Climate Feedback to Stratospheric Aerosol Forcing: The Key Role of the Pattern Effect

MORITZ GÜNTHER,[a] HAUKE SCHMIDT,[a] CLAUDIA TIMMRECK,[a] AND MATTHEW TOOHEY[b]

[a] *Max Planck Institute for Meteorology, Hamburg, Germany*
[b] *Institute of Space and Atmospheric Studies, University of Saskatchewan, Saskatoon, Saskatchewan, Canada*

(Manuscript received 29 April 2022, in final form 14 July 2022)

ABSTRACT: Volcanic aerosol forcing has previously been found to cause a weak global mean temperature response, as compared with CO_2 radiative forcing of equal magnitude: its efficacy is supposedly low, but for reasons that are not fully understood. To investigate this, we perform idealized, time-invariant stratospheric sulfate aerosol forcing simulations with the MPI-ESM-1.2 and compare them with 0.5 × CO_2 and 2 × CO_2 runs. While the early decades of the aerosol forcing simulations are characterized by strong negative feedback (i.e., low efficacy), the feedback weakens on the decadal to centennial time scale. Although this effect is qualitatively also found in CO_2-warming simulations, it is more pronounced for stratospheric aerosol forcing. The strong early and weak late cooling feedbacks compensate, leading to an equilibrium efficacy of approximately 1 in all simulations. The 0.5 × CO_2 cooling simulations also exhibit strong feedback changes over time, albeit less than in the idealized aerosol forcing simulations. This suggests that the underlying cause for the feedback change is not exclusively specific to aerosol forcing. One critical region for the feedback differences between simulations with negative and positive radiative forcing is the tropical Indo-Pacific warm-pool region (30°S–30°N, 50°E–160°W). In the first decades of cooling, the temperature change in this region is stronger than the global average, whereas it is stronger outside it for 2 × CO_2 warming. In cooling scenarios, this leads to an enhanced activation of the warm-pool region's strongly negative lapse-rate feedback.

SIGNIFICANCE STATEMENT: Large volcanic eruptions can enhance the scattering aerosol layer in the stratosphere, which leads to a global cooling for a few years. Surprisingly, Earth has been found to cool less from radiative flux perturbations from stratospheric aerosol forcing, in comparison with how much it warms as a result of increases in CO_2 concentration. We find that specific surface temperature change patterns after volcanic eruptions cause this effect. The temperature change in the tropical Indian and western Pacific Ocean determines how much global temperature change is needed to regain radiative equilibrium. Our findings contribute to understanding the climate response to volcanic eruptions and are relevant for understanding the mechanisms of climate change due to changes in CO_2 concentration.

KEYWORDS: Atmospheric circulation; Climate sensitivity; Volcanoes; Feedback; Sea surface temperature; Radiative forcing

1. Introduction

Following large volcanic eruptions, volcanic aerosols are added to the natural stratospheric aerosol layer and persist for a few years. The aerosol layer globally increases the albedo in the shortwave (SW) and absorption in the longwave (LW) spectrum. This constitutes a net negative radiative forcing, which cools Earth. We study volcanic aerosol forcing (VAF) in the broader sense of stratospheric sulfate aerosol forcing (SSAF), independent of the volcano-specific characteristic forcing evolution in time.

VAF has been found to produce smaller temperature change per unit radiative forcing than CO_2 forcing (Hansen et al. 2005; Boer et al. 2007; Gregory and Andrews 2016; Gregory et al. 2016, 2020; Marvel et al. 2016; Modak et al. 2016; Ceppi and Gregory 2019; Zhao et al. 2021). This is equivalent to the statement that the feedback parameter is more negative for VAF than for CO_2 forcing. In model simulations with a slab ocean, and with fixed sea surface temperatures (SSTs) and sea ice, the strongly negative feedback parameter has been shown to be connected to differences in atmospheric stability, which ultimately arise from different SST patterns (Ceppi and Gregory 2017, 2019). We aim to compare the climate feedbacks with SSAF and positive and negative CO_2 radiative forcing on annual to centennial time scales in a coupled atmosphere–ocean model, explain the differences between them, and identify the key regions that cause the distinct climate responses.

First, we introduce the necessary building blocks for our explanation: efficacy, the feedback parameter, and the pattern effect (section 2). This is followed by a literature review, showing that the strong feedback to VAF is a broad consensus from earlier studies (section 3).

To tackle the question of why the feedback parameter appears to be stronger for VAF, we perform three types of simulation with the coupled climate model Max Planck Institute Earth System Model 1.2 (MPI-ESM 1.2; Mauritsen et al. 2019): abrupt

---

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-22-0306.s1.

Corresponding author: Moritz Günther, moritz.guenther@mpimet.mpg.de

DOI: 10.1175/JCLI-D-22-0306.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

$0.5 \times CO_2$, abrupt $2 \times CO_2$, and an idealized SSAF simulation (section 4). With these simulations we not only shed light on the differences between the responses to stratospheric aerosol and CO_2 forcing, but also more generally elaborate disparities between the responses to negative and positive radiative forcing.

We show the different time dependencies of the feedback parameter for simulations with negative and positive radiative forcing in section 5 and establish a relationship to the pattern effect. Thereafter we discuss the origin of the modified temperature change pattern. In section 6 we shift the focus back from SSAF to VAF to show that our results also hold for more realistic simulations of volcanic eruptions. For this purpose, we compare our simulations with a large ensemble of volcanically forced simulations (MPI EVA ensemble; Azoulay et al. 2021) and simulations from phase 6 of the Coupled Model Intercomparison Project (CMIP6; Eyring et al. 2016).

2. Concepts

a. Efficacy and feedback

We describe the increased feedback parameter of VAF following Hansen et al. (2005): VAF has been claimed to have lower than unity efficacy, where efficacy is the ratio of global mean near-surface air temperature change T per unit radiative forcing F of VAF and $2 \times CO_2$ forcing:

$$\text{Efficacy} = \frac{T/F}{T_{2\times \mathrm{CO}_2}/F_{2\times \mathrm{CO}_2}}. \tag{1}$$

A forcing agent produces less temperature change than CO_2 forcing of equal magnitude if the efficacy is lower than 1.

The concept of efficacy is closely related to the feedback parameter $\lambda = dN/dT$, where N is the top-of-the-atmosphere (TOA) flux anomaly. Forcing, feedback, and surface temperature are linked by the linearized equation $N = F + \lambda T$. A more negative feedback parameter λ indicates a more stable climate system, since for a strongly negative feedback less surface temperature change is required to offset a given radiative forcing. Therefore, the efficacy of a forcing in equilibrium can be computed as the ratio of its inverse feedback parameter to the inverse feedback parameter of a $2 \times CO_2$ forcing [see Eq. (1) herein and Zhao et al. (2021)]:

$$\text{Efficacy} = \lambda_{\mathrm{CO}_2}/\lambda. \tag{2}$$

Rugenstein and Armour (2021) discuss three closely related but distinct definitions of the feedback parameter, of which only two are relevant to this work: the equilibrium and the differential feedback parameter. For an illustration, see their Table 1.

The equilibrium feedback parameter is defined as the ratio of forcing and equilibrium temperature change and therefore compares two equilibrated states. The differential feedback parameter is obtained by a regression of $N(T)$ over a certain number of years. Its value depends on the chosen period for the regression and can differ more from the equilibrium feedback parameter the more that $N(T)$ deviates from a linear relationship.

TABLE 1. Efficacy of VAF in previous studies. Values marked with an asterisk are not directly reported in the source but rather are calculated by us on the basis of the reported feedback parameters. In the case of Boer et al. (2007), the $4 \times CO_2$ feedback parameter was taken from Andrews et al. (2015). For definitions of differential and equilibrium feedback parameter, see section 2 of Rugenstein and Armour (2021).

Source	Efficacy	Method
Hansen et al. (2005)	1.00	Differential
	0.88	Equilibrium
Marvel et al. (2016)	0.73; 5%–95% CI: (−0.61, 2.06)	Differential
Ceppi and Gregory (2019)	0.45^*	Differential
	0.71^*	Differential
Boer et al. (2007)	0.81^*	Differential
Gregory et al. (2016)	0.69 ± 0.09	Differential
Merlis et al. (2014)	0.84^*–0.94^*	Equilibrium
	0.69^*–0.92^*	Differential
Zhao et al. (2021)	0.66^* (low lat)–1.16^* (high lat)	Equilibrium
Modak et al. (2016)	<1	Differential
Gregory et al. (2020)	<1	Differential

While Hansen et al. (2005) define efficacy with respect to the equilibrium efficacy, we use the term in a broader sense as a ratio of feedback parameters in general. Efficacy can depend sensitively on the employed feedback parameter definition. We use the words *equilibrium efficacy* and *differential efficacy* to refer to efficacies calculated with Eq. (2), using the definitions of the equilibrium feedback parameter and the differential feedback parameter, respectively.

b. The pattern effect

Variations of the feedback parameter have been investigated not only with respect to the forcing agent, such as CO_2, VAF/SSAF, and so on, but also with respect to changes over time in CO_2 step forcing experiments (Senior and Mitchell 2000). For most climate models, a weakening of feedback over time has been reported (Andrews et al. 2012, 2015; Armour 2017). The TOA flux anomaly N has been found to depend not only on the global average T, but also on its pattern, which gives rise to the term "pattern effect" (Stevens et al. 2016). The pattern of SST anomalies is subject to changes over time (Armour et al. 2013; Zhou et al. 2016; Ceppi and Gregory 2019; Dong et al. 2019; Andrews et al. 2018) and therefore the feedback parameter changes with time, too. Indeed, the feedback parameter spread in CMIP5 and CMIP6 models can be explained by SST pattern differences (Dong et al. 2020). This has been linked to changes in cloud feedback and lapse-rate feedback (Andrews et al. 2015; Zhang et al. 2010; Stevens et al. 2016; Andrews and Webb 2018; Andrews et al. 2018; Ceppi and Gregory 2019), which depend on the spatial distribution of SST.

A useful framework to understand the pattern effect is provided by Armour et al. (2013), who suggest that some regions intrinsically exhibit stronger feedback than others, independent of time or state. Only the temperature change pattern varies in time and activates the local feedbacks differently.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

The global feedback is the temperature change-weighted average over the local feedbacks. Therefore, the global feedback gets weaker, as the temperature change gets relatively stronger in regions with weak local feedback (Rugenstein et al. 2016). Zhou et al. (2017) and Dong et al. (2019) expand on this view by using a Green's function approach, which accounts for the fact that local SST perturbations can modify feedbacks also in remote regions (Ceppi and Gregory 2017; Andrews and Webb 2018; Lin et al. 2019). In this work, we will refer mostly to this framework.

There is evidence that the strong feedback to VAF could be related to the specific temperature change pattern after volcanic eruptions. Gregory and Andrews (2016) propose that the variations of the differential feedback parameter during the historical period could be related to specific SST patterns following volcanic eruptions, among other reasons. The efficacy of VAF has been connected to changes in tropospheric stability, which arise from different SST patterns and lead to changes in lapse-rate and cloud feedback (Ceppi and Gregory 2017, 2019). Similarly, the efficacy of anthropogenic aerosol forcing has been interpreted with respect to changes in stability that arise from specific SST patterns (Salvi et al. 2022).

More generally, the feedback parameter was found to depend on the latitude of the radiative forcing. Forcing in higher latitudes produces weaker feedbacks (Hansen et al. 1997; Forster et al. 2000; Rose et al. 2014; Rose and Rayborn 2016; Rugenstein et al. 2016; Ceppi and Gregory 2019), possibly because of the resulting high-latitude temperature change pattern (Haugstad et al. 2017; Salvi et al. 2022). In a multimodel analysis Po-Chedley et al. (2018) derive a theoretical basis for a relationship between the meridional temperature change pattern and the global feedback parameter, which was already found by Soden and Held (2006). In the model of Po-Chedley et al. (2018) more tropical temperature change leads to stronger global mean lapse-rate and water-vapor feedback. More specifically, Dong et al. (2019) show that it is the warm-pool region of the tropical Indian Ocean and the tropical western Pacific Ocean that dominates the negative feedback. There, the strong convective coupling yields a strongly negative lapse-rate feedback due to the moist adiabatic lapse rate (Manabe and Wetherald 1975; Bintanja et al. 2012; Pithan and Mauritsen 2014; Andrews and Webb 2018), partially compensated by the positive water-vapor feedback. Together with strong negative cloud feedbacks, absent positive surface-albedo feedback, and strong negative Planck feedback due to warm background conditions, this region exhibits the most negative local feedback, and controls the global feedback strength. Changes in feedback parameter over time in abrupt CO_2 forcing simulations can be attributed to changes in the relative warming of the Indo-Pacific warm-pool region to the global warming (Dong et al. 2020).

3. Earlier studies indicate that the efficacy of VAF is lower than unity

There are several modeling studies that report the efficacy or the feedback parameter of VAF. All of them point toward a low efficacy of VAF. An overview is given in Table 1.

TABLE 2. Type and number of performed simulations.

	$0.5 \times CO_2$	$2 \times CO_2$	Idealized SSAF
Years 1–10, coupled	18	18	18
Years 1–1000, coupled	1	1	1
Years 1–30, fixed SST	1	1	1

No two author teams use exactly the same method. The results stem from different models, and some use time-constant forcings (Hansen et al. 2005; Gregory et al. 2016; Ceppi and Gregory 2019; Zhao et al. 2021) while others use transient forcings (Marvel et al. 2016; Boer et al. 2007). The differential feedback parameters are regressed over a different numbers of years, between two years (Ceppi and Gregory 2019) and the entire historical period (Marvel et al. 2016; Gregory et al. 2020). In two cases, the feedback parameter of CO_2 forcing is taken from a $4 \times CO_2$ simulation instead of the $2 \times CO_2$ simulation (Boer et al. 2007; Gregory et al. 2016).

Additional studies suggest or show a lower-than-unity efficacy without reporting a specific value (Modak et al. 2016; Gregory et al. 2020). We conclude from this literature review that VAF likely exhibits a low efficacy.

4. Methods

a. Model and experiments

We employ the coupled climate model MPI-ESM 1.2 in its low-resolution version. It consists of the atmosphere component ECHAM6 (192 × 96 grid points horizontally and 47 vertical levels; Stevens et al. 2013), coupled to the ocean component MPI-OM (256 × 220 grid points horizontally and 40 vertical levels; Jungclaus et al. 2013). Furthermore, land processes and ocean biogeochemistry are simulated in the Jena Scheme for Biosphere–Atmosphere Coupling in Hamburg (JSBACH; Reick et al. 2021) and the Hamburg ocean carbon cycle model (HAMOCC; Ilyina et al. 2013), respectively. Since no interactive atmospheric chemistry processes are included, aerosols and trace gases are prescribed with monthly fields.

Based on a control simulation with preindustrial conditions (piControl), three sets of simulations are performed (Table 2). For each forcing type ($0.5 \times CO_2$, $2 \times CO_2$, and idealized SSAF) we prescribe a step-like radiative forcing and perform 18 simulations of the first 10 years, one simulation of 1000 years, and one simulation of 30 years with fixed SST.

The aerosol optical properties for the idealized SSAF are calculated with the EVA forcing generator (Toohey et al. 2016). The monthly and zonal mean fields of aerosol extinction, single scattering albedo, and the asymmetry factor are precomputed offline for a sulfur injection of 20 Tg, once for a January eruption and once for a July eruption. We then shift the July eruption by 6 months to phase match it with the January eruption and then compute the average of both. This eliminates the seasonal transport asymmetry but retains a realistic annual-average sulfate transport toward the poles. We then average over the first three post-eruption years and prescribe this profile to MPI-ESM. It is time-constant by construction, but representative of the time-averaged forcing structure

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

after a volcanic eruption. The sulfur mass is chosen so that the time-averaged global-mean effective forcing is approximately the same as in the 0.5 × CO_2 experiment to preclude forcing strength-dependent effects on the feedback (Rohrschneider et al. 2019).

The simulations with fixed SST are only used to calculate the effective radiative forcing. The 18-member ensembles of the first decade are intended to reduce the uncertainty in regressions over the early phase of the temperature change, where there are only few data points.

b. *Calculation of effective forcing, feedback parameter, and temperature change pattern*

We calculate the effective forcing as the average TOA flux imbalance of the simulations with fixed climatological SST (Forster and Taylor 2006). Although the SST are fixed, there is nonnegligible temperature change over land. Therefore, we correct the effective forcing by subtracting λT in each year from the TOA flux imbalance, where λ is the differential feedback parameter of the first decade from the coupled runs (see section 4).

The equilibrium feedback parameters are calculated as described in section 2. Since the simulations still deviate from equilibrium by about 0.3 W m^{-2} K^{-1} after 1000 years, we extrapolate $N(T)$ to zero based on a linear fit over the years 100–1000. The intercept with the T axis yields the equilibrium surface temperature change. This choice is roughly equivalent to Rugenstein and Armour's (2021) method M5, which they apply to estimate the equilibrium temperature change in almost equilibrated models. Their method of extrapolating the years 100–400 from $N(T)$ to $N = 0$ underestimates the equilibrium T in our simulations, likely because of the pronounced curvature of $N(T)$ in our cooling simulations.

In addition, we obtain differential feedback parameters by regressing over the first 10 years (early period) or years 100–400 (late period). The latter choice is based on the recommendations by Rugenstein and Armour (2021). The results are not sensitive to the exact choice of boundary years for the periods. Derived quantities for the first 10 years (e.g., feedback parameter and temperature change pattern) are calculated from the ensemble average, as opposed to first calculating them individually for each ensemble member and averaging thereafter following the recommendations of Gregory et al. (2020). The results are the same to within a few percent when the order of regression and averaging is reversed.

Analogous to the differential feedback parameter, differential temperature change patterns are calculated by regressing the local against the global average change in T. Based on the findings of Dong et al. (2019, 2020) we define the warm-pool index (WPI) as the regression slope of the temperature change in the tropical warm-pool region (30°S–30°N, 50°E–160°W) against the global surface temperature change, similar to the ratio of tropical to global temperature change of Soden and Held (2006).

All regressions are performed over annual-mean values. The method to compute errors of the feedback parameters depends on the feedback parameter definition. The uncertainty of the equilibrium feedback parameter is equal to the uncertainty of F/T_{eq}, propagated from the standard deviations of F and T_{eq}. The differential feedback parameters are computed from the slope of $N(T)$, and therefore their uncertainties are described by the least squares regression slope's standard deviation.

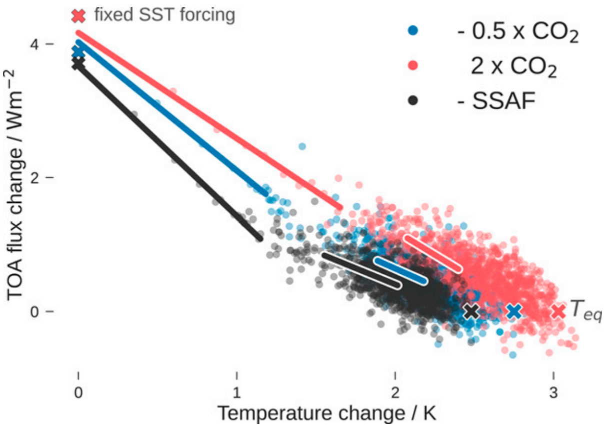

FIG. 1. The $N(T)$ (Gregory) plots. All values from 0.5 × CO_2 and SSAF experiments are multiplied by −1 so as to show them in the same quadrant as the 2 × CO_2 results. Linear regressions are shown for the early period (years 1–10) and the late period (years 100–400). Crosses mark the fixed SST effective forcings and the equilibrium temperatures, which were extrapolated from a linear regression of $N(T)$ over years 100–1000. The first 10 points of each simulation type are the ensemble averages from the 18-member ensembles of the first decade. The slope of each regression line is the differential feedback parameter.

5. Results

a. *Cooling is characterized by strong early feedback*

As pointed out in section 3, other authors have found lower-than-unity efficacies for VAF, but not under a uniform procedure. To emphasize the importance of the methodological approach, we show $N(T)$ plots (Fig. 1), the global equilibrium and differential feedback parameters (Fig. 2, left panel), and the corresponding efficacies from our simulations (Fig. 2, right panel).

The effective radiative forcing magnitude for CO_2 halving is approximately 15% lower than for CO_2 doubling, in line with the previously reported 10% from Chalmers et al. (2022). All quantities we show are normalized by effective forcing or temperature change, so that this does not compromise our findings.

The equilibrium feedback parameter is very similar in all three simulations. In equilibrium, SSAF produces only slightly less temperature change per unit forcing than a halving or doubling of CO_2. Differences between the three forcing agents are pronounced in the early differential feedback parameter, which is obtained from an ensemble average of the regression slopes of $N(T)$ over the first 10 years of each ensemble member. This method is closest to what most authors did to calculate the efficacy of VAF (see section 3). We find a more negative feedback parameter in both cooling cases for the

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

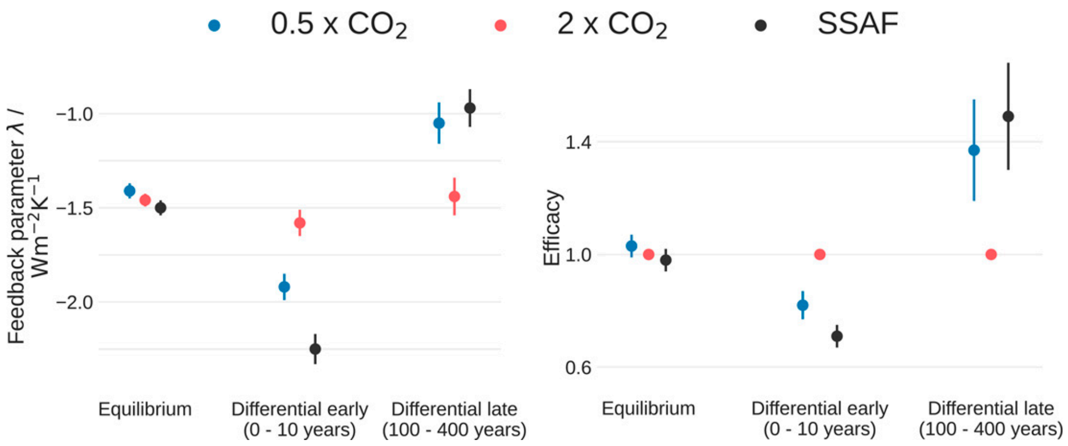

FIG. 2. (left) Feedback parameter and (right) efficacy, calculated from the equilibrium method and the differential method for the early and late period. The efficacy of 2 × CO_2 simulations is 1 by definition.

early period. By contrast, the late regression feedback parameter (years 100–400) is less negative for the cooling cases than for the 2 × CO_2 simulation.

These findings suggest the following picture: Relative to the 2 × CO_2 warming case, the cooling simulations (0.5 × CO_2 and SSAF) exhibit stronger feedback in the first decades. Therefore the efficacy is low in transient simulations of VAF, as previously reported (section 3). Later on, the feedbacks weaken and become even weaker than in the warming case. Since most authors focus on one specific time period for regressing the differential feedback parameter, this change in feedback strength has not been reported before for cooling simulations. All in all, the early stronger and the late weaker feedback almost compensate, so that the equilibrium feedback parameters of SSAF, CO_2 -induced cooling, and warming are almost equal (i.e., the equilibrium efficacy is approximately 1). Efficacy differences exist in transient states, but not in equilibrium.

Note that also in the 2 × CO_2 simulation the feedback parameter decreases over time, albeit less than in the other two simulations. The weakening of the feedback of 2 × CO_2 simulations is a well-known feature of the majority of CMIP5 and CMIP6 models (Senior and Mitchell 2000; Andrews et al. 2012; Winton et al. 2010; Armour et al. 2013; Ceppi and Gregory 2017; Armour 2017) and explained by pattern effects (see section 2).

The feedback parameter change over time is more pronounced in our cooling simulations than in the warming simulations and is largest for SSAF. Two questions arise immediately:

1) Why do the cooling simulations produce initially stronger global mean feedback than the warming simulation?
2) Which processes change over time that lead to the enhanced weakening of the feedback parameter in the cooling simulations?

In the following we will argue that these questions can be answered on grounds of different temperature change patterns and their temporal evolution.

b. *The temperature change pattern explains feedback differences*

The temperature change patterns are shown as zonal averages in Fig. 3, and spatially resolved in Fig. 4. All patterns show some common characteristics in the early period: amplified temperature change in the Arctic, intermediate temperature change in the tropics, and small temperature change in the southern mid and high latitudes, except close to the South Pole. The qualitative temperature change pattern with reduced Southern Ocean and Antarctic temperature response after volcanic eruptions was found before by Yang et al. (2019) and confirmed in a multimodel large ensemble comparison (Pauling et al. 2021). In the late period, the pattern becomes more El Niño–like, consistent with the long-term temperature change pattern in other models (Cai and Whetton 2001; Held et al. 2010).

Despite the generally similar shape of the patterns, there are important differences: While the zonal mean temperature change is smaller than or equal to the global average at almost

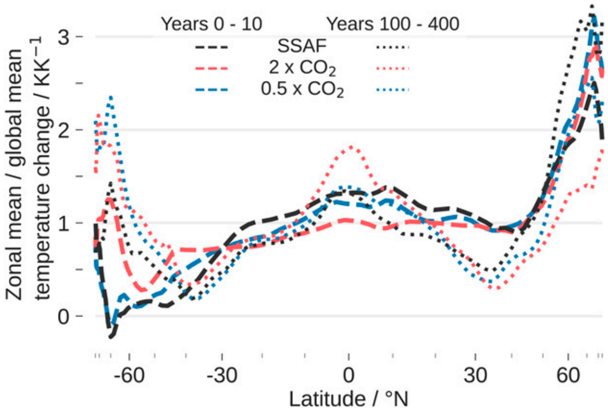

FIG. 3. Zonal mean temperature change pattern for the early (years 0–10) and late (years 100–400) periods. Values greater than 1 indicate stronger temperature change than the global average. A value of 0 indicates no temperature change.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

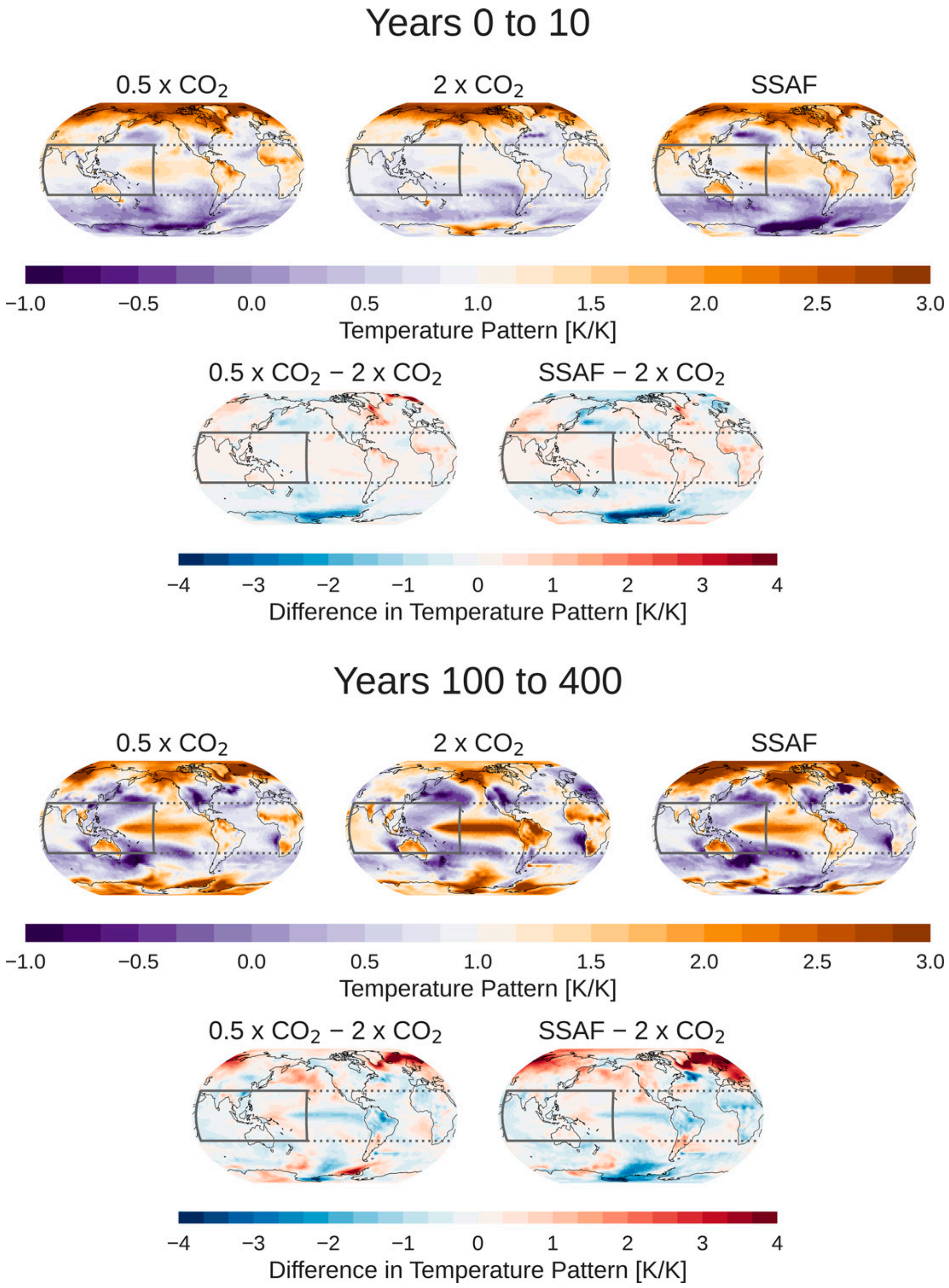

FIG. 4. Spatially resolved temperature change pattern for (top) the early period (years 0–10), and (bottom) the late period (years 100–400). The lower two panels of each subfigure show the difference to the 2 × CO_2 simulation. Values greater than 1 indicate stronger temperature change than the global average. A value of 0 indicates no temperature change. The solid and dotted lines indicate the boundaries of the warm-pool region and the tropics, respectively.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

all latitudes in the tropics in the first decade of the 2 × CO_2 simulation, the opposite is true in both cooling cases. This picture is reversed over time: In the late period, the deep tropical temperature change is enhanced relative to the global mean for the 2 × CO_2 simulation (see Fig. 3; difference between dashed and dotted line between 20°S and 20°N). The opposite is true for the cooling simulations, which have a lower temperature change in the late period than in the early period in most tropical latitudes, when compared with the global average. During the late period, in contrast with the early period, the southern high latitudes show much larger temperature changes in both cooling experiments, but especially in the 0.5 × CO_2 case.

We propose that the global feedback is strongly negative in the cooling simulations in the early period, because the temperature change is concentrated in the tropical warm-pool region (30°S–30°N, 50°E–160°W). In this region the combined lapse-rate and water-vapor feedback is stronger than in the global average (Po-Chedley et al. 2018; Dong et al. 2019, 2020), and the surface-albedo feedback is zero. Using the radiative kernels from Block and Mauritsen (2013), we perform a kernel decomposition of the total feedback into its components (Soden et al. 2008). This yields that lapse-rate, water-vapor, and surface-albedo feedback together account for 85% of the feedback differences between 2 × CO_2 and SSAF radiative forcing in the early period (not shown). The same feedback processes cause 100% of the feedback differences between 2 × CO_2 and 0.5 × CO_2, where additional small differences in cloud and Planck feedback compensate. In contrast to the multimodel comparison of Dong et al. (2019), in which the cloud feedback plays an important role for feedback parameter differences, it is approximately zero in all our simulations during the early period.

The most striking differences in the temperature patterns are not located in the warm-pool region, but rather at higher latitudes. However, temperature changes in the warm pool have much larger global radiative effects than changes in any other region, such that even moderate differences in the warm-pool region can dominate the global response (Dong et al. 2019, their Fig. 5). For this reason, we focus on the warm-pool region, although the temperature pattern differences there are only moderate.

We show the relationship between the temperature change pattern and the feedback parameter in Fig. 5, where the WPI is a simple measure for the temperature change pattern. A WPI greater than 1 means that the tropical warm-pool region warms or cools more than the global average. The feedback parameter is more negative in simulations with enhanced temperature change in the tropical warm-pool region. The scatter points of SSAF and 2 × CO_2 simulations are well separated in the early period. The negative correlation between WPI and feedback parameter exists not only for the three simulation types together, but also for each of the three ensembles individually.

As time passes, the temperature change pattern in the SSAF and 0.5 × CO_2 simulations is shifted away from the warm-pool region, and more generally from the tropics toward the extratropics. This enhances the less negative lapse-rate feedback of the non–warm-pool regions and the positive surface-albedo feedback of the high latitudes. For this reason, the overall feedback is initially stronger when cooling and becomes weaker later on.

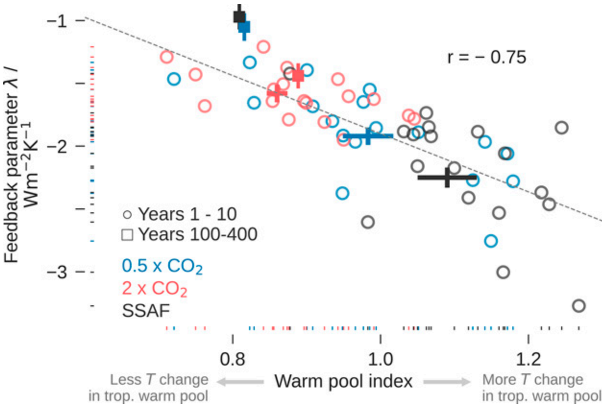

FIG. 5. Scatterplot of feedback parameter vs WPI (ratio of temperature change in the tropical warm-pool region to global mean temperature change). Each circle represents one ensemble member from the early period; crosses show ensemble mean ± standard error; squares and error bars mark values from the late period (years 100–400) ± standard error. The gray dashed line is a linear regression through all 3 × 18 = 54 ensemble members from the first decade. Marginal distributions are plotted on the axes.

However, the scatter points from the late period (squares in Fig. 5) do not fall on the regression line from the early period. Apparently, the simple WPI cannot explain completely the shift from strong to weak feedback between the decadal and centennial time scale. This shift has been found to be related to the temperature change pattern over ascent and descent regions within the tropics (Dong et al. 2019, 2020), and to the delayed southern ocean temperature change (Senior and Mitchell 2000). Additionally, the relationship between local SST and global radiation changes might not be invariant to climate change, which could alter the relationship between WPI and feedback parameter in warmer or cooler climates.

Although the WPI cannot explain the change in feedback from the decadal to the centennial time scale entirely, it does explain the efficacy differences between simulations within each period. While we point out the importance of the tropical warm-pool region for efficacy differences between 0.5 × CO_2 forcing, 2 × CO_2 forcing, and SSAF, we acknowledge that temperature change in other regions also influences the feedback. Interestingly, the cloud feedback of all our simulations is approximately the same in the early period (0.0–0.1 ± 0.1 W m^{-2} K^{-1}), although cloud feedbacks are the most important contribution to the differences between early and late feedback of CMIP5 and CMIP6 4 × CO_2 simulations (Dong et al. 2020), including MPI-ESM (Block and Mauritsen 2013). Either MPI-ESM is an outlier in this regard, or the processes that lead to differences between early and late feedback are not exactly the same processes that distinguish SSAF, 0.5 × CO_2, and 2 × CO_2 forcing in the early period.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

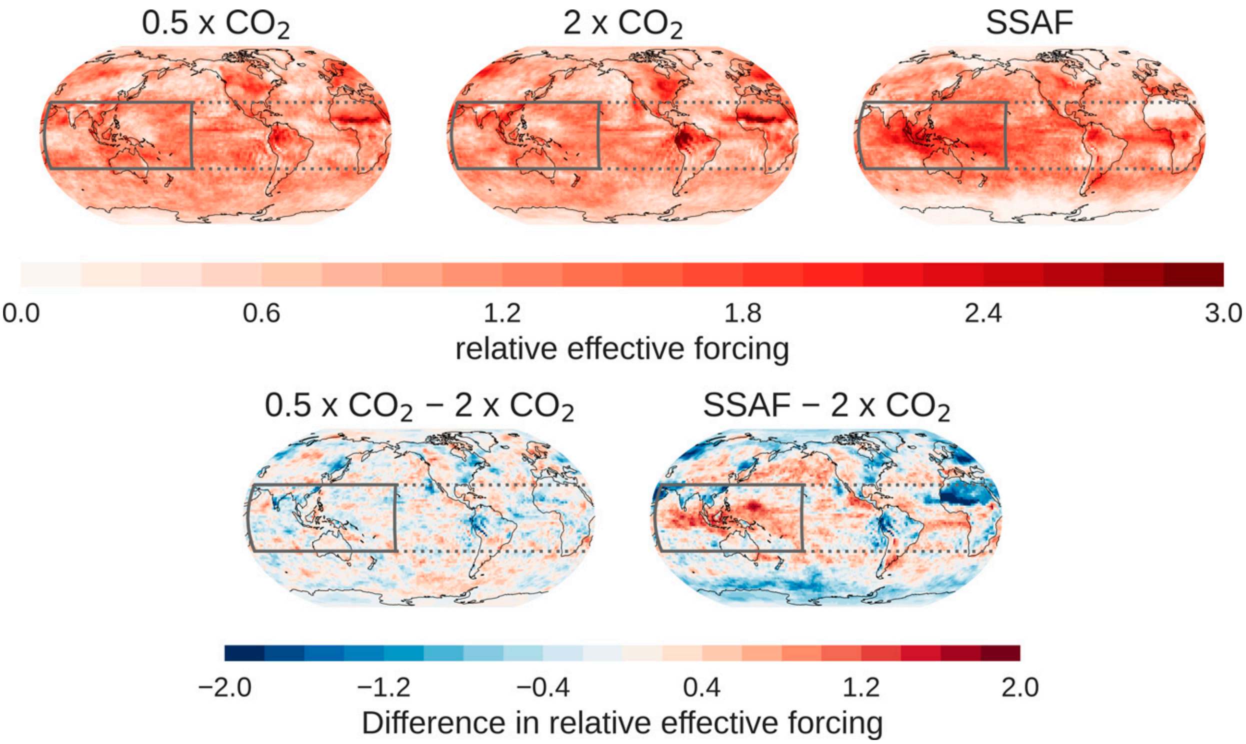

FIG. 6. (top) Relative effective forcing (effective forcing normalized by its global average). (bottom) The difference to the $2 \times CO_2$ simulation. The solid and dotted lines indicate the boundaries of the warm-pool region and the tropics, respectively.

Repeating the analysis including the whole tropics instead of only the warm-pool region, [i.e., using the “ratio of tropical to global warming” from Soden and Held (2006); Po-Chedley et al. 2018] yields similar results, albeit with weaker correlations (Fig. S1 in the online supplemental material). The conclusions are the same for both variants. While there is a theoretical foundation for the link between the ratio of tropical to global warming and the feedback parameter (Po-Chedley et al. 2018), this relation might actually originate from the dominance of the tropical warm-pool region. Of course, WPI and the ratio of tropical to global warming are correlated.

SST pattern changes can lead to changes in large-scale tropospheric stability (Ceppi and Gregory 2019). They can be quantified by the differential stability change dS/dT, where S is the area-averaged estimated inversion strength (Wood and Bretherton 2006) between 50°S and 50°N. Increased stability leads to more low clouds in the tropical subsidence regions and a more-negative lapse-rate feedback (Ceppi and Gregory 2019). Exchanging the WPI for dS/dT in Fig. 5 yields a qualitatively similar result (Fig. S2 in the online supplemental material). The SST pattern measure WPI and the stability measure dS/dT are well correlated (correlation coefficient $r = 0.81$; Fig. S3 in the online supplemental material), which corroborates the relationship between changes in SST patterns, stability, and feedback (Ceppi and Gregory 2019; Salvi et al. 2022).

c. *Sign, pattern of radiative forcing, and stratospheric heating cause temperature change pattern differences*

While we have shown that the differences in temperature change patterns explain the differences between the feedback parameters within the same period, it is not clear what causes the former. We propose that three factors combine to elicit different WPI values.

One logical explanation would be that the WPI differences are caused by different radiative forcing patterns, which we show in Fig. 6. The high aerosol concentration and the strong insolation in the tropics cause the SSAF to be stronger in the tropics than in the extratropics. Specifically, the warm-pool region experiences stronger relative effective forcing from aerosols than from either doubling or halving the CO_2 concentration. Since this pattern only weakly appears in the instantaneous forcing, the warm-pool-enhanced effective forcing pattern must largely originate from atmospheric adjustments (not shown). It could lead to more temperature change in the tropical warm-pool region. However, the two simulations with altered CO_2 concentrations have almost the same forcing pattern (see Fig. 6), yet their temperature change patterns are very different (see Figs. 3 and 4), and so are the feedback parameters (see Figs. 2 and 5). Hence, the forcing pattern alone is not sufficient.

Another conceivable explanation is that the temperature change pattern differences arise from the distinction between negative and positive radiative forcing. Ocean heat transport

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

is projected to decrease with warming (Li et al. 2013; Previdi et al. 2021), which leads to more negative radiative feedbacks (Singh et al. 2022), consistent with our interpretation of the decreased activation of more positive extratropical feedbacks. If the poleward ocean heat transport strengthens in cooler conditions, that would explain why the temperature change shifts from the tropical warm-pool region to the extratropics on the centennial time scale. The reduced ocean heat transport in warmer conditions could be the reason that the temperature change pattern, as measured by the WPI, remains relatively constant in the 2 × CO_2 simulation. Furthermore, warming stabilizes the ocean and therefore suppresses vertical mixing, whereas cooling leads to destabilization. This might cause differences in mixing time scales, surface fluxes, and consequentially SST patterns. The enhanced mixing from cooling would enhance the relative importance of ocean heat uptake, when compared with radiative feedbacks. Ocean heat uptake has been interpreted as a forcing with weak feedback (Winton et al. 2010; Rose et al. 2014). Therefore a cooling/warming asymmetry of the relative role of ocean heat uptake could affect the global mean feedback. We present no evidence corroborating or rejecting this hypothesis, but suggest it as a possible reason for a cooling versus warming feedback asymmetry. Still, if cooling versus warming were the only important distinction, the 0.5 × CO_2 and SSAF ensemble means would not be separated in Fig. 5.

Third, the altitude of the aerosol layer has been shown to influence the efficacy through stratospheric heating and water-vapor feedback (Zhao et al. 2021). Therefore the heating by absorption of LW radiation influences the feedback parameter, but our results show that this process is mediated by the temperature change pattern. This would reconcile the findings of Zhao et al. (2021) with those of Gregory and Andrews (2016), Shindell et al. (2015), and Haugstad et al. (2017), who show that the surface temperature change pattern and not the existence/nonexistence of aerosol determines the feedback strength.

Our results do not allow us to conclude what exactly causes the distinct temperature change patterns, but it is likely a combination of the sign (possibly also the magnitude) of the radiative forcing, its pattern, and the stratospheric heating from the aerosols. None of these factors is sufficient to explain the differences on its own.

6. EVA ensemble and CMIP6 simulations confirm increased warm-pool temperature change

a. *EVA ensemble*

Our SSAF simulations are highly idealized: they are based on time-independent radiative forcing and do not reflect the seasonally varying transport of stratospheric aerosol. To verify if the enhanced temperature change in the tropical warm-pool region also holds for time-dependent VAF, we compare them with the MPI "idealized volcanic forcing ensemble" (EVA-ENS). Azoulay et al. (2021) created 100-member ensembles of idealized equatorial volcanic eruption with sulfur injections of 2.5 to 40 Tg. For comparison, it has been estimated that the Mount Pinatubo eruption injected approximately 5–10 Tg of sulfur (TgS; Timmreck et al. 2018) into the stratosphere. Their aerosol optical depth distributions are, as in our case, obtained from the EVA forcing generator. The simulations of EVA-ENS assume an eruption in June 1991, and are branched off of 100 different realizations of the MPI Grand Ensemble of historical simulations (Maher et al. 2019). We determine the temperature change pattern in years 1991–93 and compute the WPI. The results are shown in Fig. 7. For low sulfur injections of 5 Tg or less, the standard error is very large and no reliable statement can be made on the spatial structure of the temperature change. For injections of 10 TgS or more, the temperature change is robustly concentrated in the tropical warm-pool region. After the idealized volcanic eruptions with injections greater than 5 TgS the WPI is approximately between 1.2 and 1.5. This is higher than the values from the idealized SSAF simulation (1.1), which could be related to the even shorter time scale of 3 years for the simulations of EVA-ENS, as compared with the time scale of 10 years of the idealized SSAF simulations.

b. *Volcanic eruptions in CMIP6 historical simulations*

For a multimodel comparison we examine the CMIP6 "historical" simulations and warming signals from the "abrupt 4 × CO_2" simulations. We use the years 1883–85 and 1991–93 from the CMIP6 historical simulations for VAF, because in these periods the radiative forcings from the eruptions of Krakatau in May–August 1883 and Mount Pinatubo in June 1991 dominate the total effective forcing from all sources. The signal-to-noise ratio is still poor, and the average simulated annual mean global mean cooling amounts to only 0.3 K for Krakatau and 0.2 K for Mount Pinatubo. This is on the order of internal variability, so that individual realizations may exhibit only small negative or even positive temperature anomalies. We compute the WPI for all ensemble members from the participating models with at least 10 historical realizations. After removing three outliers with a WPI $\gg 6$, we compare these WPI values with those from the first three years of the abrupt 4 × CO_2 simulations. The results for each model and for the whole ensemble are shown in Fig. 7. Figures showing results for each individual simulation from all models (including those with less than 10 realizations) are provided in Figs. S4 and S5 in the online supplemental material.

There are multiple sensible possibilities of averaging the results: for example, giving each realization the same weight (realization average) or giving each model the same weight (model average). Both methods yield qualitatively the same result: the temperature change pattern is more concentrated in the tropical warm-pool region in the periods of historical VAF than in the 4 × CO_2 simulations.

However, the WPI spread is large for the Pinatubo and Krakatau periods of the historical simulations, both within models and between models. Most models agree that the WPI is larger in periods of VAF, especially those with more realizations, but not all of them do. Many realizations can be

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

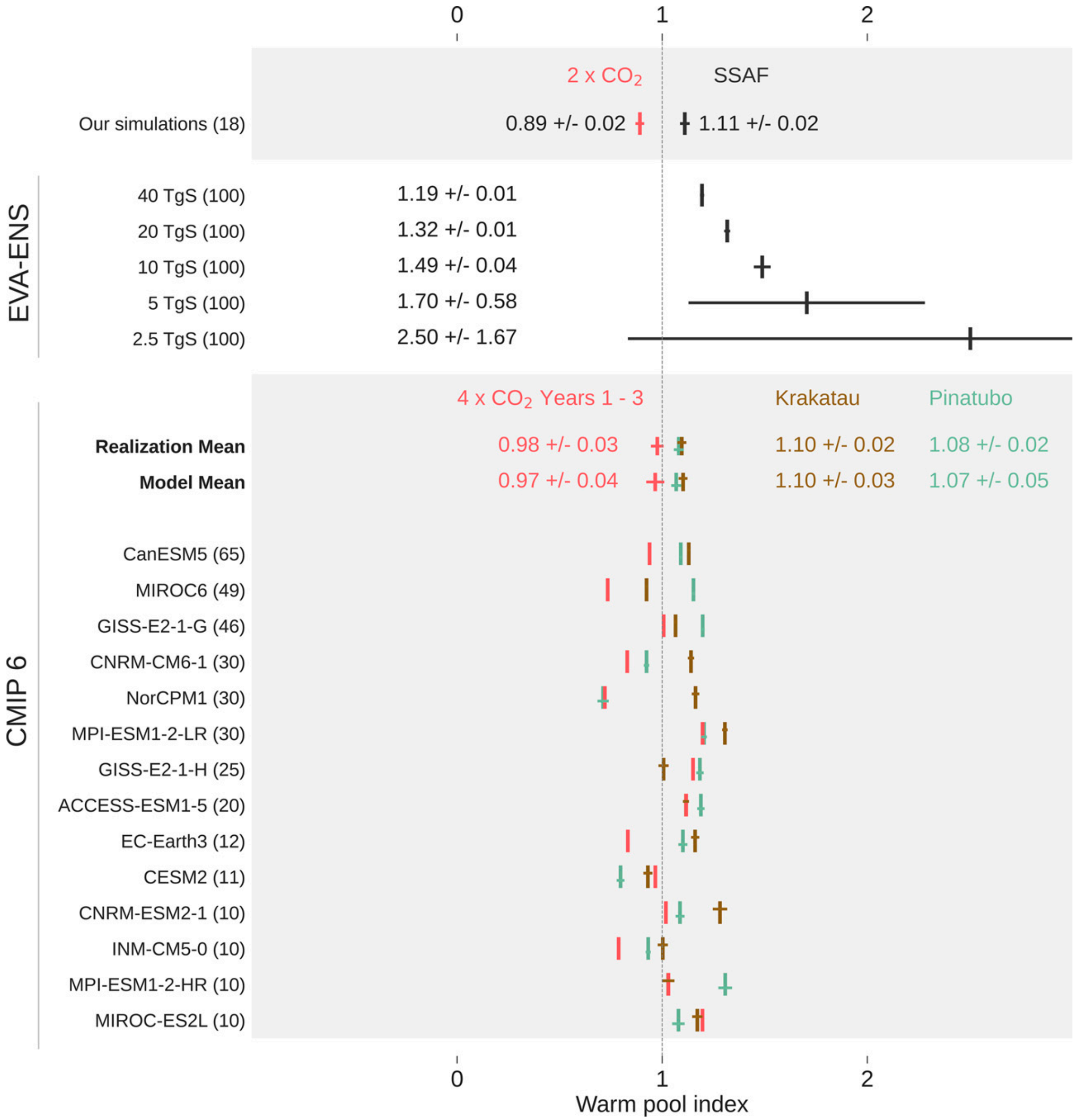

FIG. 7. Comparison of the WPI in our 18-member, 10-yr ensemble of idealized SSAF simulations, simulations from EVA-ENS, and volcanic periods of the CMIP6 historical simulations. Numbers in parentheses indicate the number of ensemble members; models with less than 10 members are not shown. Error bars represent standard errors and are slightly shifted up and down to avoid indistinguishable overlaps. For 4 × CO_2 there are generally only one or a few realizations per model. The error bar of the 2.5-TgS EVA-ENS simulations exceeds the figure limits.

found with higher WPI in the abrupt 4 × CO_2 simulations than in the historical simulations during periods of VAF. If the variability of the CMIP6 simulations is representative of the real world, observed temperature change patterns and therefore also observed feedback parameters and efficacies for any individual eruption might substantially deviate from the expected value. The high WPI of VAF is robust in the multimodel ensemble mean, and in some but not all single-model ensemble means. Even in our own ensemble of highly idealized forcing scenarios some realizations exhibit a WPI, which is substantially different from the ensemble mean, causing an overlap of the WPI values of SSAF and 2 × CO_2 simulations (but no overlap of the estimated uncertainties; see Fig. 5).

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

TABLE 3. Early WPI and feedback parameter for the simulations from CFMIP. MRI-ESM2-0 did not provide the necessary output to compute the feedback parameter. GISS-E2-1-G provides two realizations of the 2 × CO_2 experiment.

	0.5 × CO_2		2 × CO_2	
	WPI	λ (W m^{-2} K^{-1})	WPI	λ (W m^{-2} K^{-1})
MRI-ESM2-0	0.87	—	0.65	—
CNRM-CM6-1	0.75	−1.10	0.92	−1.03
IPSL-CM6A-LR	0.88	−1.16	0.82	−1.40
GISS-E2-1-G	0.79	−1.66	0.92; 0.93	−1.04; −1.60
MIROC6	0.84	−1.89	0.94	−1.61

The comparison with the CMIP historical simulations and the EVA ensemble shows that the increased WPI of SSAF in the early period is also found in simulations of VAF.

c. 0.5 × CO_2 simulations from CFMIP

To verify our results from the 0.5 × CO_2 simulations, we analyze a set of five models that performed 0.5 × CO_2 and 2 × CO_2 simulations in the framework of the Cloud Feedback Model Intercomparison Project phase 3 (CFMIP3; Webb et al. 2017). Since the CFMIP simulations span only 150 years, the feedback parameter can only be computed for the early period. Three of four models that provide the necessary output to compute the feedback parameter show at least a moderately stronger feedback in the 0.5 × CO_2 simulations than in the 2 × CO_2 simulations (Table 3). However, there is no clear relationship between WPI and feedback parameter. While this seems to indicate that our results are not supported by the CFMIP3 ensemble, we stress that these are only five models with only one realization per model (two 2 × CO_2 simulations for GISS-E2-1-G) of relatively weak forcings. In our 18-member ensembles, the ranges of feedback parameters are large and there is an overlap between the ensembles of cooling and warming. A larger ensemble from CFMIP would be necessary to conclude this more definitely. If indeed MPI-ESM is an outlier in that regard, then the conclusions on the importance of the sign (or magnitude) of the radiative forcing pattern might be wrong. Mitevski et al. (2022) found no feedback differences between 0.5 × CO_2 forcing and 2 × CO_2, although this might be explained by the fact that they examine the differential feedback parameter of years 1–150, which is exactly between our early period (years 1–10, characterized by strong feedback) and our late period (years 100–400, characterized by weak feedback). On the other hand, our results match those from Chalmers et al. (2022), who also report stronger feedbacks in 0.5 × CO_2 than in 2 × CO_2 simulations, although they find an elevated importance of cloud feedbacks in addition to lapse-rate and surface-albedo feedbacks.

7. Discussion and summary

There is a broad consensus in the literature that VAF is characterized as producing stronger feedback than 2 × CO_2 forcing (i.e., low efficacy). Using idealized simulations of 0.5 × CO_2 forcing, 2 × CO_2 forcing, and step-like SSAF, we tackle the question why the feedback parameter appears to be stronger for VAF.

The finding that the feedback to SSAF is stronger than that of 2 × CO_2 forcing holds only for differential feedback parameters in the early decade, whereas on the centennial time scale the feedback weakens considerably. This weakening is more pronounced in the 0.5 × CO_2 and SSAF simulations than in the 2 × CO_2 simulation. In equilibrium, there are almost no efficacy differences between SSAF, CO_2 warming, and CO_2 cooling, due to a compensation of early strong and late weak cooling feedbacks.

This pronounced change in feedback over time might have implications for climate engineering by solar radiation management. Because of the low early efficacy a comparably large negative radiative forcing would be necessary to balance the surface temperature change from CO_2 on the decadal time scale. On the centennial time scale the efficacy of SSAF is larger than unity. Therefore a particular CO_2 radiative forcing could be balanced by a weaker (in magnitude) negative radiative forcing from aerosols in the long term. However, it is not clear how our results might change in the presence of strong CO_2 radiative forcing. Since part of the pronounced changes in the efficacy of aerosol forcing seems to be related to the fact that Earth is cooling, this would not be observed in a world of approximately constant or slightly rising mean temperature.

The responses to cooling and warming differ and there are substantive feedback changes over time. Hence it is not possible, or at least not straightforward, to estimate the equilibrium climate sensitivity to CO_2 forcing from the observed response to volcanic eruptions. This has been argued before, but partly on different grounds (Wigley et al. 2005; Stowasser et al. 2006; Boer et al. 2007; Gregory et al. 2020; Kummer and Dessler 2014; Merlis et al. 2014). Feedback differences were often interpreted to originate from the SW/LW nature of the radiative forcing (Joshi et al. 2003; Bony et al. 2006) or from differences in ocean heat uptake (Stowasser et al. 2006; Boer et al. 2007), but less from its sign.

Changes of the feedback parameter in time are related to varying temperature change patterns. The temperature change pattern causes the differences in feedback strength between warming and cooling simulations, which explains the early low efficacy of SSAF. In comparison with a doubling of the CO_2 concentration, a halving of the CO_2 concentration—and even more so a cooling with SSAF—lead to stronger temperature changes in the tropical warm-pool region in the first decade, relative to the global mean. The slightly enhanced temperature change in the warm-pool region substantially increases near-global stability S and strengthens the global feedback parameter.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

The differences between cooling and warming simulations in the first decade mainly originate from lapse-rate, water-vapor, and surface-albedo feedback. In the cooling simulations, the temperature change progresses to the high southern latitudes on the decadal to centennial time scale, which leads to an activation of the less negative/more positive high-latitude feedbacks. Therefore, the global feedback weakens more in the cooling simulations than in the 2 × CO_2 simulation. In our simulations, this effect is common to CO_2-induced and SSAF-induced cooling, but more pronounced in the SSAF case. This highlights that the radiative forcing's pattern, its sign (possibly also the magnitude), and the stratospheric heating from the aerosols (Zhao et al. 2021) likely contribute to setting the temperature change pattern. The atmospheric and oceanic circulations blur differences in the forcing patterns when translating them to a temperature change pattern.

A comparison with effects of the Mount Pinatubo and Krakatau eruptions in the CMIP6 historical simulations shows that the enhanced temperature change in the tropical warm-pool region is an average feature of climate models after volcanic eruptions, although the spread is large. Moreover, we find this characteristic in the EVA ensemble, a large ensemble of simulations of idealized volcanic eruptions in MPI-ESM 1.2-LR.

Our simulations and the eruptions of Krakatau and Pinatubo in the CMIP6 historical simulations are characterized by tropically concentrated SSAF. Can the results be transferred to solar forcing and extratropical SSAF?

Other studies have found the efficacy of solar forcings to be lower than unity (Schmidt et al. 2012; Modak et al. 2016; Hansen et al. 2005), specifically when cooling. We speculate that this might be explained by the fact that solar forcing predominantly affects the tropics, and therefore leads to a tropically enhanced temperature change pattern, similar to SSAF. It is furthermore possible that the forcing pattern helps explain nonunity efficacies of other forcing agents (Hansen et al. 2005; Ceppi and Gregory 2019). For *anthropogenic* aerosol forcing, the link between an extratropically concentrated radiative forcing pattern and *high* efficacy has already been demonstrated (Salvi et al. 2022).

Zhao et al. (2021) found that extratropically concentrated volcanic aerosols lead to a less negative equilibrium feedback parameter than in the case of mostly tropical aerosol load. Therefore, the finding of low efficacy from SSAF might not hold for eruptions at higher latitudes. Extratropical eruptions likely elicit a temperature change pattern, which is shifted toward the extratropics, and therefore cause weaker feedback. In that case, the efficacy could increase and be on the order of or even larger than 1. While differences in LW feedbacks are also important in the simulations of Zhao et al. (2021), the efficacy is mostly explained by differences in clear-sky and cloudy-sky SW feedbacks, suggesting an elevated importance of SW-cloud and surface-albedo feedbacks.

The exact reason why the temperature patterns of cooling and warming differ remains open. In our simulations, they involve different degrees of polar amplification. The mechanisms that drive polar amplification include lapse-rate feedback, ice-albedo feedback (Pithan and Mauritsen 2014), and a changing balance of moist versus dry static poleward energy transport (Alexeev et al. 2005; Hahn et al. 2021; Armour et al. 2019). An explanation for the differences in WPI between the radiative forcings and its changes in time could be based on these processes, but is beyond the scope of this work. The physical origin of the differences between the transient temperature change patterns of warming and cooling should be the focus of further research.

Acknowledgments. This work was supported by the Deutsche Forschungsgemeinschaft Research Unit VolImpact (FOR2820, Grant 398006378) within the project VolDyn (MG) and VolClim (CT). The authors are grateful to Jiawei Bao, who gave valuable comments on an earlier version of the paper. This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project bb1093. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

Data availability statement. All primary model output and scripts used in the analysis are available online (https://www.wdc-climate.de/ui/entry?acronym=DKRZ_LTA_066_ds00004).

REFERENCES

Alexeev, V. A., P. L. Langen, and J. R. Bates, 2005: Polar amplification of surface warming on an aquaplanet in "ghost forcing" experiments without sea ice feedbacks. *Climate Dyn.*, **24**, 655–666, https://doi.org/10.1007/s00382-005-0018-3.

Andrews, T., and M. J. Webb, 2018: The dependence of global cloud and lapse rate feedbacks on the spatial structure of tropical Pacific warming. *J. Climate*, **31**, 641–654, https://doi.org/10.1175/JCLI-D-17-0087.1.

——, J. M. Gregory, M. J. Webb, and K. E. Taylor, 2012: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere–ocean climate models. *Geophys. Res. Lett.*, **39**, L09712, https://doi.org/10.1029/2012GL051607.

——, ——, and ——, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Climate*, **28**, 1630–1648, https://doi.org/10.1175/JCLI-D-14-00545.1.

——, and Coauthors, 2018: Accounting for changing temperature patterns increases historical estimates of climate sensitivity. *Geophys. Res. Lett.*, **45**, 8490–8499, https://doi.org/10.1029/2018GL078887.

Armour, K. C., 2017: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nat. Climate Change*, **7**, 331–335, https://doi.org/10.1038/nclimate3278.

——, C. M. Bitz, and G. H. Roe, 2013: Time-varying climate sensitivity from regional feedbacks. *J. Climate*, **26**, 4518–4534, https://doi.org/10.1175/JCLI-D-12-00544.1.

——, N. Siler, A. Donohoe, and G. H. Roe, 2019: Meridional atmospheric heat transport constrained by energetics and mediated by large-scale diffusion. *J. Climate*, **32**, 3655–3680, https://doi.org/10.1175/JCLI-D-18-0563.1.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

Azoulay, A., H. Schmidt, and C. Timmreck, 2021: The Arctic polar vortex response to volcanic forcing of different strengths. *J. Geophys. Res. Atmos.*, **126**, e2020JD034450, https://doi.org/10.1029/2020JD034450.

Bintanja, R., E. C. van der Linden, and W. Hazeleger, 2012: Boundary layer stability and Arctic climate change: A feedback study using EC-Earth. *Climate Dyn.*, **39**, 2659–2673, https://doi.org/10.1007/s00382-011-1272-1.

Block, K., and T. Mauritsen, 2013: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO_2. *J. Adv. Model. Earth Syst.*, **5**, 676–691, https://doi.org/10.1002/jame.20041.

Boer, G. J., M. Stowasser, and K. Hamilton, 2007: Inferring climate sensitivity from volcanic events. *Climate Dyn.*, **28**, 481–502, https://doi.org/10.1007/s00382-006-0193-x.

Bony, S., and Coauthors, 2006: How well do we understand and evaluate climate change feedback processes? *J. Climate*, **19**, 3445–3482, https://doi.org/10.1175/JCLI3819.1.

Cai, W., and P. H. Whetton, 2001: A time-varying greenhouse warming pattern and the tropical–extratropical circulation linkage in the Pacific Ocean. *J. Climate*, **14**, 3337–3355, https://doi.org/10.1175/1520-0442(2001)014<3337:ATVGWP>2.0.CO;2.

Ceppi, P., and J. M. Gregory, 2017: Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget. *Proc. Natl. Acad. Sci. USA*, **114**, 13 126–13 131, https://doi.org/10.1073/pnas.1714308114.

——, and ——, 2019: A refined model for the Earth's global energy balance. *Climate Dyn.*, **53**, 4781–4797, https://doi.org/10.1007/s00382-019-04825-x.

Chalmers, J., J. E. Kay, E. A. Middlemas, E. A. Maroon, and P. DiNezio, 2022: Does disabling cloud radiative feedbacks change spatial patterns of surface greenhouse warming and cooling? *J. Climate*, **35**, 1787–1807, https://doi.org/10.1175/JCLI-D-21-0391.1.

Dong, Y., C. Proistosescu, K. C. Armour, and D. S. Battisti, 2019: Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a Green's function approach: The preeminence of the western Pacific. *J. Climate*, **32**, 5471–5491, https://doi.org/10.1175/JCLI-D-18-0843.1.

——, K. C. Armour, M. D. Zelinka, C. Proistosescu, D. S. Battisti, C. Zhou, and T. Andrews, 2020: Intermodel spread in the pattern effect and its contribution to climate sensitivity in CMIP5 and CMIP6 models. *J. Climate*, **33**, 7755–7775, https://doi.org/10.1175/JCLI-D-19-1011.1.

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

Forster, P. M. F., and K. E. Taylor, 2006: Climate forcings and climate sensitivities diagnosed from coupled climate model integrations. *J. Climate*, **19**, 6181–6194, https://doi.org/10.1175/JCLI3974.1.

——, M. Blackburn, R. Glover, and K. P. Shine, 2000: An examination of climate sensitivity for idealised climate change experiments in an intermediate general circulation model. *Climate Dyn.*, **16**, 833–849, https://doi.org/10.1007/s003820000083.

Gregory, J. M., and T. Andrews, 2016: Variation in climate sensitivity and feedback parameters during the historical period. *Geophys. Res. Lett.*, **43**, 3911–3920, https://doi.org/10.1002/2016GL068406.

——, ——, P. Good, T. Mauritsen, and P. M. Forster, 2016: Small global-mean cooling due to volcanic radiative forcing. *Climate Dyn.*, **47**, 3979–3991, https://doi.org/10.1007/s00382-016-3055-1.

——, ——, P. Ceppi, T. Mauritsen, and M. J. Webb, 2020: How accurately can the climate sensitivity to CO_2 be estimated from historical climate change? *Climate Dyn.*, **54**, 129–157, https://doi.org/10.1007/s00382-019-04991-y.

Hahn, L. C., K. C. Armour, M. D. Zelinka, C. M. Bitz, and A. Donohoe, 2021: Contributions to polar amplification in CMIP5 and CMIP6 models. *Front. Earth Sci.*, **9**, 710036, https://doi.org/10.3389/feart.2021.710036.

Hansen, J., M. Sato, and R. Ruedy, 1997: Radiative forcing and climate response. *J. Geophys. Res.*, **102**, 6831–6864, https://doi.org/10.1029/96JD03436.

——, and Coauthors, 2005: Efficacy of climate forcings. *J. Geophys. Res.*, **110**, D18104, https://doi.org/10.1029/2005JD005776.

Haugstad, A. D., K. C. Armour, D. S. Battisti, and B. E. J. Rose, 2017: Relative roles of surface temperature and climate forcing patterns in the inconstancy of radiative feedbacks. *Geophys. Res. Lett.*, **44**, 7455–7463, https://doi.org/10.1002/2017GL074372.

Held, I. M., M. Winton, K. Takahashi, T. Delworth, F. Zeng, and G. K. Vallis, 2010: Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Climate*, **23**, 2418–2427, https://doi.org/10.1175/2009JCLI3466.1.

Ilyina, T., K. D. Six, J. Segschneider, E. Maier-Reimer, H. Li, and I. Núñez-Riboni, 2013: Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations. *J. Adv. Model. Earth Syst.*, **5**, 287–315, https://doi.org/10.1029/2012MS000178.

Joshi, M., K. Shine, M. Ponater, N. Stuber, R. Sausen, and L. Li, 2003: A comparison of climate response to different radiative forcings in three general circulation models: Towards an improved metric of climate change. *Climate Dyn.*, **20**, 843–854, https://doi.org/10.1007/s00382-003-0305-9.

Jungclaus, J. H., and Coauthors, 2013: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.*, **5**, 422–446, https://doi.org/10.1002/jame.20023.

Kummer, J. R., and A. E. Dessler, 2014: The impact of forcing efficacy on the equilibrium climate sensitivity. *Geophys. Res. Lett.*, **41**, 3565–3568, https://doi.org/10.1002/2014GL060046.

Li, C., J.-S. von Storch, and J. Marotzke, 2013: Deep-ocean heat uptake and equilibrium climate response. *Climate Dyn.*, **40**, 1071–1086, https://doi.org/10.1007/s00382-012-1350-z.

Lin, Y.-J., Y.-T. Hwang, P. Ceppi, and J. M. Gregory, 2019: Uncertainty in the evolution of climate feedback traced to the strength of the Atlantic meridional overturning circulation. *Geophys. Res. Lett.*, **46**, 12 331–12 339, https://doi.org/10.1029/2019GL083084.

Maher, N., and Coauthors, 2019: The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *J. Adv. Model. Earth Syst.*, **11**, 2050–2069, https://doi.org/10.1029/2019MS001639.

Manabe, S., and R. T. Wetherald, 1975: The effects of doubling the CO_2 concentration on the climate of a general circulation model. *J. Atmos. Sci.*, **32**, 3–15, https://doi.org/10.1175/1520-0469(1975)032<0003:TEODTC>2.0.CO;2.

Marvel, K., G. A. Schmidt, R. L. Miller, and L. S. Nazarenko, 2016: Implications for climate sensitivity from the response to

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

individual forcings. *Nat. Climate Change*, **6**, 386–389, https://doi.org/10.1038/nclimate2888.

Mauritsen, T., and Coauthors, 2019: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO_2. *J. Adv. Model. Earth Syst.*, **11**, 998–1038, https://doi.org/10.1029/2018MS001400.

Merlis, T. M., I. M. Held, G. L. Stenchikov, F. Zeng, and L. W. Horowitz, 2014: Constraining transient climate sensitivity using coupled climate model simulations of volcanic eruptions. *J. Climate*, **27**, 7781–7795, https://doi.org/10.1175/JCLI-D-14-00214.1.

Mitevski, I., L. M. Polvani, and C. Orbe, 2022: Asymmetric warming/cooling response to CO_2 increase/decrease mainly due to non-logarithmic forcing, not feedbacks. *Geophys. Res. Lett.*, **49**, e2021GL097133, https://doi.org/10.1029/2021GL097133.

Modak, A., G. Bala, L. Cao, and K. Caldeira, 2016: Why must a solar forcing be larger than a CO_2 forcing to cause the same global mean surface temperature change? *Environ. Res. Lett.*, **11**, 044013, https://doi.org/10.1088/1748-9326/11/4/044013.

Pauling, A. G., M. Bushuk, and C. M. Bitz, 2021: Robust inter-hemispheric asymmetry in the response to symmetric volcanic forcing in model large ensembles. *Geophys. Res. Lett.*, **48**, e2021GL092558, https://doi.org/10.1029/2021GL092558.

Pithan, F., and T. Mauritsen, 2014: Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nat. Geosci.*, **7**, 181–184, https://doi.org/10.1038/ngeo2071.

Po-Chedley, S., K. C. Armour, C. M. Bitz, M. D. Zelinka, B. D. Santer, and Q. Fu, 2018: Sources of intermodel spread in the lapse rate and water vapor feedbacks. *J. Climate*, **31**, 3187–3206, https://doi.org/10.1175/JCLI-D-17-0674.1.

Previdi, M., K. L. Smith, and L. M. Polvani, 2021: Arctic amplification of climate change: A review of underlying mechanisms. *Environ. Res. Lett.*, **16**, 093003, https://doi.org/10.1088/1748-9326/ac1c29.

Reick, C. H., and Coauthors, 2021: JSBACH 3—The land component of the MPI Earth System Model: Documentation of version 3.2, MPI für Meteorologie Rep., 271 pp., https://doi.org/10.17617/2.3279802.

Rohrschneider, T., B. Stevens, and T. Mauritsen, 2019: On simple representations of the climate response to external radiative forcing. *Climate Dyn.*, **53**, 3131–3145, https://doi.org/10.1007/s00382-019-04686-4.

Rose, B. E. J., and L. Rayborn, 2016: The effects of ocean heat uptake on transient climate sensitivity. *Curr. Climate Change Rep.*, **2**, 190–201, https://doi.org/10.1007/s40641-016-0048-4.

——, K. C. Armour, D. S. Battisti, N. Feldl, and D. D. B. Koll, 2014: The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophys. Res. Lett.*, **41**, 1071–1078, https://doi.org/10.1002/2013GL058955.

Rugenstein, M. A. A., and K. C. Armour, 2021: Three flavors of radiative feedbacks and their implications for estimating equilibrium climate sensitivity. *Geophys. Res. Lett.*, **48**, e2021GL092983, https://doi.org/10.1029/2021GL092983.

——, K. Caldeira, and R. Knutti, 2016: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes. *Geophys. Res. Lett.*, **43**, 9877–9885, https://doi.org/10.1002/2016GL070907.

Salvi, P., P. Ceppi, and J. M. Gregory, 2022: Interpreting differences in radiative feedbacks from aerosols versus greenhouse gases. *Geophys. Res. Lett.*, **49**, e2022GL097766, https://doi.org/10.1029/2022GL097766.

Schmidt, H., and Coauthors, 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, https://doi.org/10.5194/esd-3-63-2012.

Senior, C. A., and J. F. B. Mitchell, 2000: The time-dependence of climate sensitivity. *Geophys. Res. Lett.*, **27**, 2685–2688, https://doi.org/10.1029/2000GL011373.

Shindell, D. T., G. Faluvegi, L. Rotstayn, and G. Milly, 2015: Spatial patterns of radiative forcing and surface temperature response. *J. Geophys. Res. Atmos.*, **120**, 5385–5403, https://doi.org/10.1002/2014JD022752.

Singh, H., N. Feldl, J. E. Kay, and A. L. Morrison, 2022: Climate sensitivity is sensitive to changes in ocean heat transport. *J. Climate*, **35**, 2653–2674, https://doi.org/10.1175/JCLI-D-21-0674.1.

Soden, B. J., and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean–atmosphere models. *J. Climate*, **19**, 3354–3360, https://doi.org/10.1175/JCLI3799.1.

——, ——, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008: Quantifying climate feedbacks using radiative kernels. *J. Climate*, **21**, 3504–3520, https://doi.org/10.1175/2007JCLI2110.1.

Stevens, B., and Coauthors, 2013: Atmospheric component of the MPI-M Earth system model: ECHAM6. *J. Adv. Model. Earth Syst.*, **5**, 146–172, https://doi.org/10.1002/jame.20015.

——, S. C. Sherwood, S. Bony, and M. J. Webb, 2016: Prospects for narrowing bounds on Earth's equilibrium climate sensitivity. *Earth's Future*, **4**, 512–522, https://doi.org/10.1002/2016EF000376.

Stowasser, M., K. Hamilton, and G. J. Boer, 2006: Local and global climate feedbacks in models with differing climate sensitivities. *J. Climate*, **19**, 193–209, https://doi.org/10.1175/JCLI3613.1.

Timmreck, C., and Coauthors, 2018: The Interactive Stratospheric Aerosol Model Intercomparison Project (ISA-MIP): Motivation and experimental design. *Geosci. Model Dev.*, **11**, 2581–2608, https://doi.org/10.5194/gmd-11-2581-2018.

Toohey, M., B. Stevens, H. Schmidt, and C. Timmreck, 2016: Easy volcanic aerosol (EVA v1.0): An idealized forcing generator for climate simulations. *Geosci. Model Dev.*, **9**, 4049–4070, https://doi.org/10.5194/gmd-9-4049-2016.

Webb, M. J., and Coauthors, 2017: The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6. *Geosci. Model Dev.*, **10**, 359–384, https://doi.org/10.5194/gmd-10-359-2017.

Wigley, T. M. L., C. M. Ammann, B. D. Santer, and S. C. B. Raper, 2005: Effect of climate sensitivity on the response to volcanic forcing. *J. Geophys. Res.*, **110**, D09107, https://doi.org/10.1029/2004JD005557.

Winton, M., K. Takahashi, and I. M. Held, 2010: Importance of ocean heat uptake efficacy to transient climate change. *J. Climate*, **23**, 2333–2344, https://doi.org/10.1175/2009JCLI3139.1.

Wood, R., and C. S. Bretherton, 2006: On the relationship between stratiform low cloud cover and lower-tropospheric stability. *J. Climate*, **19**, 6425–6432, https://doi.org/10.1175/JCLI3988.1.

Yang, W., G. A. Vecchi, S. Fueglistaler, L. W. Horowitz, D. J. Luet, Á. G. Muñoz, D. Paynter, and S. Underwood, 2019: Climate impacts from large volcanic eruptions in a high-resolution climate model: The importance of forcing structure. *Geophys. Res. Lett.*, **46**, 7690–7699, https://doi.org/10.1029/2019GL082367.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC

Zhang, R., S. M. Kang, and I. M. Held, 2010: Sensitivity of climate change induced by the weakening of the Atlantic meridional overturning circulation to cloud feedback. *J. Climate*, **23**, 378–389, https://doi.org/10.1175/2009JCLI3118.1.

Zhao, M., L. Cao, G. Bala, and L. Duan, 2021: Climate response to latitudinal and altitudinal distribution of stratospheric sulfate aerosols. *J. Geophys. Res. Atmos.*, **126**, e2021JD035379, https://doi.org/10.1029/2021JD035379.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nat. Geosci.*, **9**, 871–874, https://doi.org/10.1038/ngeo2828.

——, ——, and ——, 2017: Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a Green's function approach. *J. Adv. Model. Earth Syst.*, **9**, 2174–2189, https://doi.org/10.1002/2017MS001096.

Unauthenticated | Downloaded 12/19/22 04:24 PM UTC