantaneously and concurrently beginning at model year 0. The degree to which the forcings are balanced is ascertained by evaluating the top of atmosphere net radiation change relative to *piControl*, with the goal being a difference of less than 0.1 Wm^2 between *G1* and *piControl* [*Kravitz et al.*, 2011]. The duration of *G1* varies from 50 to 100 years depending on the model (Table 1). In three of the models, the solar irradiance was restored to its control value after 50 years, and the experiment continued under the *abrupt4* × CO_2 forcing alone for 20–50 years. We restrict attention in this paper to the period of balanced forcings, leaving an analysis of the termination of geoengineering for a different study (however, see *Jones et al.* [2013] for such an examination of the GeoMIP *G2* experiment). We confine attention to the *G1* experiment for three reasons: (1) it was performed by the largest number of model groups; (2) its simplified design with balanced but large forcings results in a high signal-to-noise ratio in the regional climate response; and (3) the applied solar irradiance reduction serves as a proxy for other geoengineering methods considered by some to be more realistic—e.g., stratospheric aerosol injection—without the need to consider the details of how such methods might be simulated in each model.

2.2. The GeoMIP Multimodel Ensemble

Full details of the 12 global climate and Earth system models participating in GeoMIP may be found in *Kravitz et al.* [2013, Table 1]. We analyzed output from nine of these that have output at daily frequency (Table 1). All of the models used in our analysis, except EC-EARTH, are Earth system models, meaning that they include coupled terrestrial and oceanic carbon cycles, although the atmospheric CO_2 concentration is held fixed.

The variation in control run durations made available via GeoMIP and CMIP5 is very large (Table 1), ranging from 156 (CCSM4) to nearly 800 years (CanESM2). In the analysis, we make use of all available years for each model. This is preferable to selecting a shorter period of common length for two reasons. First, information on the degree to which different models' control states are equilibrated is not generally available, and choosing a specific period for a particular control run might lead to unexpected biases. And second, the calculation of many extreme indices relies upon threshold values determined from the distribution of values in *piControl*, and since models with longer control runs have more robustly estimated thresholds, this argues for the more inclusive approach. Results presented in the supporting information demonstrate that different models' *piControl* temperature and precipitation frequency distributions are remarkably insensitive to these differences in sample size (Figures S1 and S2 in the supporting information).

All *G1* simulations were analyzed over a common period from years 10 to 49, omitting the first decade in case of any possible initial adjustment to the forcings. However, few models show any signs of an adjustment in either global mean surface temperature or precipitation, as noted by *Schmidt et al.* [2012]. The *abrupt4* × CO_2 simulations range from 70 to 150 years in duration, starting from the same preindustrial control simulation as *G1*. For these, a common analysis period from years 30 to 69 was adopted, in order to include the maximum number of models. As this experiment begins with a sudden, large adjustment to the 4 × CO_2 forcing at year 0, followed by a slow, asymptotic approach to an equilibrium state which can only be achieved in a much longer simulation, we chose a common 40 year period as far from year 0 as possible,

Table 2. Subset of the Indices for Climate Extremes as Recommended by the Expert Team of Climate Change Detection and Indices (ETCCDI)

Index	Description	Index Definition	Units
TNn	Coldest daily T_{min}	Monthly or annual minimum value of daily minimum temperature *TN*	°C
TXx	Warmest daily T_{max}	Monthly or annual maximum value of daily maximum temperature *TX*	°C
CSDI	Cold spell duration	Number of consecutive days (> 6 days) when daily minimum temperature falls below the 10th percentile of *piControl*	days
WSDI	Warm spell duration	Number of consecutive days (> 6 days) when daily maximum temperature falls above the 90th percentile of *piControl*	days
DTR	Diurnal temperature range	Normalized sum over all days of the difference between daily maximum and minimum temperature	°C
GSL	Growing season length	Number of days between the first occurrence (> 6 days) of daily mean temperature $\overline{T} \equiv (TN + TX)/2 > 5$°C and the first occurrence (> 6 days) after 1 July (NH) or 1 January (SH) of $\overline{T} < 5$°C	days
FD	Frost days	Number of days when $TN < 0$°C	days
SU	Summer days	Number of days when $TX > 25$°C	days
Rx5day	Wettest consecutive 5 days	Maximum of consecutive 5 day (cumulative) precipitation amount	mm
CDD	Consecutive dry days	Maximum number of consecutive days when precipitation < 1 mm	days

while still including all models. The same methodology was adopted by *Schmidt et al.* [2012], who chose an even later analysis period for *abrupt4* × CO_2.

2.3. Climate Extreme Indices

The analysis of section 4 is focused on six ETCCDI indices characterizing different aspects of temperature and precipitation extremes, as summarized in Table 2. Daily maximum and minimum temperatures, *TX* and *TN*, respectively, are available directly from the model output. Absolute temperature extremes for a particular month or year are denoted by *TXx*, the maximum daily maximum, and *TNn*, the minimum daily minimum 2 m air temperatures. Although *TXx* and *TNn* are single values from the extreme tails of the temperature distribution, examination of model results reveals that they are generally not isolated. For example, in the CanESM2 *piControl* experiment, for a sample of several well-separated years over the course of the run, the next nearest values of *TX* and *TN* are within 2°C of *TXx* and *TNn* at 97 and 92% of locations on the model grid, respectively (for reference, the standard deviations of *TX* and *TN* are both near 6.7°C). The duration indices *CSDI* and *WSDI* are based on threshold estimates from *piControl* and indicate the intraannual length, in days, of a cold spell or warm spell, respectively. The precipitation index *Rx5day* provides information about the maximum 5 day precipitation amount over a month or year and can be taken as a rough indicator of changes in flooding conditions [*Frich et al.*, 2002]. The duration index *CDD* describes the maximum number of consecutive dry days in a calendar year. Results for several additional indices appearing in Table 2 are depicted graphically in the supporting information.

For the analyses presented here, all indices were regridded to a median model grid resolution of 144 × 96 (2.5° longitude × 1.9° latitude), which corresponds to the grid of the NorESM1-M model. For noninteger variables (e.g., *TXx*, *TNn*, and *Rx5day*), a first-order conservative remapping algorithm was used [*Jones*, 1999], while for integer variables (e.g., *CSDI*, *WSDI*, and *CDD*), a nearest-neighbor interpolation method was applied. Both interpolations were carried out using the Climate Data Operators package (http://code.zmaw.de/projects/cdo).

3. Probability Distributions of Monthly Temperature and Precipitation

In order to gain a general impression of the changes in extremes in the two perturbed experiments compared to *piControl*, we begin by computing probability density functions (PDFs) of temperature and precipitation for all models. At every grid point in each model, we first calculated the standardized monthly anomalies of monthly mean surface temperature in *piControl*, i.e., $\mathcal{T}_m^{\mathrm{Ctl}} = (T_m^{\mathrm{Ctl}} - \overline{T}_m^{\mathrm{Ctl}})/\sigma_{T_m}^{\mathrm{Ctl}}$, where an overline denotes the multiyear monthly mean and $\sigma_{T_m}^{\mathrm{Ctl}}$ is the interannual monthly standard deviation for month *m*. Next, we computed the monthly anomalies in *G1* and *abrupt4* × CO_2 relative to the control mean and standard deviation, i.e., $\mathcal{T}_m^{\mathrm{G1}} = (T_m^{\mathrm{G1}} - \overline{T}_m^{\mathrm{Ctl}})/\sigma_{T_m}^{\mathrm{Ctl}}$. This procedure aids the visual comparison of PDFs between the three experiments. The same algorithm was used to generate PDFs of precipitation. The results for individual models are provided in the supporting information (Figures S1 and S2), while the multimodel mean PDFs, using equal weights for each model, are displayed in Figure 1. The equal weighting was used

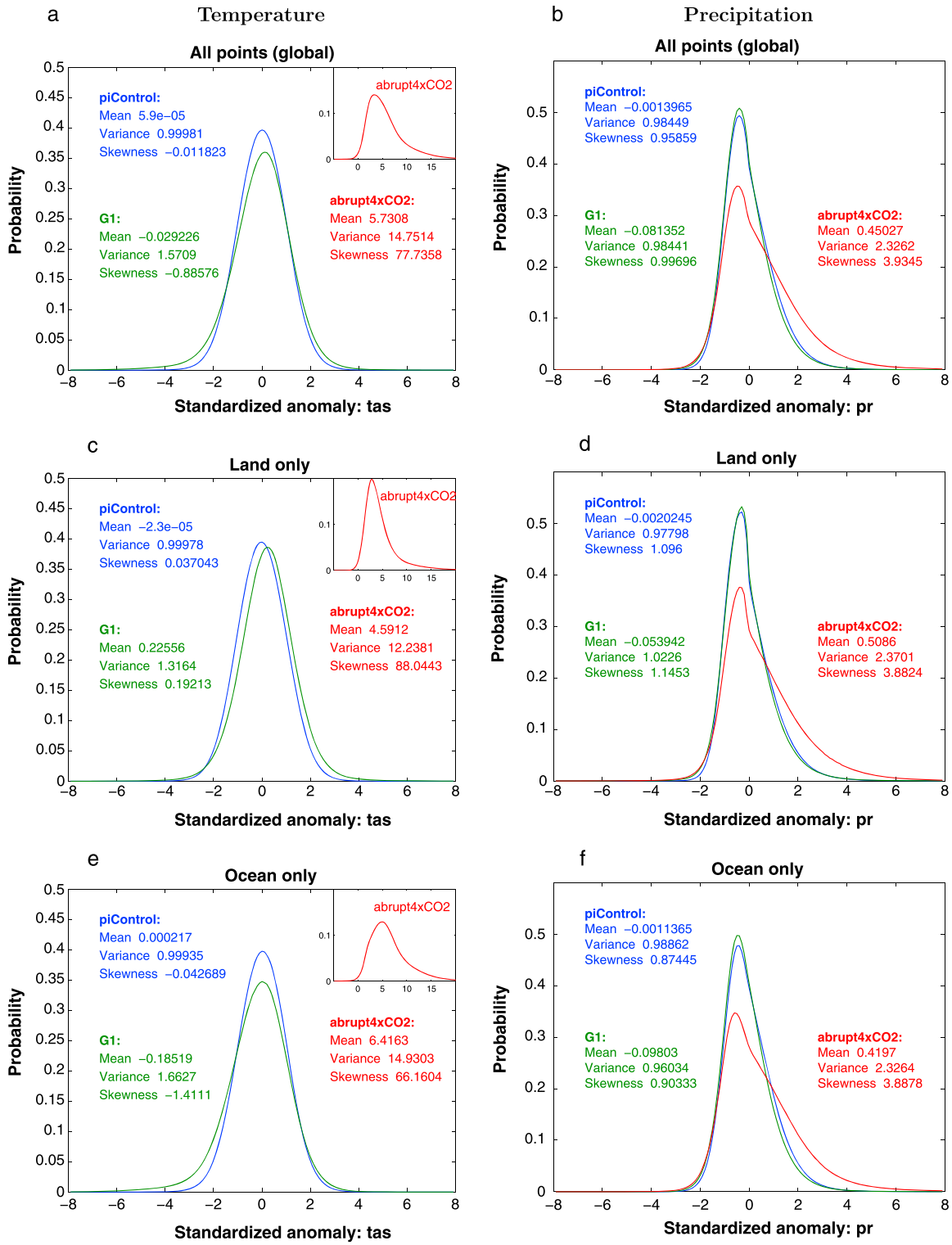

Figure 1. Normalized probability distributions of standardized monthly mean anomalies for the multimodel average and for three experiments: *piControl* (blue), *G1* (green), and *abrupt4* × CO_2 (red). Anomalies are calculated with respect to the mean and standard deviation of *piControl* for each experiment; see text for details. The PDFs of surface air temperature are shown in the left-hand panels, precipitation in the right-hand panels. Upper panels show global results, middle panels land-only, and lower panels ocean-only.

so that higher resolution models would not dominate over those with a coarser grid, while ensemble members for models with multiple realizations of *G1* or *piControl* (i.e., CanESM2 and GISS-E2-R) were averaged together before constructing the multimodel mean. The PDFs for the perturbed experiments (i.e., *G1* and *abrupt4*$\times CO_2$) were then subjected to two statistical tests in order to determine if they differ significantly from the control experiment for each model. The Wilcoxon rank sum test was used to gauge whether the median values differ, while the Ansari-Bradley test was applied to test for unequal variances [*Hollander and Wolfe*, 1973]. Both tests are nonparametric and thus do not assume a normal distribution for either sample. This is necessary since the PDFs of precipitation anomalies are non-Gaussian and positively skewed, as clearly seen in Figure 1. The results of the two tests for the *G1* experiment are tabulated for all models in Table S1 of the supporting information. In the case of *abrupt4*$\times CO_2$, the PDFs differ significantly from *piControl* in all cases, as could be anticipated from a visual inspection of Figure 1 and Figures S1 and S2 of the supporting information.

In Figure 1a, the PDFs of temperature in the global case (i.e., including all points on each model grid) display modest but discernible differences between *G1* and *piControl*, particularly for negative anomalies. The low-$\mathcal{T}$ tail in *G1* (green curve in Figure 1a) extends to lower temperatures than in the control; for example, monthly mean temperatures that are 3 standard deviations below the mean (i.e., -3σ) are more than twice as likely in *G1* than in *piControl*. By contrast, differences in the high-$\mathcal{T}$ tail in the two experiments are small. Figures 1c and 1e reveal that the latter results from a near cancellation of anomalously high $\mathcal{T}$ over land and low $\mathcal{T}$ over ocean. Stratifying the PDFs by latitude band (supporting information, Figure S3) shows that it is the low-$\mathcal{T}$ tail over the tropical oceans, in particular, that is responsible for the corresponding feature in the global PDF for the *G1* experiment (Figure 1a). These properties are consistent with the response of the mean annual and monthly temperature in *G1* as discussed by several previous authors [*Lunt et al.*, 2008; *Schmidt et al.*, 2012; *Kravitz et al.*, 2013], which is in turn a direct result of the latitude dependence of the applied forcings: i.e., negative in the tropics and positive elsewhere [*Kravitz et al.*, 2013]. Differences in the first three moments of the distribution are clearly discernable between the two experiments, *piControl* and *G1*, as also indicated in Figure 1. Over land, the mean temperature increases, the variance decreases, and skewness becomes more positive. The behavior of each of these quantities is opposite over the ocean. The results in Table S1 show that differences in the median and variance of $\mathcal{T}$ are statistically significant in all models, with a high level of confidence afforded by the large number of data points available for each model. Finally, the insets in Figures 1a, 1c, and 1e show the dramatic rightward shift of the temperature PDF in the *abrupt4*$\times CO_2$ experiment, indicating warming at nearly all locations in the models.

Compared to temperature, the PDFs of monthly precipitation anomaly $\mathcal{P}$, shown in Figures 1b, 1d, and 1f display smaller differences between the three experiments. The PDFs are positively skewed in all cases, a generic property of precipitation and other positive definite climate variables (e.g., wind speed). The largest difference between *piControl* and *G1* occurs over ocean, where both the mean and the high-$\mathcal{P}$ tail are shifted to lower values in *G1* (Figure 1f). Again, this is consistent with the known reduction of annual mean precipitation in the tropics in GCMs under similar forcings [*Bala et al.*, 2008; *Kravitz et al.*, 2013; *Tilmes et al.*, 2013]. The differences in the PDF over land are qualitatively similar (Figure 1d) but of smaller magnitude. In *abrupt4*$\times CO_2$, both the mean and variance of $\mathcal{P}$ increase strongly, with the result that both high- and low-precipitation extremes are more prevalent. While a monthly precipitation anomaly of magnitude $+3\sigma$ is only 75% as likely in *G1* compared to *piControl*, it is about 4 times more likely to occur in *abrupt4*$\times CO_2$. Analysis of the PDFs by latitude band shows that the shift of the low-$\mathcal{P}$ tail to smaller values in *abrupt4*$\times CO_2$ occurs mainly in the tropics (supporting information Figure S4). As might be suspected from a visual examination of Figure 1, the median and variance of both $\mathcal{T}$ and $\mathcal{P}$ in *abrupt4*$\times CO_2$ are significantly different from *piControl* in all models.

It is clear from an examination of all of the PDFs that the mean, variance, and skewness of temperature and precipitation perturbations that occur under *abrupt4*$\times CO_2$ are all reduced to near-*piControl* values by *G1*-type geoengineering. However, Figures 1 and S1–S4 also show that this return of climatic conditions to near-preindustrial levels is not achieved uniformly over the globe. Distinct differences in the PDFs are seen between land and ocean and at different latitudes, particularly as reflected in the tails (i.e., extremes) of $\mathcal{T}$ and $\mathcal{P}$. This motivates further investigation of these aspects in the sections which follow.

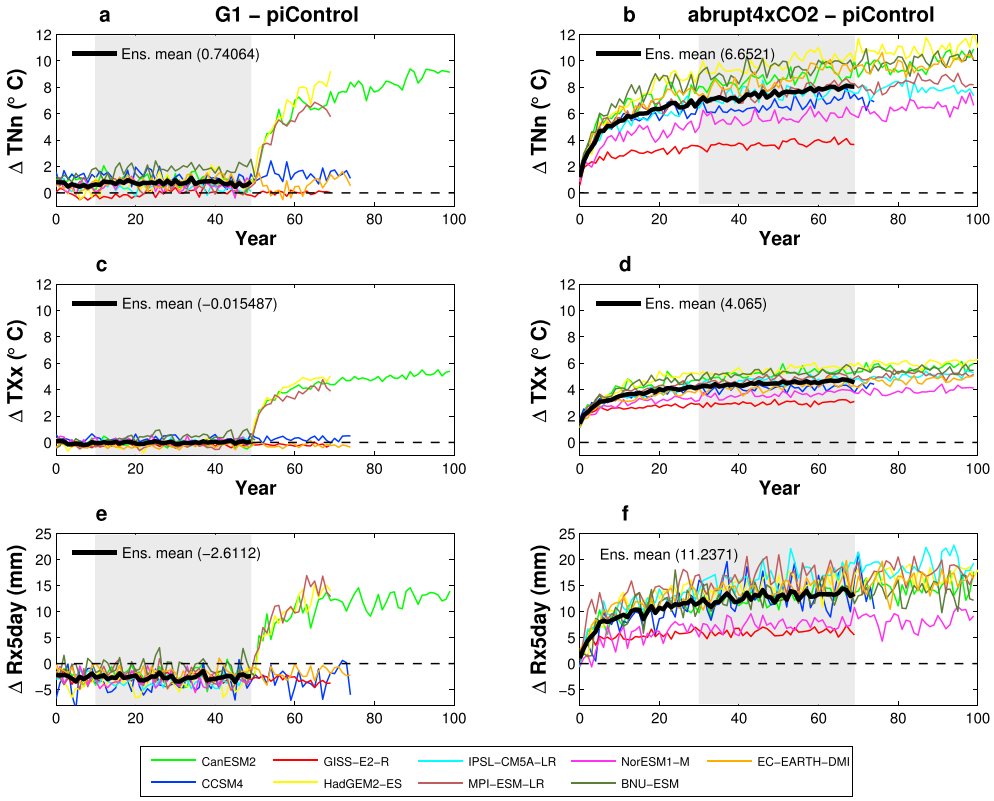

Figure 2. Time series of the difference of global mean extreme indices from *piControl* values for all models analyzed. Left-hand panels show the difference *G1*−*piControl*, right-hand panels *abrupt4* × CO_2−*piControl*, and the indices are, from top to bottom, *TNn, TXx*, and *Rx5day*. The heavy dark curve is the multimodel mean, and gray shading indicates the 40 year analysis period for each experiment used in this study. The values at the top left in each panel are the multimodel and time mean of the indicated index over the analysis period.

4. Daily Extremes Under the *G1* and *abrupt4* × CO_2 Experiments

In this section, we present results for the six ETCCDI indices describing temperature and precipitation extremes introduced in section 2.3. Global and multimodel values are reported as *mean* (*min* to *max*), where *mean* represents the all-model ensemble average (including all ensemble members of a given model), and *min*, *max* are results from the multimodel ensemble member that simulates the minimum and maximum, respectively, of the indicated quantity. The multimodel standard deviation, σ, is also reported.

4.1. Global Mean Time Series

Global mean time series of the extreme indices *TXx*, *TNn*, and *Rx5day*, expressed as differences from their control values $\langle \Delta \ldots \rangle$ (angle brackets denote global means), are shown in Figure 2. The behavior of the indices is qualitatively similar to corresponding time series of global and annual average surface air temperature $\langle \Delta T \rangle$ and precipitation $\langle \Delta P \rangle$ [*Kravitz et al.*, 2013] (supporting information Figure S1). In *G1*, $\langle T \rangle$ remains near the corresponding control value during the period of balanced forcings (years 0–49), with $\langle \overline{T} \rangle < 0.2$ K in most models. The extreme temperature index *TXx* in Figure 2b behaves similarly, with $\langle \Delta TXx \rangle = -0.021$ K (−0.17 to +0.23 K, $\sigma = 0.19$ K) over the 40 year analysis period of *G1* (shaded region in Figure 2a). By contrast, $\langle \Delta TNn \rangle$ is significantly positive in most models (Figure 2a), with a multimodel mean value of 0.61 K (−0.06 to +1.1 K, $\sigma = 0.48$ K). This is the result of systematically warmer nighttime low temperatures in *G1* compared to *piControl*, mainly over land, as demonstrated in section 4.2. This behavior might be anticipated under this form of SRM, since the reduced shortwave surface heating impacts daytime temperatures directly but not nighttime temperatures. Surface longwave forcing from increased CO_2, however, is not tied to the diurnal cycle.

In the ensemble of *G1* experiments analyzed by *Schmidt et al.* [2012] and *Kravitz et al.* [2013], $\langle \overline{P} \rangle$ is less than the corresponding control value in all models—by an average of −4.5% in the 12 models analyzed by

Kravitz et al. [2013]—indicating an overall weakening of the hydrological cycle (discussed in more detail by *Tilmes et al.* [2013]). This feature is similarly reflected in the time series of $\langle \Delta Rx5day \rangle$ in Figure 2e, which has a multimodel mean time-average value of −3.0 mm (−3.62 to −2.19 mm, $\sigma = 0.92$ mm). Compared to the *piControl* value of $\langle Rx5day \rangle = 84.3$ mm, this represents a relative difference of −3.6 (−4.3 to −2.6)%, somewhat smaller than the relative change in mean *P* cited above. The time series in Figures 2a–2c also clearly show the sudden transition from *G1* to a postgeoengineering state as SRM is removed for the three models that included this modified forcing after 50 years of geoengineering (the forcing remained unchanged in the other two models that ran beyond year 50). While the analysis of this period lies outside the scope of this paper, the termination of geoengineering in the related *G2* experiment is the topic of a recent paper by *Jones et al.* [2013], while *Alterskjaer et al.* [2013] considered its impact in the context of artificial marine cloud brightening.

Figures 2b, 2d, and 2f show corresponding time series of the same indices in the *abrupt4* × CO_2 experiment, with the analysis period again indicated by the shaded portion. It is important to note, for the comparisons between experiments which follow, that the *abrupt4* × CO_2 experiment is too short in duration to permit equilibration between atmosphere and ocean; hence, all models are in a highly transient state during the analysis period, unlike the *G1* experiment. There is a wider spread in model response to the much stronger forcing in *abrupt4* × CO_2 than in *G1*, likely a reflection of the differing magnitude of climate system feedbacks in each model. This is particularly evident for $\langle \Delta TNn \rangle$ and $\langle \Delta Rx5day \rangle$, both of which exhibit a factor of ∼4 spread amongst the different models at year 69. NorESM1-M and GISS-E2-R, both of which exhibit lower than average $\langle \Delta TNn \rangle$, also simulate significantly lower $\langle \Delta Rx5day \rangle$ than other models, as also found by *Tilmes et al.* [2013] for annual and global mean temperature and precipitation changes. This characteristic is consistent with our understanding of the relation between global mean temperature and (mean and extreme) precipitation change in increased CO_2 experiments [*Held and Soden*, 2006; *Kharin et al.*, 2007]. Large differences in the interannual variability of the models are also evident, especially for *Rx5day* (Figure 2f). The GISS-E2-R model consistently displays the lowest interannual variability amongst all the models, possibly since it is one of only two models for which three separate ensemble members were averaged, and also has the shortest *piControl* period (70 years for each ensemble member) of all the models (meaning that the control climate variability may be underestimated). Multimodel mean values of $\langle \Delta TNn \rangle$ are consistently a factor of 1.5–2 larger than those of $\langle \Delta TXx \rangle$ (heavy black curves in Figures 2d and 2e), indicating a much stronger response of nighttime low temperatures than daytime high temperatures to CO_2 quadrupling. This asymmetry between the *TXx* and *TNn* responses in both *G1* and *abrupt4* × CO_2 is discussed further in the following sections.

4.2. Spatiotemporal Structure of the Indices

4.2.1. Geographical Features

As emphasized in many recent analyses of similar experiments, the attainment and maintenance of a climate state in which global mean forcings are balanced, such as *G1*, does not particularly constrain the regional climate characteristics of that state [*Lunt et al.*, 2008; *Irvine et al.*, 2010; *Ricke et al.*, 2010; *Ban-Weiss and Caldeira*, 2010; *MacMartin et al.*, 2013]. Thus, it is instructive to examine the geographical patterns of differences (anomalies) between the three experiments: i.e., *G1* − *piControl* and *abrupt4* × *CO2* − *piControl*. Latitude-longitude variations in *TXx*, *TNn*, and *Rx5day* are shown in Figure 3 for *G1* − *piControl* (Figures 3a, 3c, and 3e) and *abrupt4* × CO_2 − *piControl* (Figures 3b, 3d, and 3f). In these maps and those that follow, regions where the change in a variable compared to its variance over the multimodel ensemble are small, as determined by Student's *t* test with a 5% *p* value, are deemed nonsignificant and are stippled. For models with more than one ensemble member (i.e., CanESM2 and GISS-E2-R *G1* experiments), the ensemble members are averaged before applying the *t* test to the multimodel ensemble.

Broadly speaking, the latitudinal distribution of the difference in temperature extremes for *G1* − *piControl* resembles that found by *Kravitz et al.* [2013] for annual mean surface temperature, with negative anomalies in the tropics and positive anomalies at higher latitude [see *Kravitz et al.*, 2013, Figure 2]. In the extratropics and everywhere over land, the warming anomaly seen in *TNn* (Figure 3a) is essentially a weaker version of that seen in the *abrupt4* × CO_2 experiment (Figure 3b), with the signature of polar amplification evident in both hemispheres. By contrast, Figure 3c shows a distinct north-south asymmetry in Δ*TXx* over land in response to the *G1* forcing. The magnitude of the extreme temperature anomalies in *G1*, however, is far smaller than in *abrupt4* × CO_2. The multimodel annual mean of the ratio of change in the two experiments, i.e., (*G1* − *piControl*)/(*G1* − *abrupt4* × CO_2), ranges from −0.2 in the tropics to +0.2 at 55°N for Δ*TXx* and from

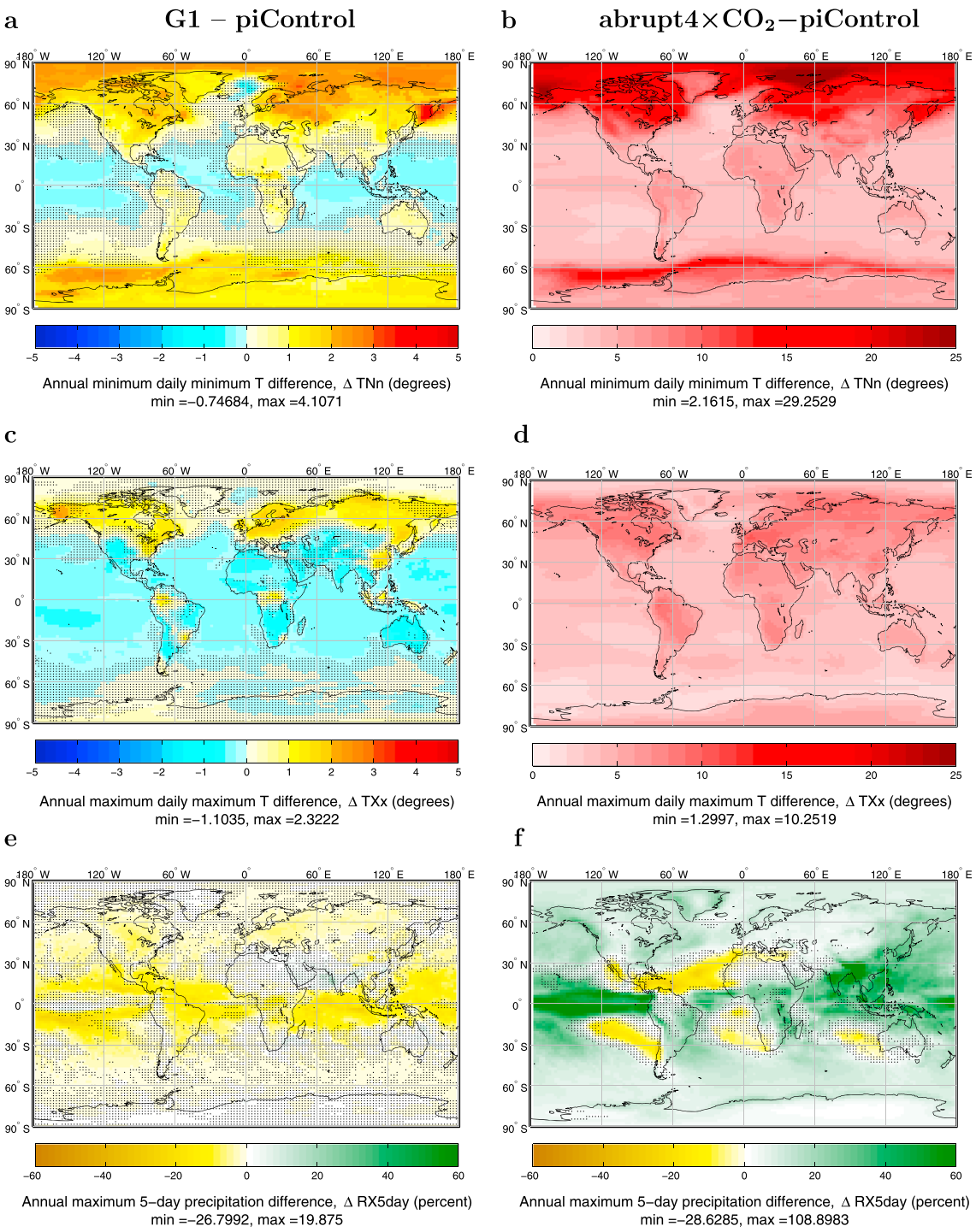

Figure 3. Geographical distributions of differences from *piControl* (multimodel means) for the extreme indices (a, b) *TNn*, (c, d) *TXx*, and (e, f) *Rx5day*, taken over the fiducial 40 year analysis periods. Results are shown for (a, c, e) *G1 − piControl* and (b, d, f) *abrupt4 × CO_2 − piControl*. The stippling pattern indicates regions where differences are *not* statistically significant at the 95% level based on the intermodel variance, as determined by Student's *t* test.

approximately 0 in the tropics to +0.3 at the South Pole for Δ*TNn* (see supporting information, Figures S6 and S7). As the magnitude of Δ*TNn* is larger than that of Δ*TXx* in regions where both are positive, it is evident that changes in *TN*, rather than *TX*, are the dominant influence on annual mean Δ*T* in the extratropics under *G1*. As Figures 3b and 3d show, this is also the case in the *abrupt4 × CO_2* experiment. Similar results

were found by *Sillmann et al.* [2013a] in their analysis of temperature extremes in CMIP5 models under the RCP8.5 increased greenhouse gas scenario, the closest analog to *abrupt4* × CO_2 [*Good et al.*, 2013].

At latitudes between 45°S and 45°N and over ocean, widespread decreases in both *TXx* and *TNn* are seen in *G1* (Figures 3a and 3c). Temperature maxima are much more strongly affected than minima, however, with multimodel mean$(\Delta TXx)_{ocean} = -0.373$ K compared with mean$(\Delta TNn)_{ocean} = +6.7 \times 10^{-4}$ K (excluding grid cells where changes are not statistically significant at the 95% level). Over land, the responses of ΔTXx and ΔTNn are largely opposite in sign over this latitude range, with $\Delta TXx < 0$ in most areas, and $\Delta TNn > 0$ but generally of smaller magnitude (<1°C in most locations where differences are significant). Exceptions to this behavior are found in equatorial South America, the eastern U.S., southern Europe, and eastern China.

These changes in ΔTNn and ΔTXx under *G1* affect the magnitude of interannual temperature variability (or "extreme temperature range," $ETR \equiv TXx - TNn$ [*Frich et al.*, 2002; *Alexander et al.*, 2006]) and also the diurnal temperature range (*DTR*) (supporting information Figure S9). *ETR* decreases noticeably over most of the globe in *G1*, consistent with the larger and more widespread increases in *TNn* versus *TXx* mentioned above. Regions of increased *ETR* correspond to "hot spots" in *TXx* where land-atmosphere feedbacks are important (see section 4.2.2). *DTR* decreases significantly over the global ocean (generally $\lesssim 0.5$ K), subtropical Africa, the Middle East, and Central Asia (by ~0.5 to 0.8 K). Statistically significant increases in *DTR* are observed over equatorial land regions in *G1*, of order 1.5 K over much of the Amazon and Congo River basins, despite the rather modest changes in *TNn* (Figures 3a and 5a) and *TXx* (Figures 3c and 6a) seen over equatorial land regions.

The pattern of annual maximum 5 day precipitation difference, $\Delta Rx5day$ (Figure 3e) reveals fewer regions of statistically significant difference between *G1* and *piControl*, but there are broad coherent patterns of reduced precipitation in the tropics, mainly over oceans, and at high northern latitudes, mainly over land. Moderate reductions in *Rx5day* ($\lesssim$ 15%) exist over large areas: Canada, Europe, Russia, China, and eastern South America. The largest decreases (reaching −28%) are seen over the equatorial oceans. Increases in $\Delta Rx5day$, such as over Egypt and Saudi Arabia, are few and isolated and, in most cases, not statistically significant. The pattern of $\Delta Rx5day$ resembles that of annual mean precipitation change presented in *Kravitz et al.* [2013, Figure 5] in the tropics and subtropics but not at high latitudes, where decreases in *Rx5day* are more evident. The magnitude of the annual $\Delta Rx5day$ is typically a factor of 2–4 larger than that of the annual ΔP at low latitudes. Finally, the increase in annual mean precipitation seen in the eastern equatorial Pacific in *Kravitz et al.* [2013, Figure 5] is not replicated in $\Delta Rx5day$. The changes in *Rx5day* under *G1* are notably different from those under *abrupt4* × CO_2: Figure 3f exhibits increased *Rx5day* nearly everywhere on the planet in *abrupt4* × CO_2, with the exception of northwest Africa, the Mediterranean Sea, and areas of the subtropical oceans.

While *TXx*, *TNn*, and *Rx5day* describe the absolute magnitude of certain extreme climate conditions, the duration of such conditions is also of interest. Three measures of persistence are shown in Figure 4: cold spell duration (*CSDI*), warm spell duration (*WSDI*), and consecutive dry days (*CDD*) (see Table 2 for precise definitions). Figures 4a–4d provide valuable complementary information to the distributions of ΔTXx and ΔTNn shown in Figures 3a–3d. While the total area of statistically significant change in *CSDI* (Figure 4a) is smaller than for *TNn* (Figure 3a), their geographic patterns closely resemble each other. The same can be said for *TXx* and *WSDI* (Figures 3c and 4c). This is to be expected from the respective definitions of *CSDI* and *WSDI* in terms of the upper and lower deciles of the daily temperature distribution (Table 2).

Outside the tropics, *CSDI* decreases in concert with increasing *TNn*, but the changes in the former are robust only north of ~40°N (mainly the northeastern U.S., Greenland, eastern Europe, and west central and northeastern Russia) and south of ~55°S (Figure 4a). The most striking feature of Figure 4a is the significant increase in *CSDI* over the tropical and subtropical oceans, with $\Delta CSDI$ exceeding 50 days per year over large regions, an increase of 140% over the preindustrial control value (36 days, by the definition of *CSDI*; Table 2). These *CSDI* increases might not be anticipated from the rather moderate reductions in annual extreme temperatures over the low-latitude oceans seen in Figures 3a and 3c. This persistent overcooling in *G1* could have consequences for tropical ocean circulation, cyclogenesis, and/or biological productivity (section 5). Most of the *CSDI* differences over land in *G1* are not robust, with the notable exception of India, which experiences an increase in cold spell duration of 10–20 days (28–56%; Figure 4a).

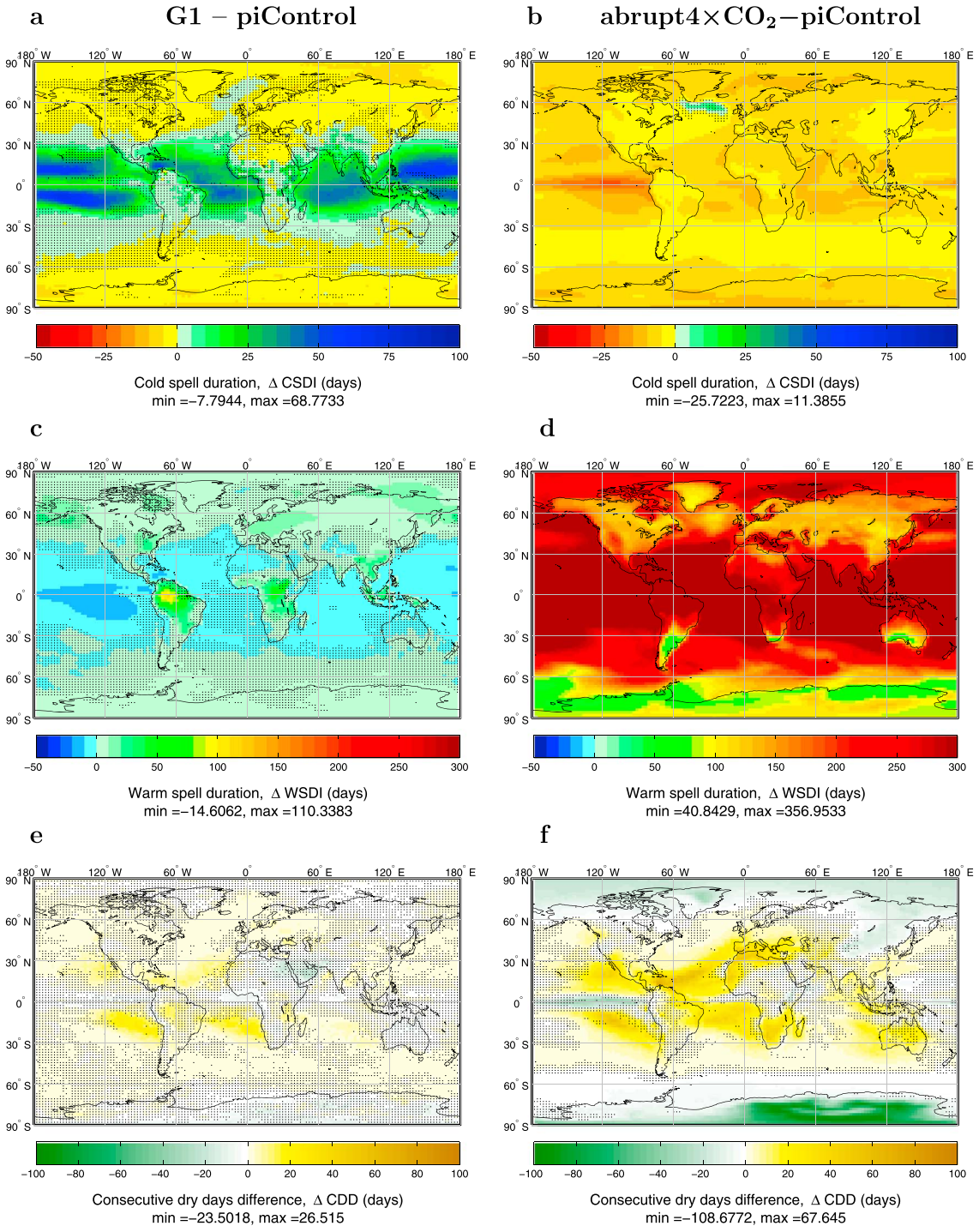

Figure 4. Geographical distributions of differences from *piControl* (multimodel means) for the extreme indices (a, b) *CSDI*, (c, d) *WSDI*, and (e, f) *CDD*, taken over the fiducial 40 year analysis periods. Results are shown for *G1* − *piControl* (Figures 4a, 4c, and 4e) and *abrupt4* × CO_2 − *piControl* (Figures 4b, 4d, and 4f). The stippling pattern indicates regions where differences are not statistically significant at the 95% level, as determined by Student's *t* test.

As expected, *CSDI* decreases worldwide in the *abrupt4* × CO_2 experiment (Figure 4b), due to the strong rightward shift of the PDF of surface temperature (Figure 1a, inset). At locations where *CSDI* decreases in both experiments, the *CSDI* reduction in *abrupt4* × CO_2 is always larger than in *G1* (where the changes are statistically significant). For instance, the *CSDI* decreases in the Nile Delta; eastern Europe and western/central Russia are up to 4 times larger in *abrupt4* × CO_2 than in *G1*.

Turning to the geographic pattern of Δ*WSDI*, Figure 4c shows that significant increases in warm spell duration, by (~50–100 days or a factor of 1.4–2.8 over *piControl* values), occur even in areas where neither Δ*TXx* nor Δ*TNn* is significant in *G1*, e.g., equatorial South America, Africa, and Southeast Asia (Figures 3a and 3c). These increases over tropical land areas are also notable for their large areal coverage. Even more widespread *WSDI* increases are seen over land areas north of 45°N in *G1*, although they are of smaller magnitude (≲15 days or 42%). Comparison of Figures 4a and 4c shows that in *G1*, time-averaged minimum temperatures, as reflected by *CSDI*, decrease much more strongly over the tropical and subtropical oceans than do time-averaged maximum temperatures, as reflected by *WSDI*. On the other hand, Figures 4b and 4d reveal a notable asymmetry in the response of these two indices under the *abrupt4* × CO_2 forcing, with increases in *WSDI* being much larger (by a factor of 13) than decreases in *CSDI*. This is a direct result of the marked rightward shift of the temperature PDF in *abrupt4* × CO_2 and the definitions of the two indices. Decreases in *CSDI* are limited by the decreasing area to the left of the 10th percentile as the temperature PDF shifts rightward, while the corresponding area increase in the high-temperature tail is essentially unbounded.

The mostly positive changes in *CDD* (Figure 4e) are broadly consistent with the decreases in *Rx5day* noted above but are more robust over ocean than land, where most changes are not statistically significant. Significant lengthening of dry periods occurs over much of Europe into western Russia (by up to 3 days), southern Canada (up to 5 days), and Amazonia and Tanzania/Congo (up to 17 days). Decreases in *CDD* are far less common but are seen in Egypt and southern Saudi Arabia (up to 23 days). In *abrupt4* × CO_2, large areas of the subtropical oceans display increases in *CDD*, while strong decreases are evident at high latitudes, particularly in Antarctica (Figure 4f). These features are either much weaker or entirely absent in *G1*.

Geographic distributions showing experimental differences for a number of other ETCCDI indices which might be of interest in a variety of contexts (e.g., agricultural impacts) are provided in the supporting information (Figures S9 and S10).

4.2.2. Seasonality and Zonal Mean Changes Over Land

We now turn to an examination of the zonal structure and seasonality of changes in the climate extreme indices. In order to facilitate comparison between results for *G1* and *abrupt4* × CO_2, the same scales are employed on figure panels in this section. However, we provide alternate versions highlighting the detailed structure of the *G1* − *piControl* differences in the supporting information (Figures S5–S8), as certain of these are described below. In this section, we restrict attention to land areas only, where significant changes in extremes might be expected to have the greatest socioeconomic impact, e.g., in populated areas. The corresponding results for land and ocean combined may be found in the supporting information (Figures S6–S8).

Figure 5 (left) displays the zonal and annual mean behavior of anomalies of *TNn*. In the *G1* experiment (Figure 5a and supporting information Figure S5a), most models simulate a positive zonal and annual mean Δ*TNn* response, with a multimodel mean annual peak value of 1.95 K near 60°N, with the exception of GISS-E2-R, which simulates a near-zero response everywhere except near the South Pole. BNU-ESM features a larger Δ*TNn* than any other model north of at 30°N, at more than double the multimodel mean value at 60°N, and with a similar behavior in the Southern Hemisphere. Three models also simulate slightly negative Δ*TNn* between 20°S and 10°N. The response in *abrupt4* × CO_2 is of course much larger (peak value of +12.8 K at 65°N) and uniformly positive (Figure 5b), in agreement with previous studies of extremes under strong CO_2 forcing [*Orlowsky and Seneviratne*, 2012; *Sillmann et al.*, 2013a]. Again, however, the GISS model is notable for its subdued response in the Northern Hemisphere (less than half of the annual, multimodel mean Δ*TNn* at 65°N).

It is evident from Figure 5 (right) (see also supporting information Figure S5a) that Δ*TNn* exhibits significant seasonality at high latitudes: in the Arctic and Antarctic, positive *TNn* anomalies peak in the winter hemisphere in both *G1* and *abrupt4* × CO_2. On top of the annual cycle in the unperturbed temperature, this represents a reduction in the seasonality of *TN* at high northern latitudes. In the context of coupled global climate model experiments with forcing from increased greenhouse gases alone (generally milder analogs of *abrupt4* × CO_2), this response has been largely attributed to an effective heat capacity increase of the surface ocean due to sea ice loss [*Mann and Park*, 1996; *Dwyer et al.*, 2012; *Donohoe and Battisti*, 2013], which damps the strong seasonality of solar absorption. This could be a factor in the reduced seasonality

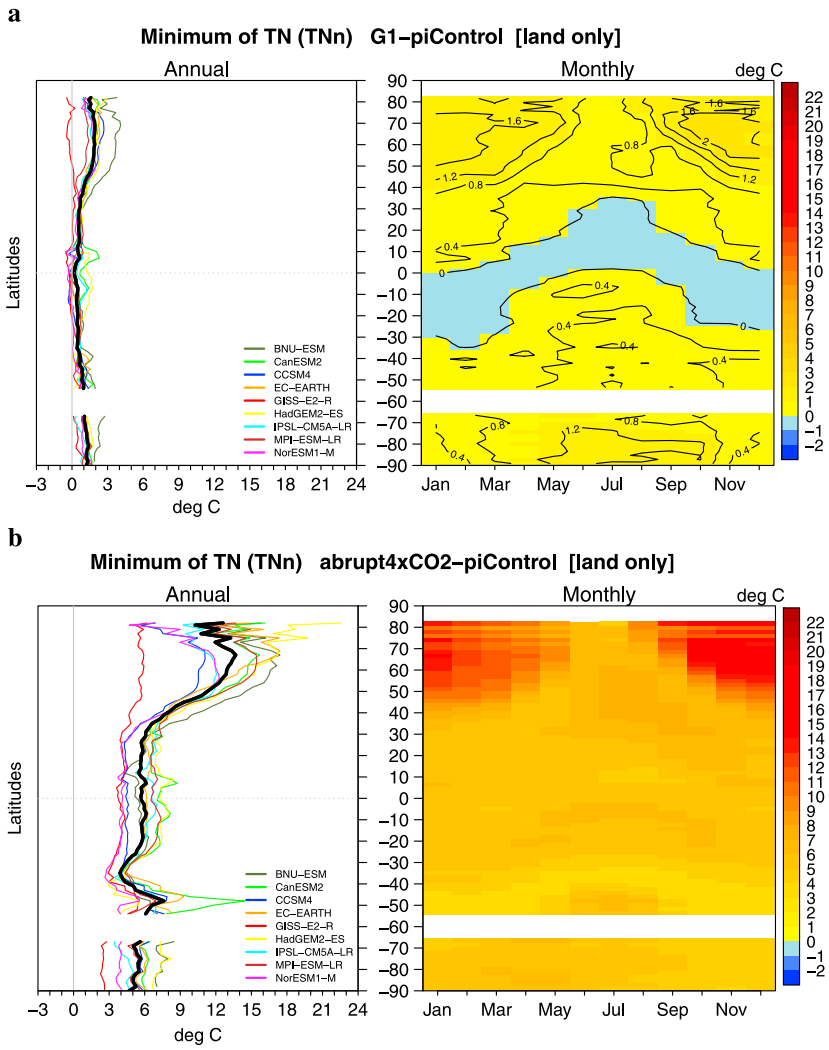

Figure 5. Multimodel, zonal mean (left column) annual and (right column) monthly anomalies in the extreme index *TNn* for the differenced experiments: (a) *G1 − piControl* and (b) *abrupt4 × CO_2 − piControl*, taken over the fiducial 40 year analysis periods. Units are degrees Celsius in all panels. The zonal averages are computed over land points only, which accounts for the gaps in the Arctic and Antarctic oceans. An alternate version of Figure 5a showing the more detailed structure of the *G1 − piControl* differences is provided in supporting information Figure S5a.

of *TN* in the Southern Hemisphere in *G1*, since some of the GeoMIP models do exhibit sea ice loss there [*Kravitz et al.*, 2013].

The reduced seasonality of *TN* in the Arctic, however, cannot be as readily attributed to the consequences of sea ice loss, since the latter is small in most models under *G1* [*Kravitz et al.*, 2013; *Moore et al.*, 2014]. In this case, enhanced wintertime heating might be provided by a combination of direct longwave heating of the surface by increased CO_2 and atmospheric water vapor (due to enhanced evaporation) and increased poleward energy transport from midlatitudes to high latitudes, both of which have been demonstrated to occur under enhanced greenhouse gas forcing [*Donohoe and Battisti*, 2013]. In a multimodel study of the Arctic response under *G1* specifically, *Moore et al.* [2014] found evidence of increased cyclonic activity entering the Barents Sea region in spring, leading to reduction of sea ice area in that season as warm air is advected into that region. Finally, significant seasonal changes in cloud cover, which are seen in certain of the GeoMIP models (see section 4.3 and Figure 9c), may also contribute to the *TN* response.

At lower latitudes under *G1*, $\Delta TNn < 0$ at the equator and in the subtropical summer hemisphere, where the reduced solar irradiance has the greatest impact. This temperature response pattern essentially mirrors the

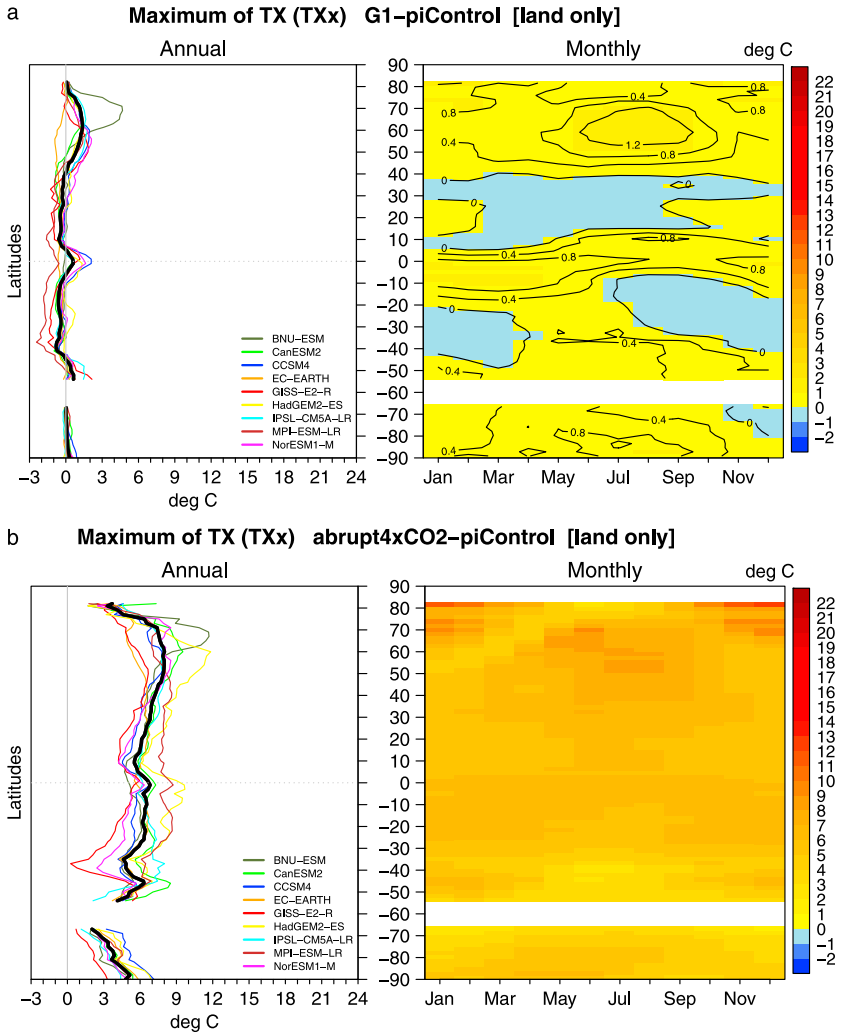

Figure 6. As in Figure 5 but for the extreme index *TXx*. An alternate version of Figure 6a showing the more detailed structure of the *G1 − piControl* differences is provided in supporting information Figure S5b.

spatial and temporal dependence of the forcing, as seen in other SRM experiments of this type [e.g., *Govindasamy and Caldeira*, 2000, Figure 1], and does not depart in any significant way from the corresponding plot of monthly, zonal mean temperature in *G1* (see Figure 8a). Clearly, for *TNn*, SRM as implemented in *G1* is less effective at compensating for the *abrupt4* × CO_2 forcing as one approaches high latitudes: Δ*TNn* reaches a maximum of +2.36 K in November at 71°N.

Turning now to differences in *TXx*, it is evident from Figure 6 that the multimodel mean Δ*TXx* in both *G1 − piControl* and *abrupt4* × CO_2 *− piControl* is of smaller magnitude than Δ*TNn* (Figure 5). Consistent with Figure 3c, *TXx* for *G1 − piControl* exhibits a more complex difference pattern with latitude and season than *TNn*. The multimodel mean Δ*TXx* (Figure 6a, left) is generally small and negative ($-1 \lesssim \Delta TXx < 0$°C) at low to midlatitudes, changing to positive at higher latitudes, with the transition occurring at ~40°N and ~50°S. Notable peaks of positive sign in Δ*TXx* occur at the equator (in two thirds of the models) and between 50–70°N (all models except EC-EARTH). Negative *TXx* anomalies are present in *G1* in subtropical spring and also in Southern Hemisphere summer from 30–45°S. The high-latitude seasonality of Δ*TXx* is opposite to that of Δ*TNn* (Figure 5a and supporting information Figure S5), with a large *TXx* anomaly occurring from May through October at 45–70°N, of peak magnitude +1.7°C in July at around 60°N. This feature is difficult to explain on the basis of the spatial pattern and seasonality of the forcings alone. However, as the same feature appears in the corresponding plot for *abrupt4* × CO_2 *− piControl* (Figure 6b), it likely reflects a higher

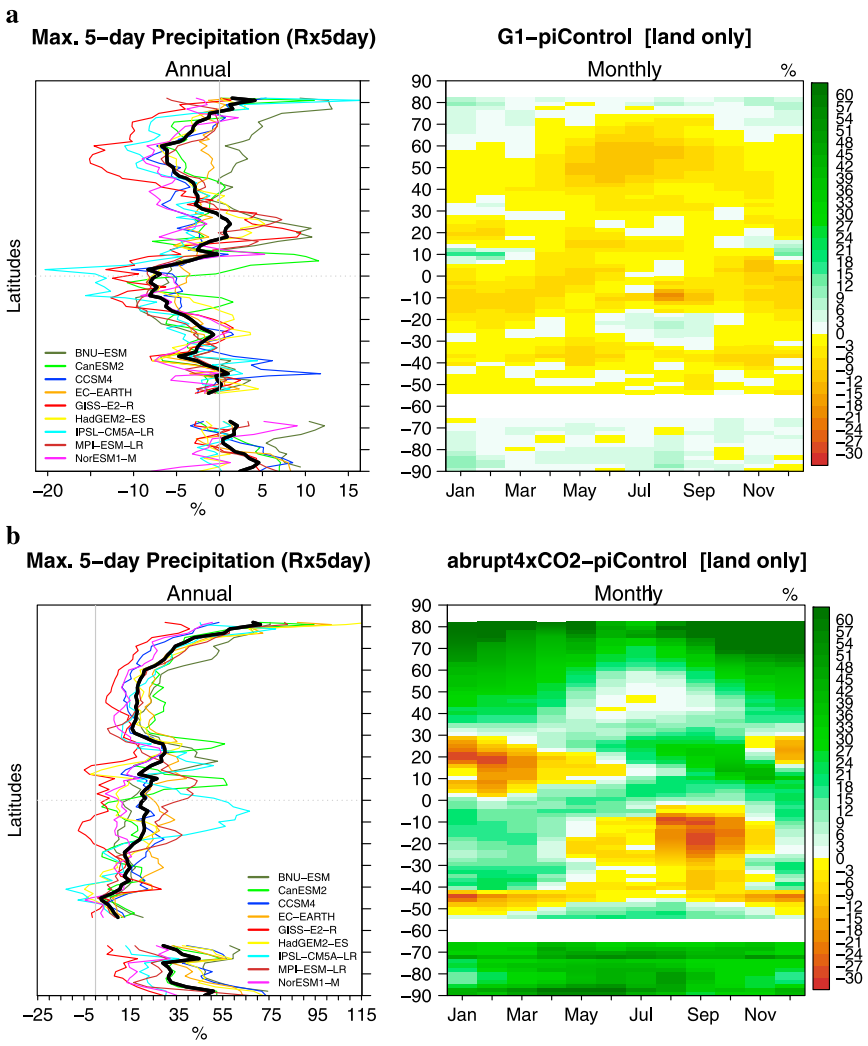

Figure 7. As in Figure 5 but for the extreme index *Rx5day*. Differences are in percent in all panels.

sensitivity of the models to the *abrupt4* × CO_2 forcing compared with the reduced solar forcing over those latitudes and months. This point is discussed at greater length in section 4.3. Again, Figure 6a shows that BNU-ESM has a markedly larger Δ*TXx* response than the other models north of 55°N, more than 4 times the multimodel mean value at 70°N.

Corresponding results for the extreme precipitation index *Rx5day* are presented in Figure 7. The most prominent feature in the monthly *G1 − piControl* plot (Figure 7a) is a 8–13% reduction of *Rx5day* in Northern Hemisphere midlatitudes from about April to September, approximately coincident with the *TXx* maximum discussed above but beginning 1 to 2 months earlier. Other than a comparable decrease just south of the equator, this is the only regional feature in Δ*Rx5day* for which all the models are in agreement (Figure 7a, left). As shown in Figure 7b, the annual, multimodel mean land-only response of *Rx5day* to the *abrupt4* × CO_2 increase is positive at all latitudes, with most models simulating smaller increases at low (~20%) compared to high latitude (a factor of 2–3 larger). At high latitudes, this increase in extremely wet episodes persists year-round. These results are consistent with those found by *Sillmann et al.* [2013a] for the related indices *R95p* (95th percentile of precipitation *P* on days with $P \geq 1$ mm) and *PCPTOT* (total annual or monthly precipitation *P* summed over all days with $P \geq 1$ mm) in the RCP8.5 scenario. Seasonally, however, the pattern of Δ*Rx5day* in *abrupt4* × CO_2 is more complex. Increased occurrence of extreme rainfall is seen in summer and autumn at nearly all latitudes, but prominent decreases (~25–30%) are observed in subtropical winter and spring.

4.3. Forcing and Response Under SRM Geoengineering

In order to better understand the origin of the Northern Hemisphere midlatitude summer anomalies in *TXx* and *Rx5day* (Figures 6a and 7a), and also the prominent *WSDI* increases over the continental tropics (Figure 4c), we examined the response of a larger set of monthly mean variables readily available for the model CanESM2. As can be seen in Figures 5–7 (left), CanESM2 lies relatively close to the multimodel mean for temperature and precipitation extremes and hence should be reasonably representative of the ensemble. In this model, both the spatiotemporal pattern of monthly mean surface air temperature T (Figure 8a) and the geographical pattern of summertime T (Figure 8b) are similar to those of *TXx* (Figures 3c and 6a, respectively); hence, it is plausible that both anomalies share the same origin. As shown in Figure 8b, the regions of largest ΔT in Northern Hemisphere summer (JJAS mean) are concentrated over land in northwestern Eurasia, especially the Baltics, Quebec and Atlantic Canada, Alaska, and the Yukon and Northwest Territories of Canada. Other large ΔT increases in the JJAS map, which agree with the ΔTXx distribution, are seen in eastern Russia, Japan, and China. Notable ΔT peaks over Amazonia and equatorial Africa in Figure 8b also correspond to features in the ΔTXx (not statistically significant) and $\Delta WSDI$ maps (significant).

Over the same regions in JJAS, corresponding decreases in precipitation and evaporation (of order 20 to 40%) are seen (Figures 9b, 10a, and 10b), as part of broader features that encompass much of Asia and North America, for example. Large, and much more localized, decreases in total cloud fraction are also simulated in this season (Figures 9c and 10c), ranging in magnitude from −5 to −20%, with the largest decrease coinciding with the maximum $\Delta T = 2.3$ K over Scandinavia (Figure 10c). Similar results were found for the annual mean cloud cover change in the four-model *G1* ensemble of *Schmidt et al.* [2012]. As shown in the remaining panels of Figures 9 and 10, coincident patterns of change are found in absorbed shortwave radiation at the surface (positive), soil moisture (negative), and springtime (FMAM mean) snow cover (negative).

The above results can be understood within the context of previous studies in which the effects of solar and high-CO_2 forcings were examined either individually [*Gregory and Webb*, 2008; *Andrews et al.*, 2009; *Bala et al.*, 2010; *Andrews et al.*, 2011] or in combination [*Bala et al.*, 2008; *Cao et al.*, 2012; *Fyfe et al.*, 2013]. This prior work established that the observed patterns are the sum of distinct fast and slow responses of the system to these forcings. Here the term "fast" refers to those responses that occur in the absence of annual and global mean surface temperature change $\langle \Delta T \rangle$, typically on the time scale of 1 year or less, as distinguished from the "slow" or "feedback" response that occurs on longer timescales as a result of nonzero ΔT [*Hansen et al.*, 2005; *Andrews et al.*, 2009; *Bala et al.*, 2010]. The fast response is typically isolated by conducting experiments in which the atmospheric model is exposed to fixed sea surface temperatures (SSTs), so that ΔT is guaranteed to be zero over ocean and small over land. Because $\langle \Delta T \rangle \simeq 0$ in *G1*, the results of these fixed SST experiments are a relevant and valuable analog to the current situation of globally balanced forcings under *G1*.

There are three rapid responses of interest in the context of the *G1* experiment. First, the change in atmospheric radiative heating rates due to the large CO_2 increase induces a response of the vertical temperature profile in the troposphere, with consequent changes in atmospheric stability, moisture, and clouds [*Gregory and Webb*, 2008]. An overall reduction of cloud amount in experiments with fixed SSTs has been shown to cause a significant increase in shortwave forcing and a decrease in longwave forcing at the surface [*Gregory and Webb*, 2008]. Moreover, a recent study by N. Huneeus et al. (Solar irradiance reduction and increased CO_2 radiative forcings and feedbacks in the GeoMIP ensemble, submitted to *Journal of Geophysical Research: Atmospheres*, 2013) applied the Gregory method to the difference of *abrupt4*$\times CO_2$ and *G1* specifically and found changes in both shortwave and longwave cloud forcing consistent with the operation of the fast response in the GeoMIP simulations. Second, as shown by *Bala et al.* [2008] and *Andrews et al.* [2009], the dominant effect of reduced solar forcing on the surface energy budget is a decrease in latent heat flux that weakens the hydrological cycle and causes a reduction in both precipitation and evaporation. And finally, the large CO_2 increase in *G1* prompts a strong physiological response of terrestrial vegetation, wherein plants reduce their stomatal conductance with a consequent decrease in evapotranspiration over land [*Andrews et al.*, 2011; *Cao et al.*, 2012]. In coupled models including an interactive terrestrial carbon cycle, the latter two effects have recently been shown to be of comparable magnitude [*Fyfe et al.*, 2013; *Tilmes et al.*, 2013].

Figures 9b, 10a, and 10b reflect these expected changes in precipitation and evaporation over both high latitude and equatorial land areas and further suggest a strong coupling of the hydrological cycle with

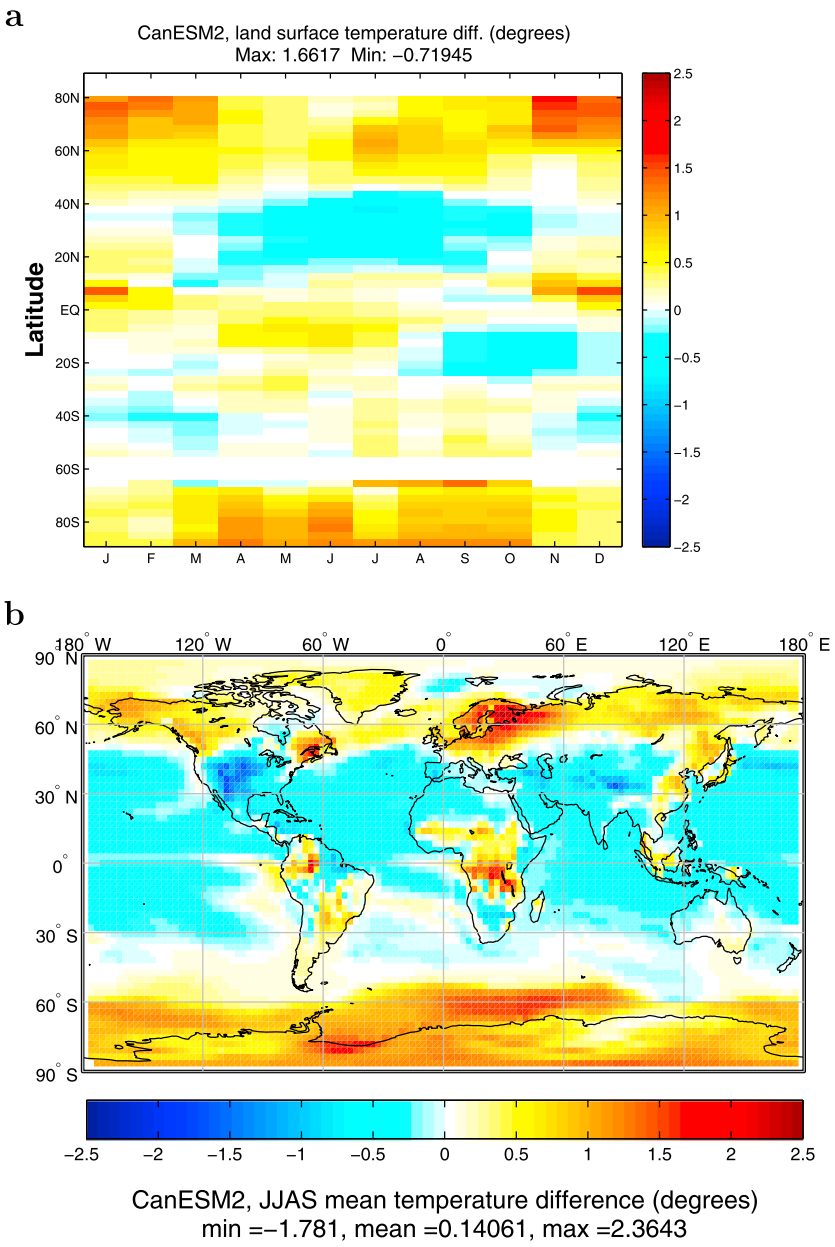

Figure 8. (a) Zonal and monthly mean land-only anomaly (*G1* minus *piControl*) in surface air temperature and (b) June–September surface air temperature anomaly (in degrees Celsius) from the three-member ensemble of CanESM2. The *G1* means are taken over the fiducial 40 year analysis period, while the *piControl* means are for the entire control period of CanESM2 (2015–2810).

cloud cover (Figures 9c and 10c), through reduced water availability to the atmosphere over land areas. A reduction of mean (not shown) and extreme precipitation in springtime evident at both low and high northern latitudes in Figure 7a contributes to a reduction in soil moisture (Figure 9f) and a decrease in evaporation propagating into summer (Figure 9b). This particular land-atmosphere feedback mechanism was discussed for instance in *Fischer et al.* [2007], *Seneviratne et al.* [2010], and by *Mueller and Seneviratne* [2012] in the context of observed heat waves. Reduced cloud cover enhances solar absorption at the increasingly arid land surface during the summer months, resulting in hotter daytime temperatures, i.e., increased *TXx* in *G1* (Figures 3b and 6a).

It is reasonable that increased absorption of shortwave radiation preferentially increases *TXx* over *TNn*, since the former is directly linked to daytime temperature. Indeed, the effect of decreased cloud cover alone would be expected to lower *TNn*, since nighttime cooling to space is increased with fewer clouds. However,

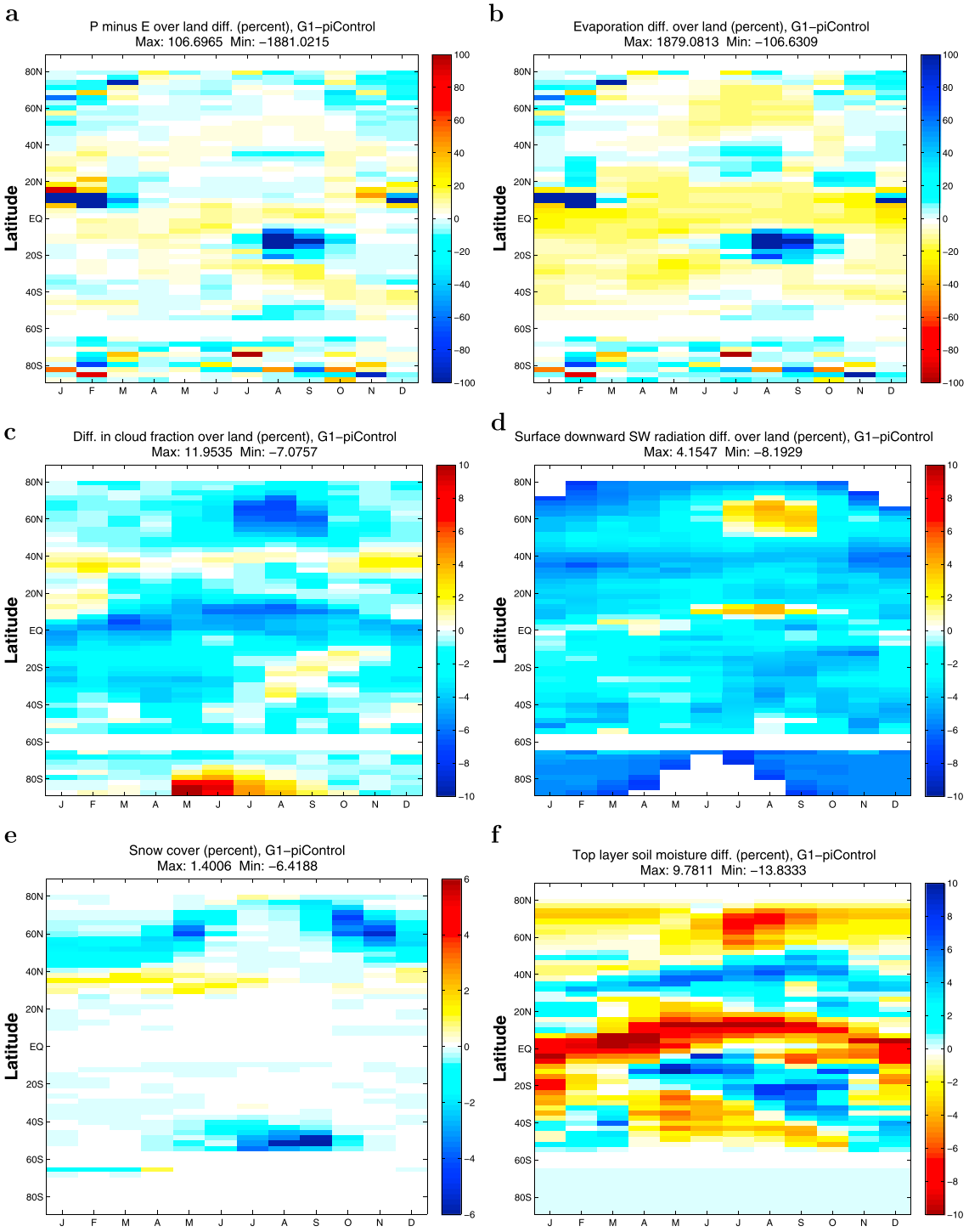

Figure 9. Zonal and monthly mean land-only anomalies (*G1* minus *piControl*) in selected monthly variables from the three-member ensemble of CanESM2: (a) surface air temperature; (b) precipitation minus evaporation; (c) total cloud fraction; (d) absorbed shortwave radiation at the surface; (e) snow cover; and (f) top layer (0–10 cm) soil moisture. The *G1* means are taken over the fiducial 40 year analysis period, while the *piControl* means are for the entire control period of CanESM2 (2015–2810). Units for each quantity are given in the title of each panel.

this cooling is evidently small compared to the large increase in longwave heating from the $4 \times CO_2$ forcing, judging from the positive *TNn* anomalies north of 35°N shown in Figure 5a.

In summary, while reduced solar irradiance in *G1* largely offsets the warming and drying effects of quadrupled atmospheric CO_2 concentration in most areas, it also decreases latent heat flux and precipitation over

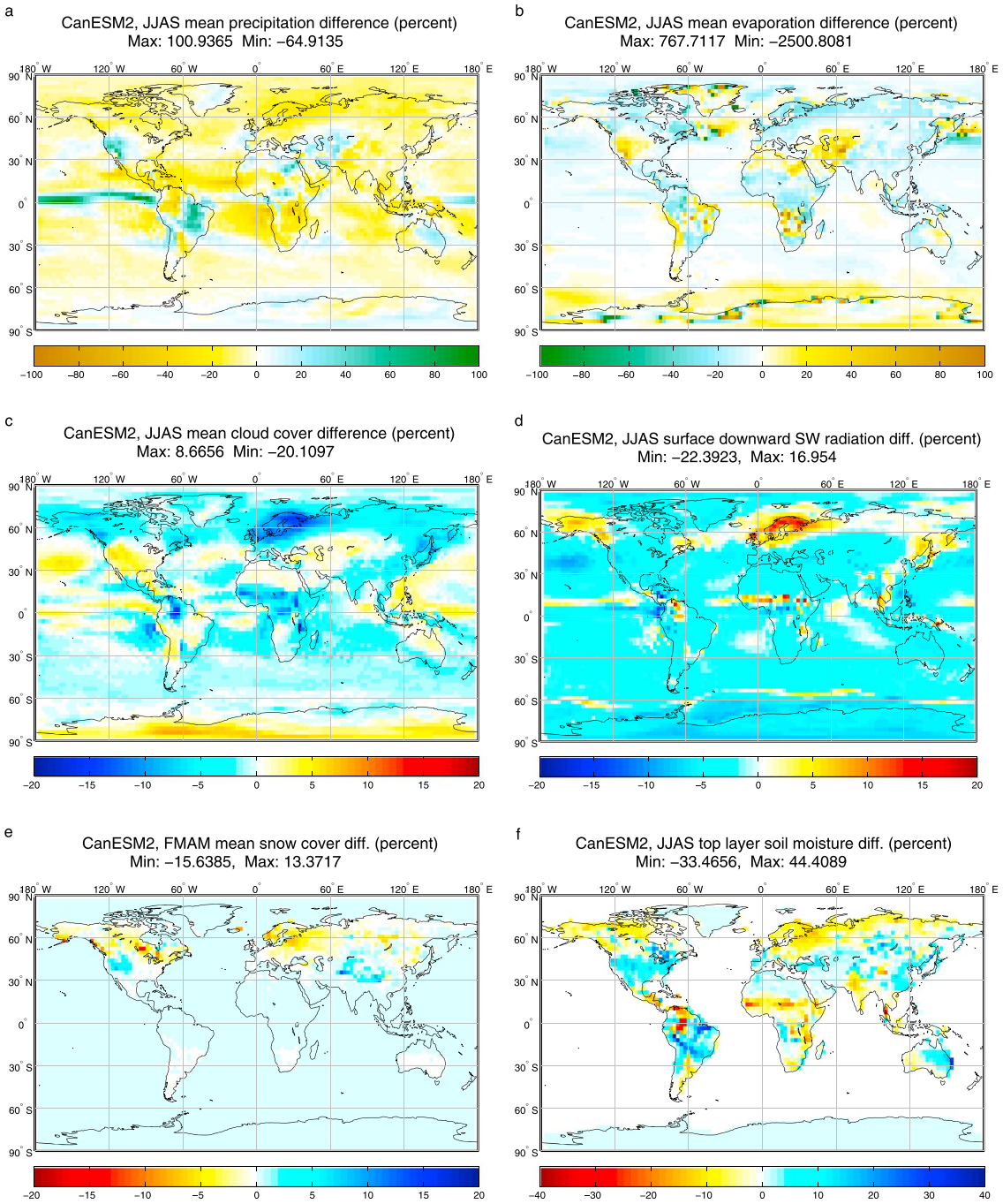

Figure 10. Geographical distributions of differences (*G1* − *piControl*) for selected monthly variables from the three-member ensemble average of CanESM2: (a) precipitation; (b) evaporation; (c) total cloud fraction; (d) absorbed shortwave radiation at the surface; (e) snow cover; and (f) top layer (0–10 cm) soil moisture. The *G1* means are taken over the fiducial 40 year analysis period, while the *piControl* means are for the entire control period of CanESM2 (2015–2810). Units for each quantity are given in the title of each panel.

land, leading to preferential increases of *TXx* compared to *TNn* over land areas in summer. While quantifying each of these mechanisms and their associated feedbacks is not possible without additional experiments, it is an interesting topic for further research.

4.4. Efficacy of SRM Geoengineering With Regard to Climate Extremes

The results of the preceding sections make clear that, despite the maintenance of a state wherein the global and annual mean surface temperature is approximately equal to that of *piControl*, complex regional patterns

Table 3. Efficacy[a] of *G1* for Neutralizing Annual Extremes in *abrupt4* × CO_2

Variable	Global	Land	Ocean
TNn	0.87	0.86	0.88
TXx	0.89	0.88	0.91
Rx5day	0.71	0.77	0.70
CSDI	−1.34	0.0743	−1.55
WSDI	0.96	0.92	0.98
CDD	0.71	0.77	0.67

[a]The efficacy *e*, defined in equation (2), is the fraction of RMS change in *abrupt4* × CO_2 relative to *piControl* that is offset by the solar irradiance reduction in *G1*. The value of *e* ranges from negative values (indicating differences are larger in *G1* than in *abrupt4* × CO_2) through 0 (low efficacy) to 1 (high efficacy). RMS values for each experiment are provided in Table 2 of the supporting information.

of extreme temperature and precipitation persist in the *G1* experiment. The extent to which SRM is able to reduce the change in a particular climate extreme index in any location of interest can be assessed by examining the figures presented thus far in the paper.

However, it can also be helpful to summarize this efficacy of climate engineering using a simpler quantitative measure. To do so, we adopt a metric similar to that used by *Kravitz et al.* [2013], who calculated the root-mean-squared differences between experiments at each model grid point:

$$RMS_{\mathrm{Exp-Ctl}}(X) = \sqrt{\frac{\Sigma(X_{\mathrm{Exp}} - X_{\mathrm{Ctl}})^2 \mathrm{d}A}{\Sigma \mathrm{d}A}}, \qquad (1)$$

where X is the extreme index of interest, Exp is the perturbed experiment (*G1* or *abrupt4* × CO_2), Ctl is the *piControl* experiment, and Σ is a sum over the grid cells (dA) of a specified area A (e.g., global, land-only, ocean-only). We can then define the efficacy of *G1* as a geoengineering strategy as

$$e(X) \equiv 1 - \frac{RMS_{\mathrm{G1-Ctl}}(X)}{RMS_{4\times CO2-\mathrm{Ctl}}(X)} = 1 - r(X), \qquad (2)$$

where $r(X)$ is the metric used by *Kravitz et al.* [2013] in their analysis of monthly mean climatic variables. Where SRM geoengineering is most effective at offsetting the *abrupt4* × CO_2 response, *G1* values approach those of *piControl* and $e(X) \rightarrow 1$, while $e(X) = 0$ indicates no difference from the *abrupt4* × CO_2 state. Negative values of $e(X)$ can result if *RMS* differences from *piControl* in *G1* exceed those in *abrupt4* × CO_2. In the present context, such values indicate that the climate response in *G1* overshoots the target of a mere restoration of *piControl* conditions. A compilation of global, land-only, and ocean-only *e* values for all six extreme indices is presented in Table 3. In the following, we refer to these values in terms of percentages.

The efficacies in *G1* for offsetting changes in annual *TNn* and *TXx* under *abrupt4* × CO_2 conditions in the model ensemble are 87% and 89%, respectively, over the entire globe. Comparing with the value $e(T) = 91\%$ obtained by *Kravitz et al.* [2013] from monthly mean temperatures, we conclude that *G1* is slightly less effective at neutralizing extreme temperature changes under strong CO_2 forcing than for moderating monthly mean temperature changes. On the other hand, while the efficacy of *G1* for offsetting the sizable *RMS* 5 day precipitation change under *abrupt4* × CO_2 is smaller than for temperature extremes, i.e., $e(Rx5day) = 71\%$ in the global case, it is considerably larger than that obtained by *Kravitz et al.* [2013] for mean precipitation [i.e., $e(P) = 55\%$].

With regard to the duration indices, Table 3 reveals that the large increases in *WSDI* seen in *abrupt4* × CO_2 (Figure 4d) are very effectively neutralized in *G1* ($e = 96\%$), especially over ocean (98%). This is consistent with the high efficacy for mean temperature differences found by *Kravitz et al.* [2013], insofar as *WSDI* may be viewed as a time-integrated version of *TXx*, and thus more akin to monthly mean than maximum annual temperature. Moreover, the overcooling phenomenon seen at low latitudes in *G1* and discussed in section 4.2 is expressed clearly in the negative values of $e(CSDI)$ in Table 3. By this metric, overcooling is essentially an ocean-only phenomenon. This could be the result of the reduced evaporative cooling over land [*Tilmes et al.*, 2013] coupled with the lower albedo of the ocean (meaning that SRM has a larger direct effect there). Finally, the efficacy of *G1* for moderating dry spells in *abrupt4* × CO_2 is $e(CDD) = 71\%$ globally, equal to the figure obtained for *Rx5day*. Thus, the frequencies of extreme high precipitation and extended drought conditions are reduced roughly equally in *G1* compared to *abrupt4* × CO_2.

While $e(X)$ is undoubtedly a highly simplified measure of the complex regional differences in extremes exhibited in the foregoing sections, it nevertheless leads to two useful conclusions regarding the efficacy of the *G1* experiment as a strategy for reducing climate extremes encountered in a *abrupt4* × CO_2 world: (1) *G1* is more effective in reducing changes in temperature extremes compared to changes in precipitation extremes; (2) *G1* is less effective for reducing changes in temperature extremes than mean temperature

changes but more effective for reducing changes in precipitation extremes versus means (where the mean changes were examined by *Kravitz et al.* [2013]).

5. Summary and Concluding Remarks

We have examined temperature and precipitation changes with a particular focus on extremes in a multimodel, idealized geoengineering experiment (GeoMIP *G1*), in which the global radiative forcing due to an instantaneous quadrupling of CO_2 is compensated by a commensurate reduction of the solar irradiance. While *G1* is a highly idealized experiment, it represents an effective sensitivity test of model-simulated climate in the context of geoengineering and serves to illustrate certain mechanisms by which the Earth system might respond to a deliberate manipulation of external climate forcings.

In previous investigations of the *G1* experiment with ensembles composed of many of the same models, it was found that the global mean surface temperature is approximately restored to preindustrial levels but at the cost of a global reduction in precipitation and significant residual temperature changes at the regional scale (e.g., cooling of the tropical ocean and warming at high latitudes [*Schmidt et al.*, 2012; *Kravitz et al.*, 2013]). By analyzing probability density functions of monthly surface temperature ($\mathcal{T}$) and precipitation ($\mathcal{P}$) standardized anomalies, we found that the mean and variance of both distributions differed significantly between the *G1* and *piControl* simulations (Figure 1). The differences are most apparent in the tails of the distribution and reflect distinct regional patterns, including an extension of the high-$\mathcal{T}$ tail over land, of the low-$\mathcal{T}$ tail over ocean, and a general shift of $\mathcal{P}$ to drier conditions. The changes in these tails, which have direct implications for the occurrence of extreme events, were further investigated in the remainder of the paper.

Using several of the indices for climate extremes defined by the ETCCDI (Table 2), we analyzed a number of quantities of interest, namely, absolute extreme events, specifically the coldest night (*TNn*) and warmest day (*TXx*); the duration of extreme conditions, specifically cold spell (*CSDI*), warm spell (*WSDI*), and consecutive dry day (*CDD*) durations; and the maximum consecutive 5 day precipitation amount (*Rx5day*). For global mean quantities in the *G1* experiment, changes in *TXx* can be more effectively neutralized than those in *TNn* as a consequence of the fact that reducing the solar irradiance has a more direct effect on daytime compared to nighttime temperatures. The global average *Rx5day* is reduced in *G1* compared to *piControl* in accordance with the general reduction of precipitation, as discussed in the case of monthly mean extremes by *Tilmes et al.* [2013].

Examination of the geographical and seasonal patterns of anomalies in the extreme indices reveals further interesting features (Figures 3–7). For *TNn*, the strong heating of the northern high latitudes simulated under *abrupt4* × CO_2 is very much alleviated under *G1*, but a significant residual warming remains in that region compared to *piControl* and also over most land areas, particularly in winter. By contrast, and consistent with the dominance of the solar forcing pattern at low latitudes, significant overcooling of the tropical oceans is simulated. This overcooling is also manifested in the increase of cold spell duration as represented by *CSDI*. This index is based on the climatology of daily minimum temperatures in *piControl* and thus indicates that the tropical ocean is significantly cooler over an extended period of time (about 2 months) in *G1* compared to *piControl*. This could influence several processes, including ocean circulation patterns, thermally driven atmospheric flows in coastal areas, the frequency of cyclogenesis, and marine biology. As these are areas of high biological productivity in the contemporary climate, where high SSTs support rapid nutrient recycling [*Schmittner et al.*, 2008], a sizeable decrease in upper ocean heat content due to reduced tropical SSTs in *G1* might therefore have deleterious consequences for pelagic ecosystems (however, dynamical responses might also be important [*Sarmiento and Gruber*, 2006]).

While the warming anomaly pattern of *TNn* in *G1* is essentially a weakened version of that seen in the *abrupt4* × CO_2 experiment, the same cannot be said for the geographic distribution of ΔTXx. First, the cooling anomalies in *TXx* over both land and ocean in *G1* extend into the midlatitudes, while they are essentially an ocean-only phenomenon in *TNn*. Second, the pattern of residual warming of *TXx* in *G1* over northern extratropical land, with peaks over central and northern Europe, eastern Canada, and northeast Asia, does not resemble the *abrupt4* × CO_2 warming pattern.

By examining the spatiotemporal patterns of change for a range of variables in a single model (CanESM2), we concluded that a number of previously identified rapid feedback processes are likely operating in

Acknowledgments
We thank all participants of the Geoengineering Model Intercomparison Project and their model development teams, CLIVAR/WCRP Working Group on Coupled Modeling for endorsing GeoMIP, and the scientists managing the Earth System Grid (ESG) data nodes who have assisted with making GeoMIP output available. Most of the model output used in this study was obtained from the ESG node at PCMDI, http://pcmdi9.llnl.gov/esgf-web-fe/, or from the BADC and BNU peer nodes accessible from that location. The *G1* output for the IPSL, MPI, and NorESM models was obtained from the IMPLICC project via http://implicc1.dkrz.de:8080/thredds/catalog/. The EC-EARTH output was supplied by one of the coauthors (S.Y.), to whom requests should be directed. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. C.C. is supported by a Canadian NSERC grant (CRDPJ 403886-10) and J.S. by the German Research Foundation (grant Si 1659/1-1) and the AERO-CLO-WV project (184714/S30) funded by the Norwegian Research Council. K.A. is funded by the European Commission's 7th Framework Program through IMPLICC (FP7-ENV-2008-1-226567) and by the Norwegian Research Council through the EarthClim project (207711/E10) and its program for supercomputing (NOTUR). D.J. and J.C.M. thank all members of the BNU-ESM model group and support from the Joint Center for Global Change Studies (JCGCS) as well as the Center of Information and Network Technology at Beijing Normal University for assistance in publishing the GeoMIP data set. J.E.K. and H.M. acknowledge support from the IMPLICC project and the EuTRACE project (grant agreement 306395). Simulations with the MPI-ESM, NorESM, and IPSL-CM5A models were supported through the IMPLICC project. B.K. is supported by the Fund for Innovative Climate and Energy Research (FICER). The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830. Simulations performed by B.K. were supported by the NASA High-End Computing (HEC) Program through the NASA Center for

G1, prompted by the combination of reduced solar irradiance and increased CO_2 concentration. These responses occur both at the surface and in the troposphere and act in concert to decrease available soil moisture and cloud cover. In CanESM2, and likely in the other GeoMIP models as well, it is the reduction in midlatitude to high-latitude Northern Hemisphere cloud cover in spring and summer that leads to increased incident solar radiation at the surface, thereby elevating *TXx* preferentially over *TNn*.

While a more detailed understanding of the cloud response is beyond the scope of this work, there is an important lesson here for the possibility of unexpected changes in the climate system under proposed geoengineering experiments. In this case, despite the overall reduction in solar irradiance occurring under *G1*, internal feedbacks within the system lead to regional *increases* in absorbed solar radiation at the surface, which in turn prompt discernable increases in daily maximum and monthly mean temperatures over large land areas at northern latitudes. In these northern continental locations, the increased incidence of dry spring and hot summer conditions represents an instance of where the application of global SRM might incur significant regional climate impacts.

We also considered the efficacy of the *G1* experiment as a strategy for reducing climate extremes encountered in a high-CO_2 world by quantifying the differences between the *piControl*, *G1*, and *abrupt4*$\times CO_2$ multimodel experiments. It was found that while *G1* is more effective in reducing changes in temperature extremes compared to precipitation extremes and for reducing changes in precipitation extremes versus means, it is less effective at reducing changes in temperature extremes compared to means.

Our results reinforce previous findings regarding the regional disparity in climatic response that occurs as a result of conceptually simple SRM measures such as *G1* [*Irvine et al.*, 2010; *Ricke et al.*, 2010]. As shown here, the operation of climate feedbacks creates regional anomalies that cannot be anticipated from a consideration of the forcings alone [*Armour et al.*, 2013]. Indeed, this is one justification for the simplified experimental design of GeoMIP, which allows this anomalous behavior to be clearly detected in the absence of other complicating factors. Minimizing these regional anomalies would require the manipulation of forcings at the regional scale, an idea that is beginning to be explored [*Ban-Weiss and Caldeira*, 2010; *MacMartin et al.*, 2013]. In general, however, the mechanisms behind the myriad robust regional changes seen in *G1* and similar experiments have yet to be fully investigated.

Finally, it is important to recognize that any climate engineering scheme that depends entirely on SRM, no matter how complex, will leave many other problems associated with historically unprecedented levels of CO_2 in Earth's atmosphere untouched. Among the best studied of these is ocean acidification [*Doney et al.*, 2009], but there are others associated with biogeochemical processes on land (as discussed in section 4.2.2 [*Andrews et al.*, 2011]) and anomalous cooling of the stratosphere [*Tilmes et al.*, 2009], to name but two. Hence, we join with other researchers in not only highlighting the caution that should be taken with respect to geoengineering the Earth's climate but also with extrapolating the results of simplified sensitivity experiments such as *G1* to other contexts. SRM via stratospheric aerosol injection, for example, presents a host of additional factors that modify the climate response beyond the first-order reduction of incoming solar radiation. These have been explored by way of sensitivity studies, but the results are highly dependent on the model treatment of the aerosol, adding additional uncertainty to the simulated response [*Robock et al.*, 2008; *Tilmes et al.*, 2009; *Heckendorn et al.*, 2009]. While our study provides a preliminary assessment of expected changes in surface temperature, precipitation, and related extremes in the context of a simple geoengineering sensitivity experiment, further work is needed to improve our understanding of the internal processes governing these changes and in exploring the effects of SRM on additional climatic variables.

Author Contributions

J.S. conceived the original idea for the paper; C.L.C. and J.S. designed and performed the research; D.B., C.L.C., and J.S. analyzed the model output; C.L.C. and J.S. wrote the paper; and the remaining authors performed the simulations (including C.L.C.) and contributed critical remarks on the manuscript.

References

Alexander, L. V., et al. (2006), Global observed changes in daily climate extremes of temperature and precipitation, *J. Geophys. Res.*, *111*, D05109, doi:10.1029/2005JD006290.

Alterskjaer, K., J. E. Kristjansson, O. Boucher, H. Muri, U. Niemeier, H. Schmidt, M. Schulz, and C. Timmreck (2013), Sea-salt injections into the low-latitude marine boundary layer: The transient response in three Earth system models, *J. Geophys. Res. Atmos.*, *118*, 12,195–12,206, doi:10.1002/2013JD020432.

Climate Simulation (NCCS) at Goddard Space Flight Center. A.R. is supported by NSF grants AGS-1157525 and CBET-1240507. S.T. is funded through the National Center for Atmospheric Research and the National Science Foundation. Finally, we are grateful to Peter Irvine, Adam Monahan, Phil Rasch, and three anonymous reviewers whose comments helped improve the manuscript.

Andrews, T., P. M. Forster, and J. M. Gregory (2009), A surface energy perspective on climate change, *J. Clim.*, *22*, 2557–2570, doi:10.1175/2008JCLI2759.1.

Andrews, T., M. Doutriaux-Boucher, O. Boucher, and P. M. Forster (2011), A regional and global analysis of carbon dioxide physiological forcing and its impact on climate, *Clim. Dyn.*, *36*, 783–792.

Angel, R. (2006), Feasibility of cooling the Earth with a cloud of small spacecraft near the inner Lagrange point (L1), *Proc. Natl. Acad. Sci. U.S.A.*, *103*(46), 17,184–17,189.

Armour, K. C., C. M. Bitz, and G. H. Roe (2013), Time-varying climate sensitivity from regional feedbacks, *J. Clim.*, *26*, 4518–4534, doi:10.1175/JCLI-D-12-00544.1.

Bala, G., P. B. Duffy, and K. E. Taylor (2008), Impact of geoengineering schemes on the global hydrological cycle, *Proc. Natl. Acad. Sci. U.S.A.*, *105*(22), 7664–7669, doi:10.1073/pnas.0711648105.

Bala, G., K. Caldeira, and R. Nemani (2010), Fast versus slow response in climate change: Implications for the global hydrological cycle, *Clim. Dyn.*, *35*, 423–434.

Ban-Weiss, G. A., and K. Caldeira (2010), Geoengineering as an optimization problem, *Environ. Res. Lett.*, *5*, 034009, doi:10.1088/1748-9326/5/3/034009.

Caesar, J., and J. A. Lowe (2012), Comparing the impacts of mitigation versus non-intervention scenarios on future temperature and precipitation extremes in the HadGEM2 climate model, *J. Geophys. Res.*, *117*, D15109, doi:10.1029/2012JD017762.

Cao, L., G. Bala, and K. Caldeira (2012), Climate response to changes in atmospheric carbon dioxide and solar irradiance on the time scale of days to weeks, *Environ. Res. Lett.*, *7*, 034015, doi:10.1088/1748-9326/7/3/034015.

Donat, M., L. Alexander, H. Yang, I. Durre, R. Vose, and J. Caesar (2013), Global land-based datasets for monitoring climatic extremes, *Bull. Am. Meteorol. Soc.*, *94*, 997–1006, doi:10.1175/BAMS-D-12-00109.1.

Doney, S. C., V. J. Fabry, R. A. Feely, and J. A. Kleypas (2009), Ocean acidification: The other CO_2 problem, *Annu. Rev. Mar. Sci.*, *1*, 169–192, doi:10.1146/annurev.marine.010908.163834.

Donohoe, A., and D. S. Battisti (2013), The seasonal cycle of atmospheric heating and temperature, *J. Clim.*, *26*, 4962–4980, doi:10.1175/JCLI-D-12-00713.1.

Dwyer, J. G., M. Biasutti, and A. H. Sobel (2012), Projected changes in the seasonal cycle of surface temperature, *J. Clim.*, *25*, 6359–6374, doi:10.1175/JCLI-D-11-00741.1.

Fischer, E. M., S. I. Seneviratne, D. Luthi, and C. Schär (2007), Contribution of land-atmosphere coupling to recent European summer heat waves, *Geophys. Res. Lett.*, *34*, L06707, doi:10.1029/2006GL029068.

Frich, P., L. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. Klein Tank, and T. Peterson (2002), Observed coherent changes in climate extremes during the second half of the twentieth century, *Clim. Res.*, *19*, 193–212.

Fyfe, J. C., J. N. S. Cole, V. K. Arora, and J. F. Scinocca (2013), Biogeochemical carbon coupling influences global precipitation in geoengineering experiments, *Geophys. Res. Lett.*, *40*, 651–655, doi:10.1002/grl.50166.

Gardiner, S. M. (2010), Is "Arming the future" with geoengineering really the lesser evil? Some doubts about the ethics of intentionally manipulating the climate system, in *Climate Ethics: Essential Readings*, edited by S. M. Gardiner et al., pp. 284–314, Oxford Univ. Press, Oxford, U. K., and New York.

Good, P., J. M. Gregory, J. A. Lowe, and T. Andrews (2013), Abrupt CO_2 experiments as tools for predicting and understanding CMIP5 representative concentration pathway projections, *Clim. Dyn.*, *40*, 1041–1053.

Govindasamy, B., and K. Caldeira (2000), Geoengineering Earth's energy balance to mitigate CO_2-induced climate change, *Geophys. Res. Lett.*, *27*(14), 2141–2144.

Gregory, J. M., and M. Webb (2008), Tropospheric adjustment induces a cloud component in CO_2 forcing, *J. Clim.*, *21*, 58–71, doi:10.1175/2007JCLI1834.1.

Hansen, J., et al. (2005), Efficacy of climate forcings, *J. Geophys. Res.*, *110*, D18104, doi:10.1029/2005JD005776.

Heckendorn, P., D. Weisenstein, S. Fueglistaler, B. P. Luo, E. Rozanov, M. Schraner, L. W. Thomason, and T. Peter (2009), The impact of geoengineering aerosols on stratospheric temperature and ozone, *Environ. Res. Lett.*, *4*, 045108, doi:10.1088/1748-9326/4/4/045108.

Held, I. M., and B. J. Soden (2006), Robust responses of the hydrological cycle to global warming, *J. Clim.*, *19*, 5686–5699.

Hollander, M., and D. A. Wolfe (1973), *Nonparametric Statistical Methods*, 503 pp., John Wiley, New York.

IPCC (2012), *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*, Cambridge Univ. Press, Cambridge, U. K., and New York.

Irvine, P. J., A. Ridgwell, and D. J. Lunt (2010), Assessing the regional disparities in geoengineering impacts, *Geophys. Res. Lett.*, *37*, L18702, doi:10.1029/2010GL044447.

Jones, A., et al. (2013), The impact of abrupt suspension of solar radiation management (termination effect) in experiment G2 of the Geoengineering Model Intercomparison Project (GeoMIP), *J. Geophys. Res. Atmos.*, *118*, 9743–9752, doi:10.1002/jgrd.50762.

Jones, P. W. (1999), First- and second-order conservative remapping schemes for grids in spherical coordinates, *Mon. Weather Rev.*, *127*, 2204–2210.

Kharin, V. V., F. W. Zwiers, X. Zhang, and G. C. Hegerl (2007), Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, *J. Clim.*, *20*, 1419–1444, doi:10.1175/JCLI4066.1.

Kravitz, B., A. Robock, O. Boucher, H. Schmidt, K. E. Taylor, G. Stenchikov, and M. Schulz (2011), The Geoengineering Model Intercomparison Project (GeoMIP), *Atmos. Sci. Lett.*, *12*, 162–167, doi:10.1002/asl.316.

Kravitz, B., et al. (2013), Climate model response from the Geoengineering Model Intercomparison Project (GeoMIP), *J. Geophys. Res. Atmos.*, *118*, 8320–8332, doi:10.1002/jgrd.50646.

Lenton, T. M., and N. E. Vaughan (2009), The radiative forcing potential of different climate geoengineering options, *Atmos. Chem. Phys.*, *9*, 5539–5561.

Lunt, D. J., A. Ridgwell, P. J. Valdes, and A. Seale (2008), "Sunshade World": A fully coupled GCM evaluation of the climatic impacts of geoengineering, *Geophys. Res. Lett.*, *35*, L12710, doi:10.1029/2008GL033674.

MacMartin, D. G., D. W. Keith, B. Kravitz, and K. Caldeira (2013), Management of trade-offs in geoengineering through optimal choice of non-uniform radiative forcing, *Nat. Clim. Change*, *3*, 365–368, doi:10.1038/nclimate1722.

Mann, M. E., and J. Park (1996), Greenhouse warming and changes in the seasonal cycle of temperature: Model versus observations, *Geophys. Res. Lett.*, *23*, 1111–1114.

Moore, J. C., et al. (2014), Arctic sea ice and atmospheric circulation under the GeoMIP *G1* scenario, *J. Geophys. Res. Atmos.*, *119*, 567–583, doi:10.1002/2013JD021060.

Mueller, B., and S. Seneviratne (2012), Hot days induced by precipitation deficits at the global scale, *Proc. Natl. Acad. Sci. U.S.A.*, *109*(31), 12,398–12,403, doi:10.1073/pnas.1204330109.

Orlowsky, B., and S. I. Seneviratne (2012), Global changes in extreme events: Regional and seasonal dimension, *Clim. Change*, *110*(3), 669–696, doi:10.1007/s10584-011-0122-9.

Peterson, T., and M. Manton (2008), Monitoring changes in climate extremes: A tale of international collaboration, *Bull. Am. Meteorol. Soc.*, *89*, 1266–1271, doi:10.1175/2008BAMS2501.1.

Ricke, K. L., M. G. Morgan, and M. R. Allen (2010), Regional climate response to solar-radiation management, *Nat. Geosci.*, *3*, 537–541, doi:10.1038/ngeo915.

Robock, A., L. Oman, and G. L. Stenchikov (2008), Regional climate responses to geoengineering with tropical and Arctic SO_2 injections, *J. Geophys. Res.*, *113*, D16101, doi:10.1029/2008JD010050.

Robock, A., A. Marquardt, B. Kravitz, and G. Stenchikov (2009), Benefits, risks, and costs of stratospheric geoengineering, *Geophys. Res. Lett.*, *36*, L19703, doi:10.1029/2009GL039209.

Royal Society (2009), *Geoengineering the Climate: Science, Governance, and Uncertainty*, 82 pp., The Royal Society, London.

Sarmiento, J. L., and N. Gruber (2006), *Ocean Biogeochemical Dynamics*, 528 pp., Princeton Univ. Press, Princeton, N. J.

Schmidt, H., et al. (2012), Solar irradiance reduction to counteract radiative forcing from a quadrupling of CO_2: Climate responses simulated by four earth system models, *Earth Syst. Dyn.*, *3*, 63–78, doi:10.5194/esd-3-63-2012.

Schmittner, A., A. Oschlies, H. D. Matthews, and E. D. Galbraith (2008), Future changes in climate, ocean circulation, ecosystems, and biogeochemical cycling simulated for a business-as-usual CO_2 emission scenario until year 4000 AD, *Global Biogeochem. Cycles*, *22*, GB1013, doi:10.1029/2007GB002953.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth Sci. Rev.*, *99*, 125–161, doi:10.1016/j.earscirev.2010.02.004.

Seneviratne, S. I., et al. (2012), Changes in climate extremes and their impacts on the natural physical environment, in *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, edited by C. B. Field et al., pp. 109–230, A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC), Cambridge Univ. Press, Cambridge, U. K., and New York.

Sillmann, J., V. V. Kharin, F. W. Zwiers, X. Zhang, and D. Bronaugh (2013a), Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, *J. Geophys. Res. Atmos.*, *118*, 2473–2493, doi:10.1002/jgrd.50188.

Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh (2013b), Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, *J. Geophys. Res. Atmos.*, *118*, 1716–1733, doi:10.1002/jgrd.50203.

Tebaldi, C., K. Hayhoe, J. M. Arblaster, and G. A. Meehl (2006), Going to extremes. An intercomparison of model-simulated historical and future changes in extreme events, *Clim. Change*, *79*(3-4), 185–211, doi:10.1007/s10584-006-9051-4.

Tilmes, S., R. R. Garcia, D. E. Kinnison, A. Gettelman, and P. J. Rasch (2009), Impact of geoengineered aerosols on the troposphere and stratosphere, *J. Geophys. Res.*, *114*, D12305, doi:10.1029/2008JD011420.

Tilmes, S., et al. (2013), The hydrological impact of geoengineering in the Geoengineering Model Intercomparison Project GeoMIP, *J. Geophys. Res. Atmos.*, *118*, 11,036–11,058, doi:10.1002/jgrd.50868.

Zhang, X., L. Alexander, G. C. Hegerl, P. Jones, A. K. Tank, T. C. Peterson, B. Trewin, and F. W. Zwiers (2011), Indices for monitoring changes in extremes based on daily temperature and precipitation data, *WIREs Clim. Change*, *2*, 851–870, doi:10.1002/wcc.147.