JGR Atmospheres

RESEARCH ARTICLE
10.1029/2020JD034450

The Arctic Polar Vortex Response to Volcanic Forcing of Different Strengths

Alon Azoulay[1,2], Hauke Schmidt[1], and Claudia Timmreck[1]

[1]Max Planck Institute for Meteorology, Hamburg, Germany, [2]Now at Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany

Key Points:

- Model simulations suggest a threshold below which stratospheric sulfur injections from volcanoes do not affect the Arctic polar vortex
- Statistically significant winter warming in Northern Eurasia is simulated for large eruptions

Correspondence to:
H. Schmidt,
hauke.schmidt@mpimet.mpg.de

Citation:
Azoulay, A., Schmidt, H., & Timmreck, C. (2021). The Arctic polar vortex response to volcanic forcing of different strengths. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD034450. https://doi.org/10.1029/2020JD034450

Received 16 DEC 2020
Accepted 17 MAY 2021

Author Contributions:

Conceptualization: Alon Azoulay, Hauke Schmidt, Claudia Timmreck
Data curation: Alon Azoulay
Formal analysis: Alon Azoulay, Hauke Schmidt
Funding acquisition: Hauke Schmidt, Claudia Timmreck
Investigation: Alon Azoulay, Hauke Schmidt, Claudia Timmreck
Methodology: Alon Azoulay, Hauke Schmidt
Project Administration: Hauke Schmidt, Claudia Timmreck
Resources: Hauke Schmidt
Software: Alon Azoulay
Supervision: Hauke Schmidt, Claudia Timmreck
Validation: Hauke Schmidt
Visualization: Alon Azoulay
Writing – original draft: Hauke Schmidt
Writing – review & editing: Alon Azoulay, Claudia Timmreck

Abstract Tropical volcanic eruptions injecting sulfur into the stratosphere are assumed to not only scatter radiation and cool Earth's surface but also alter atmospheric circulation and in particular to strengthen the stratospheric polar vortex in boreal winter. The exact impact is difficult to estimate because of the small number of well observed eruptions and the high internal variability of the vortex. We use large (100-member) ensembles of simulations with an Earth system model for idealized volcanic aerosol distributions resulting from sulfur injections between 2.5 and 20 Tg. We suggest the existence of a threshold somewhere between 2.5 and 5 Tg(S) below which the vortex does not show a detectable response to the injection. This nonlinearity is introduced partly through the infrared aerosol optical density which increases much stronger than linear with increasing particle size occurring for increasing injection amount. Additionally, the dynamical mechanism causing the vortex strengthening seems not to set in for small aerosol loading. Furthermore, we add to the recent discussion concerning a possible downward propagation of the circulation response leading to a winter warming in Northern Eurasia. At latitudes northward of about 50°N, our simulations do show such an average warming pattern that is statistically significant for injections of 10 Tg(S) or more.

Plain Language Summary Large volcanic eruptions can inject sulfur containing gases into the stratosphere where they build sulfate aerosols. These particles (a) scatter incoming sunlight away from the Earth, resulting in a temporary global mean cooling at the surface, and (b) absorb infrared radiation and thereby warm the lower stratosphere. This heating is thought to strengthen the Arctic polar vortex, circumpolar westerly winds in the winter stratosphere. The exact effect of volcanic aerosol on the polar vortex is, however, unknown. Here, we aim to understand the dependence of the vortex strengthening on the amount of injected sulfur. For five different eruption strengths, we simulate the atmospheric response 100 times. We show that the simulated vortex response depends nonlinearly on eruption strength and is indistinguishable from zero for the smallest injection, suggesting a threshold below which the dynamical mechanism leading to the vortex strengthening does not work. The stratospheric Arctic vortex is of interest because there are strong indications that it influences wintertime climate in the troposphere. Starting at an eruption strength similar to the Pinatubo eruption in 1991, our model simulates an increased likelihood of a warmer than normal winter in Northern Eurasia despite the global cooling.

1. Introduction

Large explosive volcanic eruptions can inject sulfur into the stratosphere, which forms sulfate aerosols and leads to global mean surface temperature reductions and changes in atmospheric circulation in the following years (Robock, 2000; Timmreck, 2012). Particular attention has been paid to a possible strengthening of the polar vortex in the Arctic winter stratosphere because via stratosphere-troposphere dynamical coupling this has been linked to tropospheric circulation changes leading, for instance, to a surface winter warming signal in Northern Eurasia (Robock & Mao, 1992; Shindell et al., 2004), although the latter has been disputed recently (Polvani et al., 2019).

Stratospheric volcanic sulfate aerosols not only scatter solar radiation but also absorb solar near-infrared and terrestrial infrared radiation and thereby heat the stratosphere. Stronger aerosol heating at low latitudes and weaker heating during polar night is supposed to lead to a meridional temperature gradient anomaly in the lower stratospheric winter hemisphere and, via thermal wind balance, to a strengthening of the stratospheric winter westerlies. Bittner, Timmreck, et al. (2016) argue that the strengthening of the polar vortex

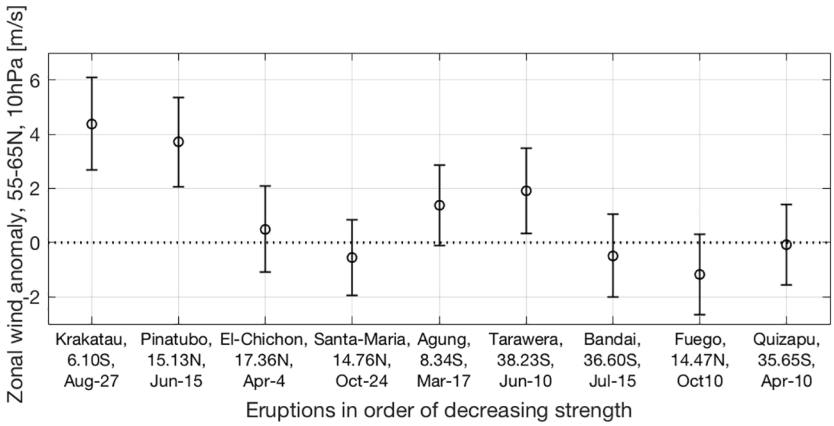

Figure 1. Zonal mean zonal wind anomalies at 10 hPa averaged over 55°N–65°N for the first winter (December–February) after eruptions of the nine largest volcanoes since 1850, as simulated in a 100-member ensemble of historical simulations (Maher et al., 2019). The anomalies for the individual eruptions are calculated with respect to volcanically quiescent periods of different length before the respective eruptions as described by Bittner, Schmidt, et al. (2016). The eruptions are ordered by decreasing strength as given by global tropical mean AOD at 0.55 μm from a reconstruction (Stenchikov et al., 1998). Error bars indicate twice the standard error, that is, the 95% probability range of the mean anomaly. AOD, aerosol optical depth.

(with a center around 60°N) is likely not directly related to the gradient in the heating rate anomaly, which is strongest around 35°N, but involves dynamical propagation of the signal via wave-mean flow interaction. The exact contribution of direct and indirect thermal wind effects to the vortex strengthening is unclear. Additionally, it is unclear to what extent potential effects of volcanic eruptions on stratospheric ozone and on the propagation of planetary waves into the stratosphere may contribute to a vortex strengthening (Graf et al., 2007; Marshall et al., 2009; Stenchikov et al., 2002).

Direct observations of stratospheric circulation anomalies after large volcanic eruptions are limited to the satellite era which includes only two eruptions with large stratospheric sulfur injections, El Chichón in 1982 and Mt. Pinatubo in 1991. Due to the high internal variability of the winter stratosphere, it is unclear how representative the observed cases are for an average response to volcanic eruptions. Studies of circulation changes, therefore, rely heavily on simulations with general circulation models.

Some early modeling studies indicated a strengthening of the Arctic polar vortex (e.g., Stenchikov et al., 2002) as a response to volcanic eruptions. Later, several studies analyzing the "historical" simulations (1850 to present day) in different phases of the Coupled Model Intercomparison Project (CMIP) concluded that recent generations of climate models do not reproduce well such a strengthening (Charlton-Perez et al., 2013; Driscoll et al., 2012; Stenchikov et al., 2006). More recently, Bittner, Schmidt, et al. (2016) showed, however, that the Coupled Model Intercomparison Project Phase 5 (CMIP5) models on average do simulate a forced vortex strengthening if the analysis is limited to the first post-eruption winter and the two strongest eruptions since 1850, Krakatoa (1883) and Pinatubo (1991).

Bittner, Schmidt, et al. (2016) analyzed the historical simulations of the 100-member Max Planck Institute-Grand Ensemble (MPI-GE) performed with the Max Planck Institute Earth System Model (MPI-ESM). They found that if their analysis is limited to winter seasons after the eruptions of Krakatoa and Pinatubo a significant vortex strengthening can be diagnosed from only seven ensemble members, but this number increases when more smaller eruptions are included. Figure 1 shows ensemble mean polar vortex anomalies from the same set of simulations for the nine strongest volcanic eruptions since 1850. The strength of an eruption is measured in this case by tropical mean volcanic aerosol optical depth (AOD) anomalies prescribed in the simulations. It is clear that for smaller eruptions, no significant vortex strengthening is simulated, for some of them even a vortex weakening is indicated. This explains the less robust vortex strengthening for an analysis including a larger number of eruptions. It provokes, however, the question why smaller eruptions do not lead to a forced vortex strengthening. The nine analyzed volcanic eruptions differ not only in terms of eruption strength but also by their location, time of year, and boundary conditions during the

time of eruption. Toohey et al. (2014) suggested that even relatively small differences in the distribution of aerosols may influence the response of atmospheric circulation. To exclude all other factors except for the eruption strength, here we present an analysis of 100-member ensemble simulations with idealized distributions of volcanic aerosol produced by the Easy Volcanic Aerosol (EVA) forcing generator differing only in the amount of injected sulfur. The main goal is to identify if there may be an eruption strength threshold for a vortex strengthening, and what mechanisms could be involved in producing nonlinearities.

A strengthening of the Arctic polar vortex has often been linked to anomalous surface winter warming in Northern Eurasia observed after several large eruptions (Robock, 2002; Shindell et al., 2004). Recently, Polvani et al. (2019), however, disputed the causal link between the volcanic effect on stratospheric circulation and the Pinatubo winter warming pattern. They argued that the forced polar vortex strengthening would be too weak (they showed that the polar vortex in the winter 1991/1992 was not anomalously strong), and the Northern Hemisphere winter climate too variable, such that very likely the Pinatubo eruption would have played no significant role in the occurrence of the winter warming. This view was supported by Polvani and Camargo (2020) who analyzed the boreal winter following the Krakatoa eruption of 1883 in surface temperature reconstructions, reanalysis data, and model ensembles. While reconstructions and reanalysis data confirm a Eurasian warming in December–February of the winter 1883/1884, this is within two standard deviations of the observed variability, and neither the MPI-ESM Grand Ensemble nor the CMIP5 ensemble of high-top models shows a significant forced Eurasian warming. Zambri and Robock (2016), however, detected a warming statistically significant at the 90% level in a CMIP5 ensemble when averaging over the first winters after the Pinatubo and Krakatoa eruptions. Here, we analyze in our idealized ensembles if the occurrence of a winter warming pattern may depend on the strength of an eruption. Using a sufficiently large model ensemble with 100 realizations for each scenario, we will be able to distinguish between forced response and internal variability.

In Section 2, we describe the model used for the simulations, the EVA forcing generator, and the setup of the simulations. Results are provided in Section 3 and conclusions and a summary in Section 4.

2. Tools and Experiments

2.1. MPI-ESM

All simulations analyzed in this study have been performed with the Max Planck Institute Earth System Model in its low resolution (MPI-ESM-LR) configuration. We are using version 1.1.00p2 which is the successor of the MPI-ESM version used in CMIP5 as described by Giorgetta et al. (2013) and contains already many of the updates included in the CMIP6 version of the model (Mauritsen et al., 2019).

The MPI-ESM contains the component models ECHAM6.3 (Stevens et al., 2013) for the atmosphere, JSBACH3.0 (Reick et al., 2013) for the land surface, MPIOM1.6 (Jungclaus et al., 2013) for the ocean, and HAMOCC (Ilyina et al., 2013) for the ocean biogeochemistry. The LR configuration uses for the atmosphere a spectral dynamical core with truncation at wavenumber 63 (T63; corresponding to horizontal resolution of ∼1.9°) and 47 hybrid pressure levels up to 0.01 hPa. The ocean component uses a GR15L40 grid which has 40 vertical levels and a horizontal resolution of nominally about 1.5° which is, however, getting finer toward the poles.

The MPI-ESM does not include interactive representations of atmospheric composition. Aerosol effects on radiation are represented by prescribing monthly mean climatologies of their optical properties (Giorgetta et al., 2013, see below). Similarly, radiatively active chemical species are prescribed, except for water vapor which is interactive. This means in particular that ozone is prescribed as a monthly zonal mean climatology. In the historical simulations analyzed here, the (Cionni et al., 2011) climatology is used that does not account for potential influences of volcanic aerosol. Due to the relatively low vertical resolution of the atmospheric component, the MPI-ESM-LR does not generate a quasi-biennial oscillation of tropical stratospheric winds (H. Schmidt et al., 2013). This may be important because observations indicate that the NH polar vortex is weaker on average in the easterly QBO phase than in the westerly phase (Holton & Tan, 1980), an effect that may interact with that of volcanic aerosols.

2.2. MPI-GE

The MPI-GE (Maher et al., 2019) is a 100-member ensemble of several CMIP5 experiments. Here, we are only using the post-Pinatubo period of the historical simulations of the MPI-GE which were started in 1850 from restart files for different years of a corresponding preindustrial control simulation. To account for volcanic aerosol forcing, the simulations use an extended version of the Stenchikov reconstruction (Stenchikov et al., 1998) as described by H. Schmidt et al. (2013). This data set provides aerosol optical properties (extinction, single scattering albedo, and asymmetry factor) calculated for the 40 wavelengths of the model's radiative transfer scheme on the basis of satellite observations of AOD and effective radius. A description of the aerosol distribution is provided in Section 2.4. Volcanic aerosol signals in these simulations have already been analyzed by Bittner, Schmidt, et al. (2016).

2.3. The Idealized Volcanic Forcing Ensemble (EVA-ENS)

As mentioned above, it is difficult to interpret, for example, CMIP historical simulations for the dependence of circulation responses on the strength of a volcanic eruption. The main reason is that the eruptions in these simulations additionally differ in particular in timing, location, and the altitude of its stratospheric injection. To avoid this complication, we are using the EVA forcing generator (Toohey et al., 2016), see Section 2.4, to generate four idealized data sets of zonally and monthly mean volcanic aerosol optical properties (extinction, asymmetry factor, single scattering albedo) with the magnitude of sulfur injections into the lower stratosphere being the only varying input parameter. The idealized forcing data sets are used as an input for the MPI-ESM-LR described above. We produced 100-member ensembles of simulations for each data set. The 100 simulations for each forcing data set start from the 100 different meteorological situations of the historical simulations of the MPI-GE for January 1, 1991, and run for 3 years using exactly the same setup and boundary conditions as the original historical simulation of the MPI-GE, except for the stratospheric volcanic forcing.

The chosen injections for our idealized experiments are 2.5, 5, 10, and 20 Tg of sulfur (Tg[S]). The location of all eruptions is chosen to be at the equator and all injections are taking place at June 15, which corresponds to the day of the Pinatubo eruption in 1991. From satellite observations, Guo et al. (2004) have estimated the stratospheric SO_2 injection of the Pinatubo eruption to be of 18–19 Tg (corresponding to 9–9.5 Tg[S]), a number which is, however, considered uncertain (Kremser et al., 2016). In addition to the four ensembles for different idealized volcanic aerosol forcing fields for comparison, we also run a 100-member ensemble without volcanic emission (0 Tg(S)) as control run. Six-hourly output of all simulations was used to compute transformed Eulerian mean diagnostics (Andrews et al., 1987).

2.4. Characteristics of Volcanic Aerosol Distributions

We are using the EVA forcing generator (Toohey et al., 2016) to compute idealized representations of volcanic aerosol forcing. EVA generates stratospheric aerosol optical properties as a function of time, latitude, height, and wavelength for a given input of location, time, and amount of sulfur injected into the stratosphere from a volcanic eruption. By construction, EVA-generated AOD at 0.55 μm depends linearly on the injected sulfur amount in the magnitude range considered in our experiments, which is a common assumption (Crowley & Unterman, 2013; G. A. Schmidt et al., 2011). Based on observations and simple scaling arguments, EVA assumes that the aerosol effective radius, used in the calculation of optical properties, is locally proportional to the mass of sulfate to the power of 1/3. Build-up and decay of sulfate aerosols is simulated using observation-based assumptions on lifetimes of stratospheric SO_2 and SO_4 of 180 and 330 days, respectively.

The aerosol distributions resulting from EVA are necessarily idealized because stratospheric transport is represented in a very simplified approach between only three boxes (one for the tropics plus extratropical boxes for both hemispheres). This means that, for instance, the transport barrier provided by the polar vortex in winter hemispheres is ignored as well as the potential impact of different injection heights. The upper panels of Figure 2 provide a comparison of AOD at 0.55 μm from EVA for a tropical injection of 9 Tg sulfur in June and the Stenchikov reconstruction (Stenchikov et al., 1998) for the Pinatubo period as used in the historic simulations of the MPI-GE. While it will be shown below that the EVA-produced AOD for

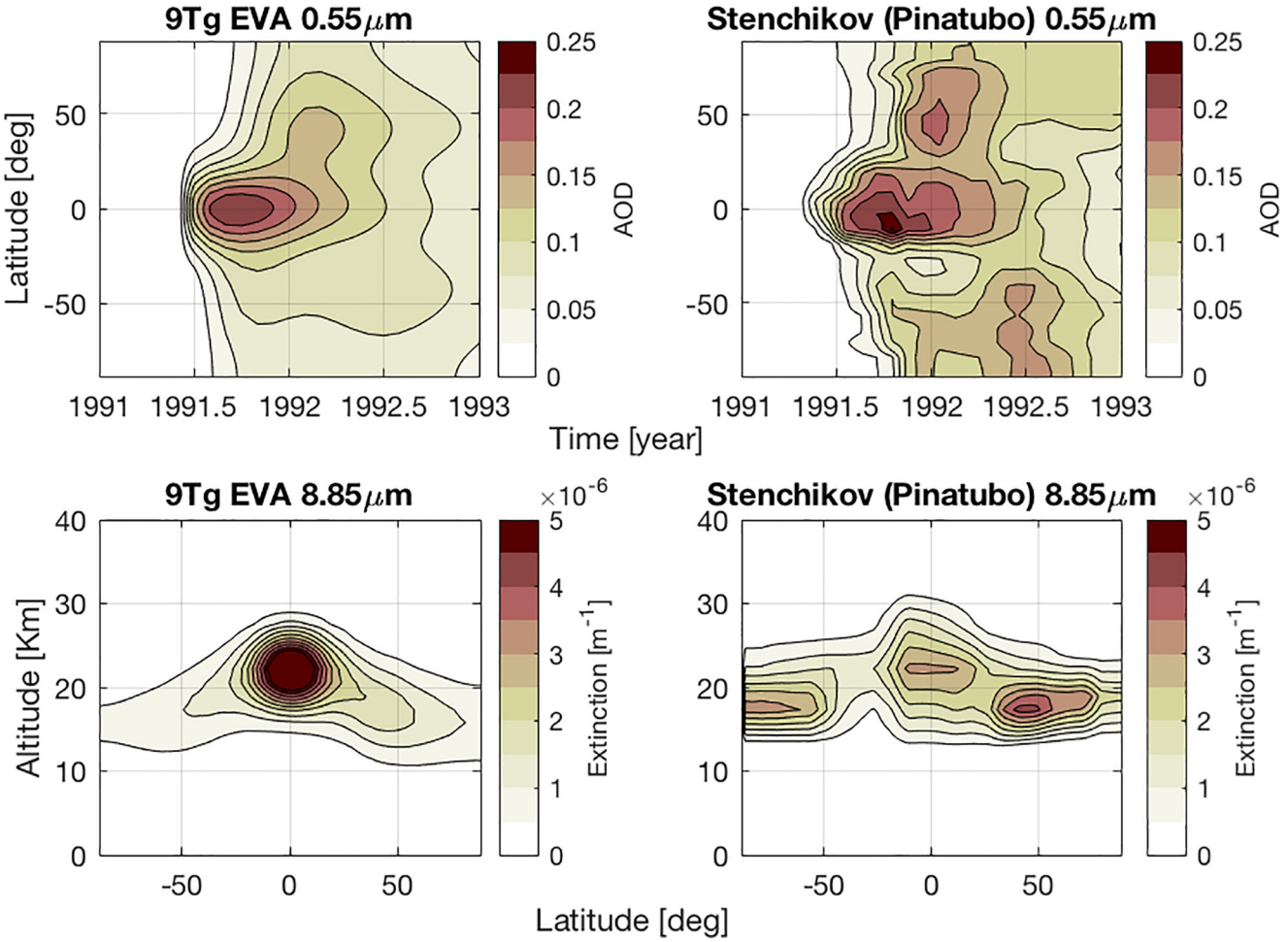

Figure 2. Optical properties of volcanic aerosol distributions (left panels) generated by EVA assuming an injection of 9 Tg of sulfur into the stratosphere at June 15 and (right panels) from the Stenchikov reconstruction for the period of the eruption of Mt. Pinatubo. Upper panels show time-latitude evolutions of volcanic AOD for radiation of 0.55 μm wavelength and lower panels show zonal averages of extinction (10^{-6} m^{-1}) of radiation of 8.85 μm wavelength averaged over the first posteruption boreal winter (DJF). EVA, Easy Volcanic Aerosol; AOD, aerosol optical depth; DJF, December–February.

some wavelength bands does not depend strictly linearly on the injected amount of sulfur, the AOD distribution patterns for the injection amounts chosen in our experiments (not shown) are to first order well approximated by the pattern of the 9 Tg(S) injection given in Figure 2 with just the magnitude scaled by the injection amount.

As shown in Figure 2, assuming an injection of 9 Tg(S) EVA would produce a volcanic AOD distribution similar, but by no means equal to the Stenchikov reconstruction. Differences in the Southern Hemisphere are partly related to the eruption of the Chilean volcano Cerro Hudson (45°S) in October 1991 which is only represented in the Stenchikov reconstruction. A difference in the Northern Hemisphere is that the Stenchikov reconstruction has larger first-winter AOD values in the extratropics resulting in a larger meridional AOD gradient in high latitudes while the EVA gradient is smoother. This difference is clearly visible also in the zonal mean extinction for 8.85 μm radiation, presented for the first posteruption boreal winter in the lower panels of Figure 2. Also the vertical profiles differ slightly. In the tropics, for instance, the magnitude of the EVA peak is larger, but the Stenchikov climatology extends further to higher and lower altitudes. Ångström exponents used in EVA and the Stenchikov reconstruction to derive extinction at other wavelengths than 0.55 μm apparently differ slightly such that the tropical extinction peak in the Stenchikov reconstruction is rather comparable to an assumed EVA injection of about 6 Tg(S).

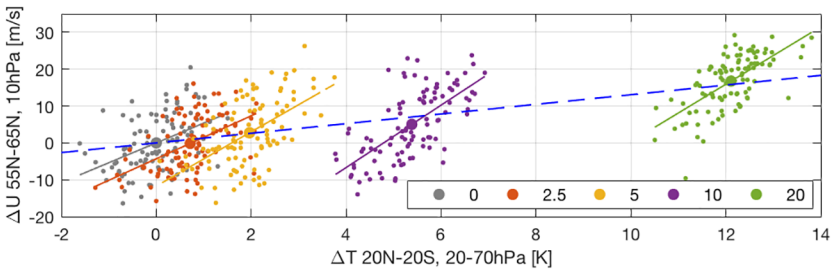

Figure 3. Scatter plot of simulated polar vortex zonal mean zonal wind anomalies versus tropical temperature anomalies averaged over altitude and latitude ranges indicated at the axes for the first winter (DJF) after the EVA-generated volcanic eruptions. Small dots mark anomalies for all individual eruptions of the ensembles for injections of 0 (gray), 2.5 (red), 5 (yellow), 10 (purple), and 20 (green) Tg(S). Anomalies are calculated with respect to the ensemble mean of the 0 Tg(S) simulations. Large dots show ensemble mean anomalies. The blue dashed line is the result from a linear regression taking into account all anomalies from all ensembles; solid lines are results from regressions for the ensembles for individual injection strengths. DJF, December–February; EVA, Easy Volcanic Aerosol.

In addition to volcanic aerosols, EVA can also account for stratospheric background sulfate aerosols for which we choose typical values from industrial volcanically quiescent periods. These background values are, however, orders of magnitude smaller than the aerosol loading from even the smallest eruptions considered in our experiments.

3. Results

3.1. From Sulfur Injection to Vortex Strengthening

For the purpose of identifying potential nonlinearities between the amount of injected sulfur and the Arctic vortex anomalies, we first analyze DJF (December–February) mean bulk quantities characteristic for different processes involved in the vortex strengthening.

Figure 3 is a scatter plot of the anomalies of zonal mean zonal wind at 10 hPa averaged over 55°N–65°N, as a proxy for the vortex strengthening, versus the zonal mean temperature anomaly at 50 hPa averaged over 20°S–20°N, as a proxy for aerosol induced tropical warming. Anomalies are shown for every ensemble member of the experiments with injection amounts from 0 to 20 Tg(S) calculated with respect to the average zonal wind and temperature of the reference ensemble with 0 Tg(S) injection. The figure not only provides an impression of the variability of the individual ensemble members but also gives an indication of differences caused by the injection amount. It is obvious that regression slopes calculated for the individual ensembles (about 7 m s^{-1} K^{-1}) are much steeper than the slope resulting from the regression over all simulations (about 1.3 m s^{-1} K^{-1}). The spread of the members of each ensemble is caused by the internal variability of the model dynamics which is, hence, characterized by the steep slope. We assume that the strong high latitude internal variability in the Arctic winter stratosphere entails a relatively weak variability of the stratospheric mean meridional circulation that adiabatically influences tropical lower stratospheric temperature. The weaker overall slope represents the effect of increasing volcanic injections, and the cause and effect-relation is rather in the opposite direction. Tropical temperature anomalies caused by volcanic aerosol change the meridional temperature gradient and lead to a vortex strengthening.

From the scatter plot, it is difficult to clearly identify potential nonlinearities although it indicates, for example, that the average temperature anomaly of about +5.3 K in the 10 Tg(S) ensemble is more than double than in the 5 Tg(S) ensemble (+1.9 K). Figure 4 shows relations between ensemble mean bulk quantities. Panel (a) shows the relation between the Arctic vortex strengthening and the amount of sulfur injected. Error bars represent 2 times the standard error, that is, with a probability of 95% the forced response of the individual ensemble experiments is within the error bars. The average vortex anomaly of the 2.5 Tg(S) experiments is -0.2 ± 1.9 m/s while for all experiments with higher injection amounts a zero vortex response is outside the 95% error range. This indicates a possible threshold for a nonzero forced winter mean vortex response somewhere between injections of 2.5 and 5 Tg(S). This threshold implies nonlinearity of the response. A regression line restricted to include the origin, that is, a zero vortex anomaly for zero injection,

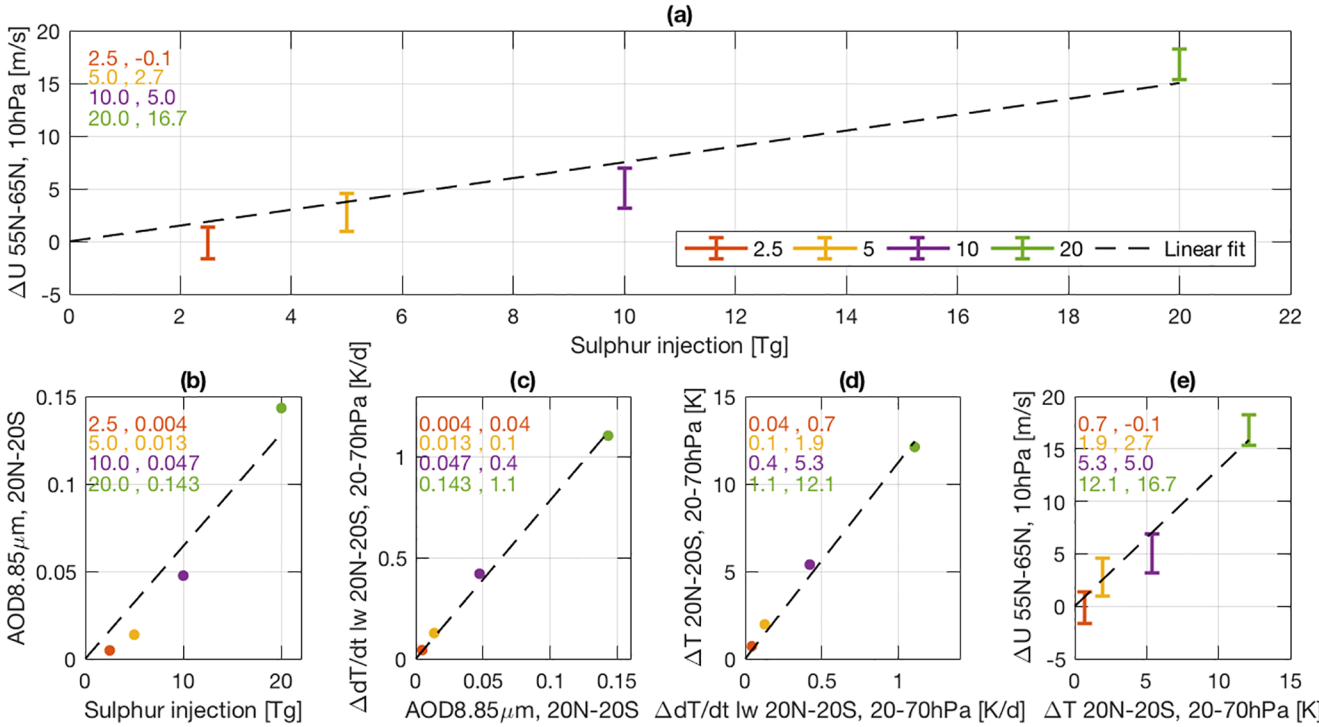

Figure 4. Scatter plots of quantities characterizing the simulated ensembles for volcanic eruptions of different strength and simulated ensemble mean anomalies calculated with respect to the ensemble mean of the 0 Tg(S) simulation. (a) Zonal mean zonal wind (at 10 hPa averaged over 55°N–65°N) anomaly versus amount of sulfur injected into the stratosphere; (b) Global mean 8.85 μm AOD averaged over 20°S–20°N versus amount of injected sulfur; (c) Adiabatic longwave heating rate anomaly averaged over 20°S–20°N and 20–70 hPa versus AOD as in (b); (d) Tropical temperature anomaly (averaged over 20°S–20°N and 20–70 hPa) versus heating rate anomaly as in (c); (e) Zonal wind anomaly as in (a) versus temperature anomaly as in (d). The error bars of the wind anomalies are showing twice the standard error. Error bars for all other quantities are neglected due to their smallness. All values are ensemble mean averages over DJF, except the heating rate anomalies which are calculated just for one member for each experiment with the same initial conditions but different forcing, and for the 10th time step of a simulation starting on January 1. Dashed lines show results from linear regressions restricted to include the point of origin. Colored numbers indicate the coordinates of individual ensemble means. AOD, aerosol optical depth; DJF, December–February.

lies outside the 95% error ranges of some of the injections (Figure 4a). An unrestricted regression (not shown) would indicate an unphysical negative vortex response for zero forcing. Of course, we cannot say if the threshold is indeed a limit for a zero forced response or if it just represents a detection threshold for the given, very large ensemble size.

The further panels of Figure 4 break down the relation between vortex response and injected sulfur into a chain of processes. To characterize the tropical heating agent, we show the average AOD between 20°S and 20°N for the model's infrared radiation band centered at 8.85 μm which is the strongest contributor to aerosol heating. The AOD increases stronger than linear with the injected sulfur amount (Figure 4b). While an injection of 2.5 Tg(S) causes a tropical mean AOD in the first winter after the injection of 0.005 the four times stronger injection of 10 Tg(S) causes an about 10 times larger AOD of 0.047, and the eight times stronger injection an almost 30 times larger AOD. As mentioned above, EVA-generated AOD at 0.55 μm depends linearly on the injected sulfur amount in the magnitude range considered here. The stronger than linear increase of AOD at 8.85 μm is a feature of the Mie calculation-based lookup tables used in EVA for deriving extinction at wavelengths other than 0.55 μm. With increasing injection amount, the effective particle radius increases causing the infrared AOD to increase stronger than linear. The longwave diabatic heating rates (also taken at 50 hPa and averaged over 20°S–20°N) caused by the volcanic aerosols increase almost linearly with the AOD (Figure 4c). The same is true for the increase of the temperature anomaly with the heating rate anomaly (Figure 4d). The relation of polar vortex wind anomaly and tropical temperature (Figure 4e; already discussed for Figure 3) is again linear within the relatively large error bars caused by the high internal variability of the vortex. The split into a chain of processes shows that the relatively weak

infrared AOD for weak injections (as shown in Figure 4b) contributes to the vortex response being indistinguishable from zero for the injection of 2.5 Tg(S).

3.2. Monthly Mean Wind, Temperature, and Planetary Waves

The analysis of winter mean bulk quantities above has indicated the existence of a threshold of stratospheric aerosol injections below which the vortex does not strengthen. To better understand this behavior, we analyze ensemble mean responses depending on latitude and altitude and broken down into individual months of boreal winter.

Figure 5 shows zonal mean temperature responses for the Pinatubo ensemble and all EVA ensembles for the months of October–March. Lower stratospheric temperature anomalies reach a maximum of almost 5 K toward late winter and are probably biased high as observations indicate anomalies closer to 2–3 K (Labitzke & McCormick, 1992). The EVA anomalies in Figure 5 are scaled to an injection of 10 Tg(S). As mentioned above, the Pinatubo aerosol distribution would match closest to an EVA injection of about 9 Tg(S) which at least partly explains the overall slightly weaker anomalies for the Pinatubo ensemble in comparison to the anomalies scaled to 10 Tg(S). In general, anomaly patterns look very similar for all ensembles with one maximum in the tropical lower stratosphere, near the maximum of the aerosol loading, and another maximum near the Arctic stratopause which can be attributed to adiabatic warming through increased downwelling (not shown). In general, scaled temperature anomalies are larger for high than for low injection rates indicating a stronger than linear increase of temperature anomalies with the injection amount that is, as discussed above (Figure 4b), largely due to the behavior of infrared AOD. Significant negative temperature anomalies in the lower polar stratosphere are analyzed for all months in the strongest injection experiments and a few months of the weaker injection experiments. This is related to adiabatic cooling as discussed by Toohey et al. (2014). Overall, and as can be expected, the area of temperature anomalies significant at the 95% confidence level increases with the injection amount. While for the 20 Tg(S) ensemble only very small regions show no significant response, in the 2.5 Tg(S) ensemble the tropical lower stratosphere is the only region that shows significant responses for all months. Polar temperature anomalies are insignificant in this experiment except for a small altitude range near the stratopause showing a negative anomaly, that is, opposite to the other experiments, in January.

Zonal mean zonal wind anomalies scaled by the injection amount and corresponding to the temperature anomalies via the thermal wind relation are presented in Figure 6. In October, all experiments show the strongest positive zonal wind in the subtropical lower to middle stratosphere, that is, in the regions of the strongest diabatic heating rate gradients. Bittner, Timmreck, et al. (2016) suggested that this early winter anomaly contributes to the vortex strengthening in later winter months through its influence on wave propagation. Anomalies simulated for November indicate nonlinear dependence on the injection amount concerning the timing of the propagation of the maximum wind anomaly to high latitudes. While for the 2.5 and 5 Tg(S) cases the maximum anomaly is still clearly located in the subtropics, larger injection experiments, including Pinatubo, already show at least a second maximum in the polar vortex area. These large injection experiments show no clear seasonal evolution in later winter months, unlike the downward propagation demonstrated for responses to solar or ENSO forcing (Kodera & Kuroda, 2002; Manzini et al., 2006). In the smaller injection experiments, the behavior is less clear. Mean polar vortex zonal wind anomalies in the 2.5 Tg(S) case are negative from November to January, but not statistically significant. Differences between the experiments are statistically significant, however, as presented in Figure 7 which shows differences between scaled zonal wind anomalies for early winter (October, November), midwinter (December, January), and late winter (February, March) in the 2.5, 5, and 10 Tg(S) experiments with respect to 20 Tg(S). The strongest, statistically significant difference is simulated for 2.5 Tg(S) compared to 20 Tg(S) for midwinter. This again suggests the existence of a threshold as discussed above. Scaled polar vortex zonal wind anomalies are statistically significantly weaker at least in midwinter to late winter also in the 5 and 10 Tg(S) cases compared to the largest injection.

Why is there such a threshold of sulfur injection below which the polar vortex seems not to react? Why should not also a relatively weak warming in the tropical lower stratosphere cause a forced vortex response detectable in such a large ensemble? Global stratospheric circulation patterns are largely controlled by wave-mean flow interactions. Figure 8 shows Eliassen-Palm (EP)-flux anomalies, that is, anomalies of wave

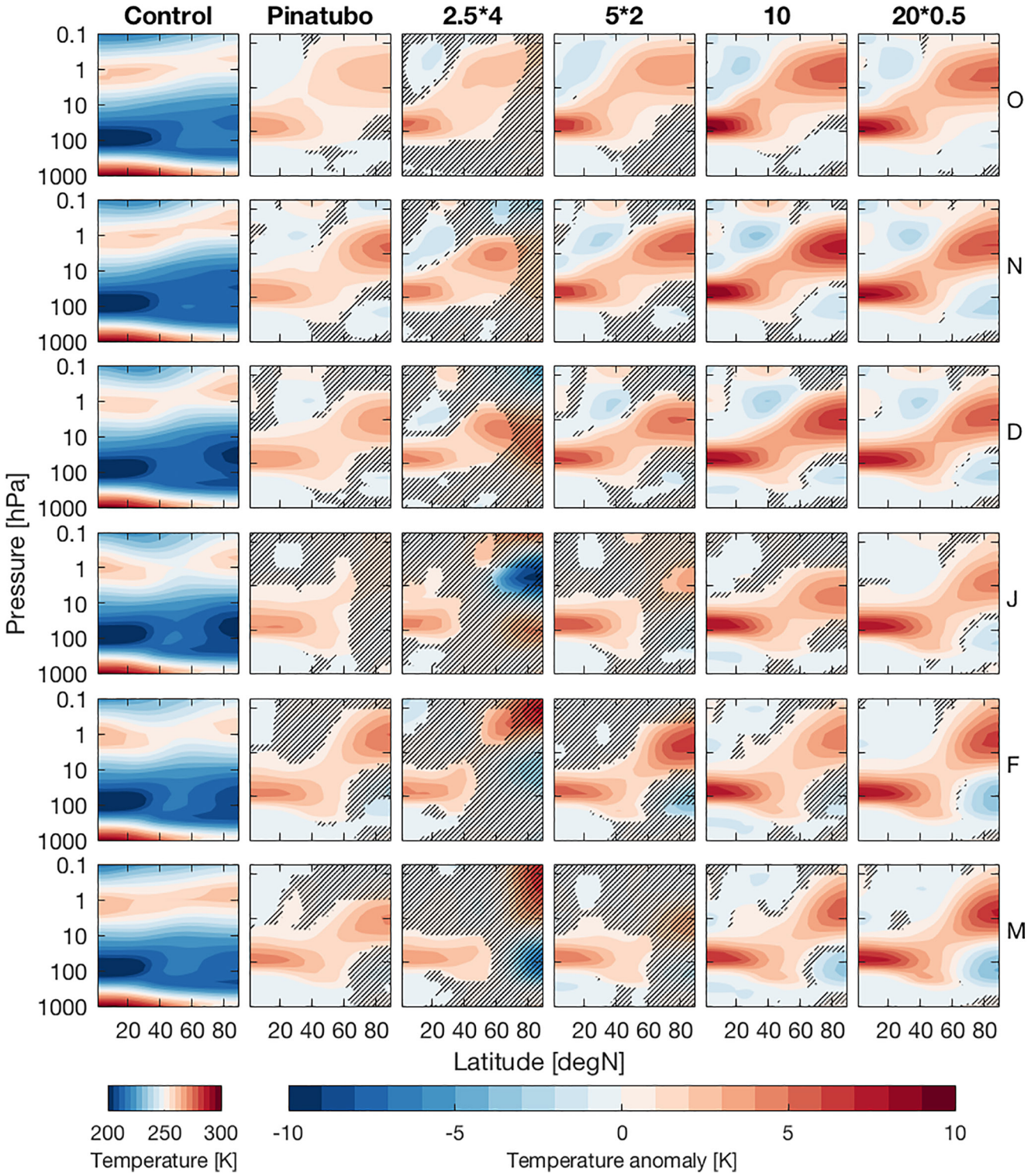

Figure 5. Monthly, zonal, and ensemble mean temperatures (left column) and temperature anomalies (columns 2–6) in K simulated for the months from October to March (top to bottom rows). The control temperature (left column) is from the 0 Tg(S) experiment. Anomalies are from the first posteruption winters of the MPI-GE historical simulation (second column) and from the 2.5, 5, 10, and 20 Tg(S) experiments (columns 3–6) and calculated with respect to the control. Anomalies in columns 3–6 are linearly scaled by $10/n$ where n is the amount of sulfur (Tg[S]) injected in the respective experiment. In unhatched regions, anomalies are statistically significantly different from zero at the 95% level according to a two-tailed t test. MPI-GE, Max Planck Institute Grand Ensemble.

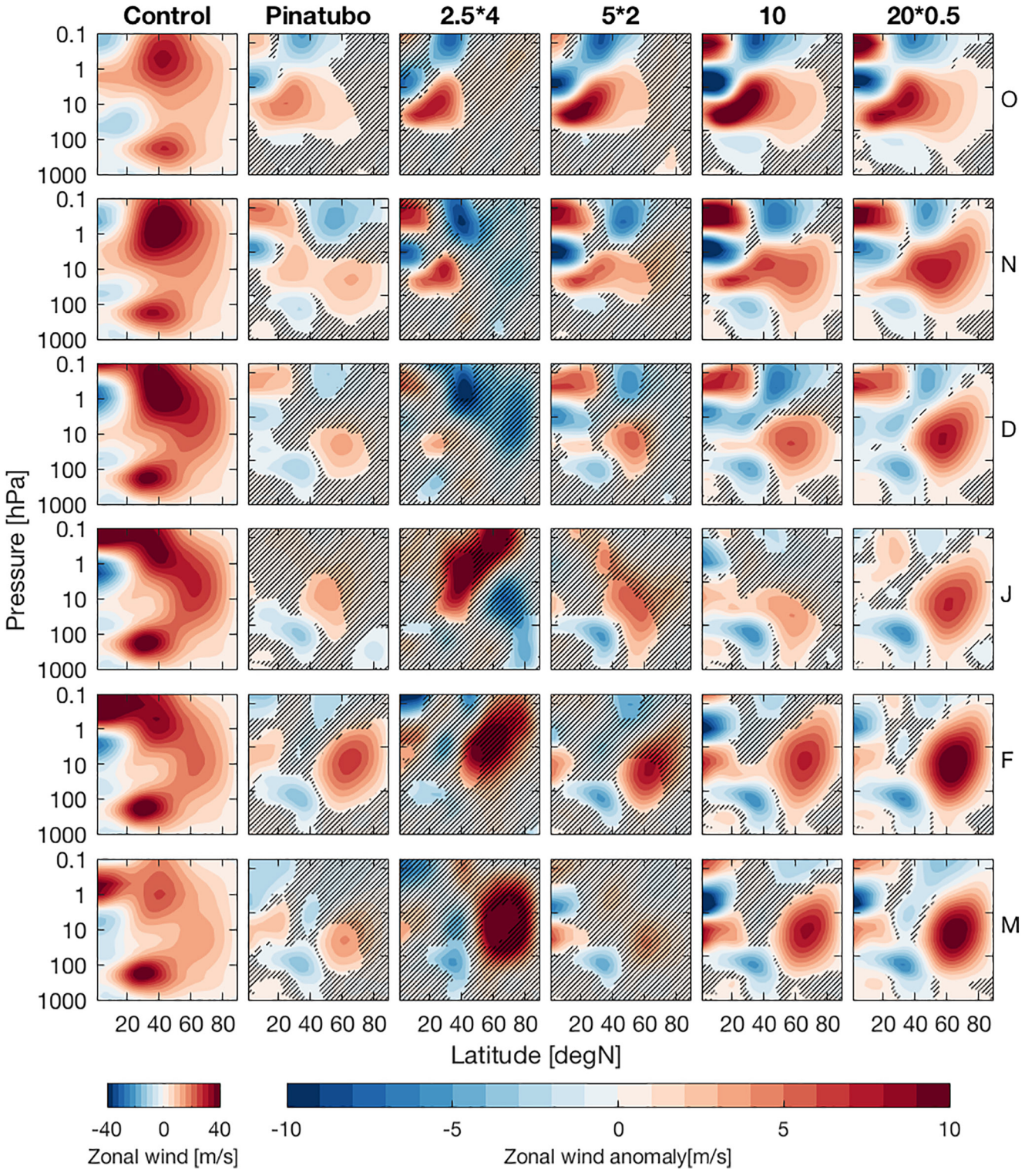

Figure 6. Same as Figure 5 but for zonal wind (m/s).

activity fluxes resolved by the model, and the zonal wind forcing exerted by the EP-flux divergence. Both quantities are, again, linearly scaled to an injection of 10 Tg(S). Bittner, Timmreck, et al. (2016) suggested that the westerly wind anomaly at low to middle latitudes in the lower to middle stratosphere causes a stronger equatorward propagation of planetary waves that leads to less wave convergence at high latitudes

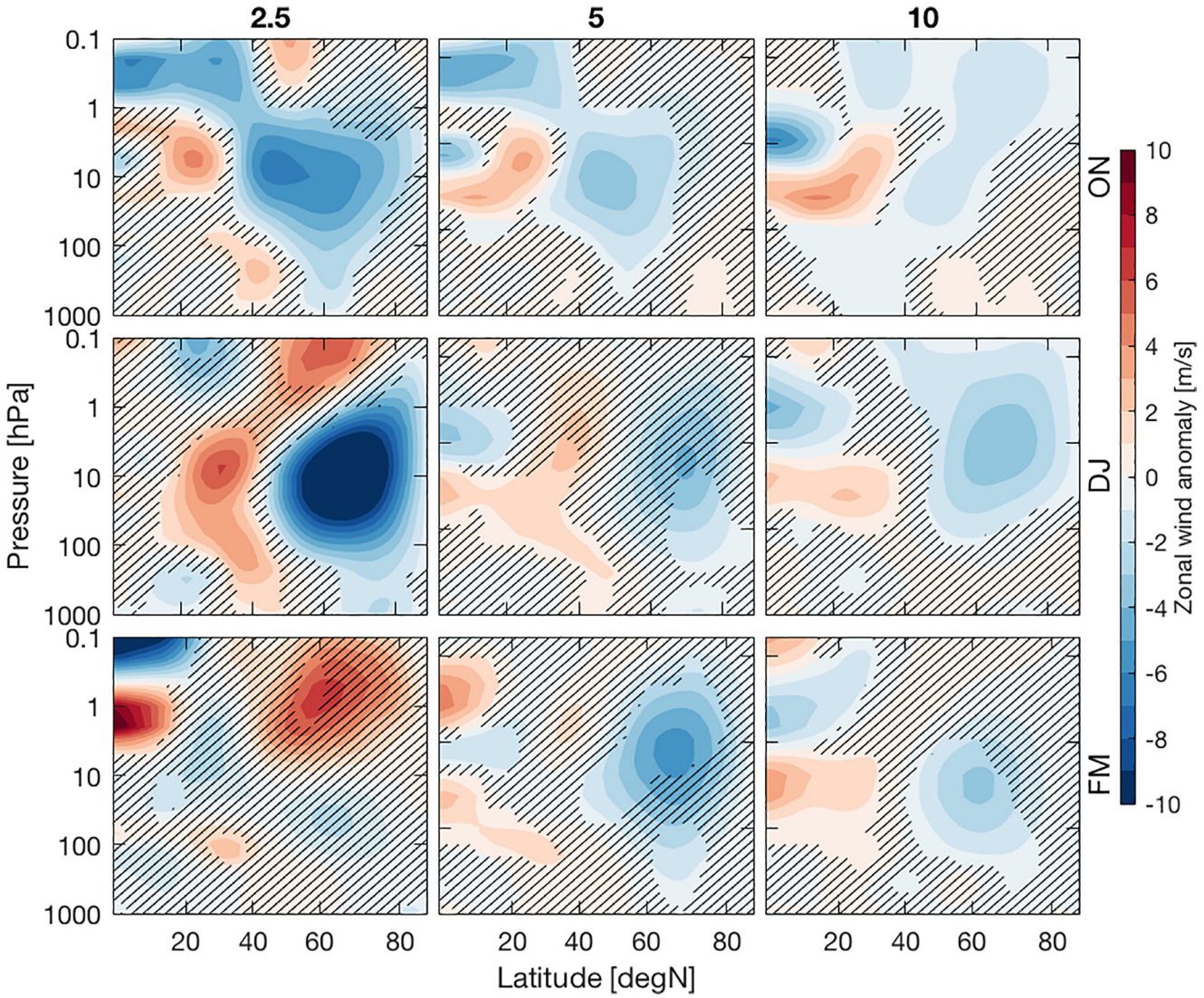

Figure 7. Differences of scaled anomalies of bimonthly, zonal, and ensemble mean zonal winds (m/s). Differences are calculated by subtracting the anomalies of the 20 Tg(S) experiment from the anomalies of the 2.5, 5, and 10 Tg(S) experiments (columns 1–3, respectively) scaled by $20/n$ Tg(S), where n is the amount of sulfur injected in the respective experiment. Top to bottom rows show the bimonthly means of October–November, December–January, and February–March, respectively. In unhatched regions, differences are statistically significantly different from zero at the 95% level according to a two-tailed t test.

and a stronger polar vortex over the course of the winter. EP-flux vectors in Figure 8 indeed indicate a stronger equatorward wave propagation in the lower to middle stratosphere at latitudes lower than about 50°N for all injection amounts. In the experiments for injections of 5 Tg(S) or more, around 10 hPa, this is related to a dipole of anomalous EP-flux convergence at latitudes lower and divergence at latitudes higher than about 40°N. In all these experiments and all through the winter, the anomalous divergence, indicating less deceleration of westerly winds, is statistically significant close to the polar vortex near 10 hPa. In the 2.5 Tg(S) experiment, there is no such significant high-latitude divergence simulated. The equatorward wave propagation anomaly is statistically significant only in a small altitude range and obviously too weak to have a strengthening impact on the polar vortex. EP-flux anomalies in the 2.5 Tg(S) case differ from the other experiments strongly above 1 hPa. It is possible that the easterly wind anomaly related to the strong divergence in early and middle winter propagates downward and also influences the vortex in the middle stratosphere. However, it is not easy to identify a clear connection of this high altitude anomaly to small injections of volcanic aerosol.

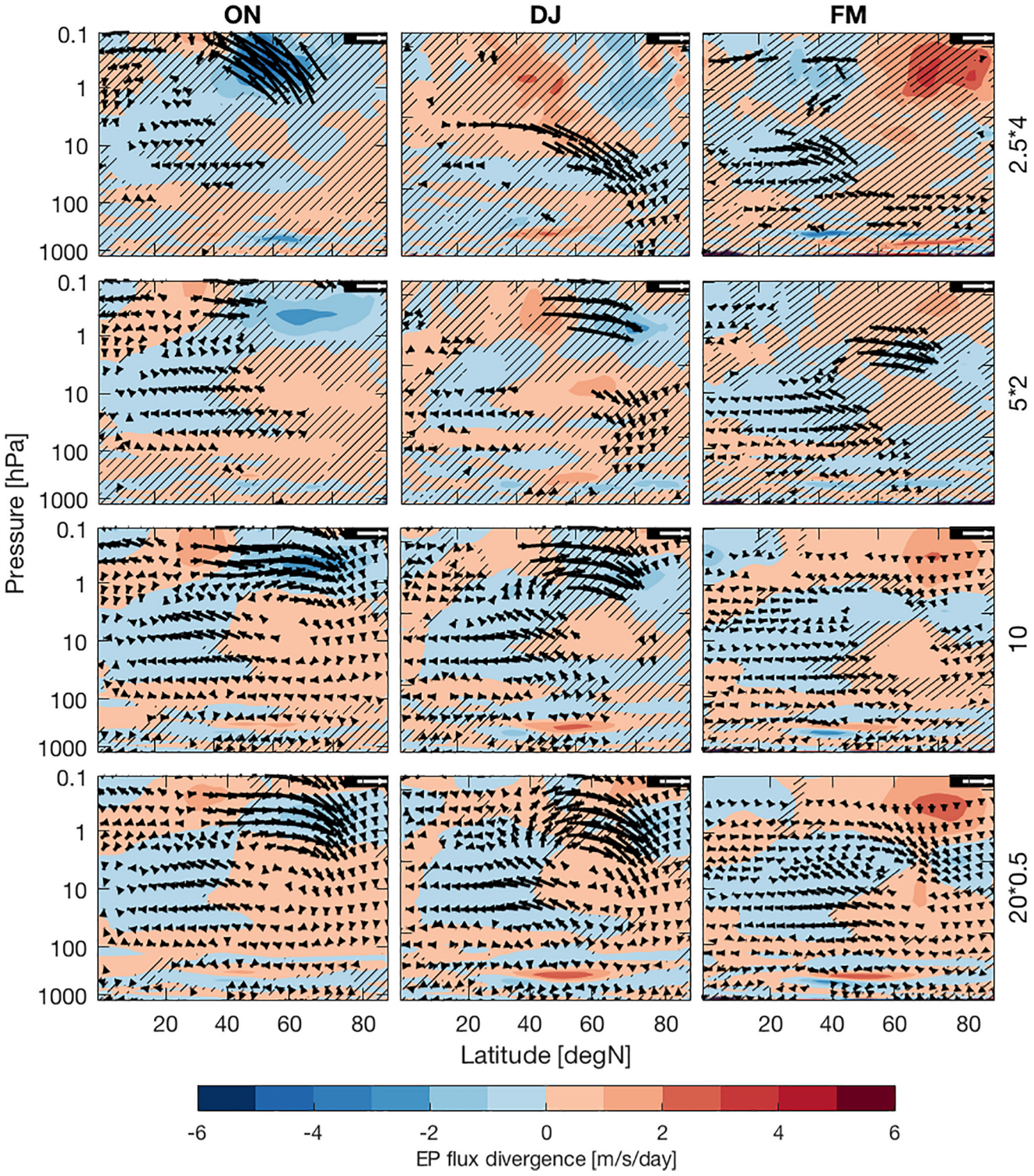

Figure 8. Bimonthly, zonal, and ensemble mean anomalies of EP-fluxes (arrows) and zonal wind tendencies (color shading, m s^{-1} day^{-1}) from EP-flux divergences and divergence anomalies for all experiments. All anomalies are linearly scaled by 10/*n*, where *n* is the amount of sulfur injected (Tg[S]) in the respective experiment. For better visualization, EP-fluxes are scaled by the inverse density and plotted only for selected levels. The length of the white reference arrows in the upper right corners indicate a value of 10^8 m^3 s^{-2}. Top to bottom rows show the experiments with 2.5, 5, 10, and 20 Tg(S) injections, respectively. Left to right columns show bimonthly means of October–November, December–January, and February–March, respectively. In unhatched regions, EP-flux divergence anomalies are statistically significantly different from zero at the 95% level according to a two-tailed *t* test. EP-flux anomaly vectors are only plotted where the statistical significance exceeds 95%. EP, Eliassen-Palm.

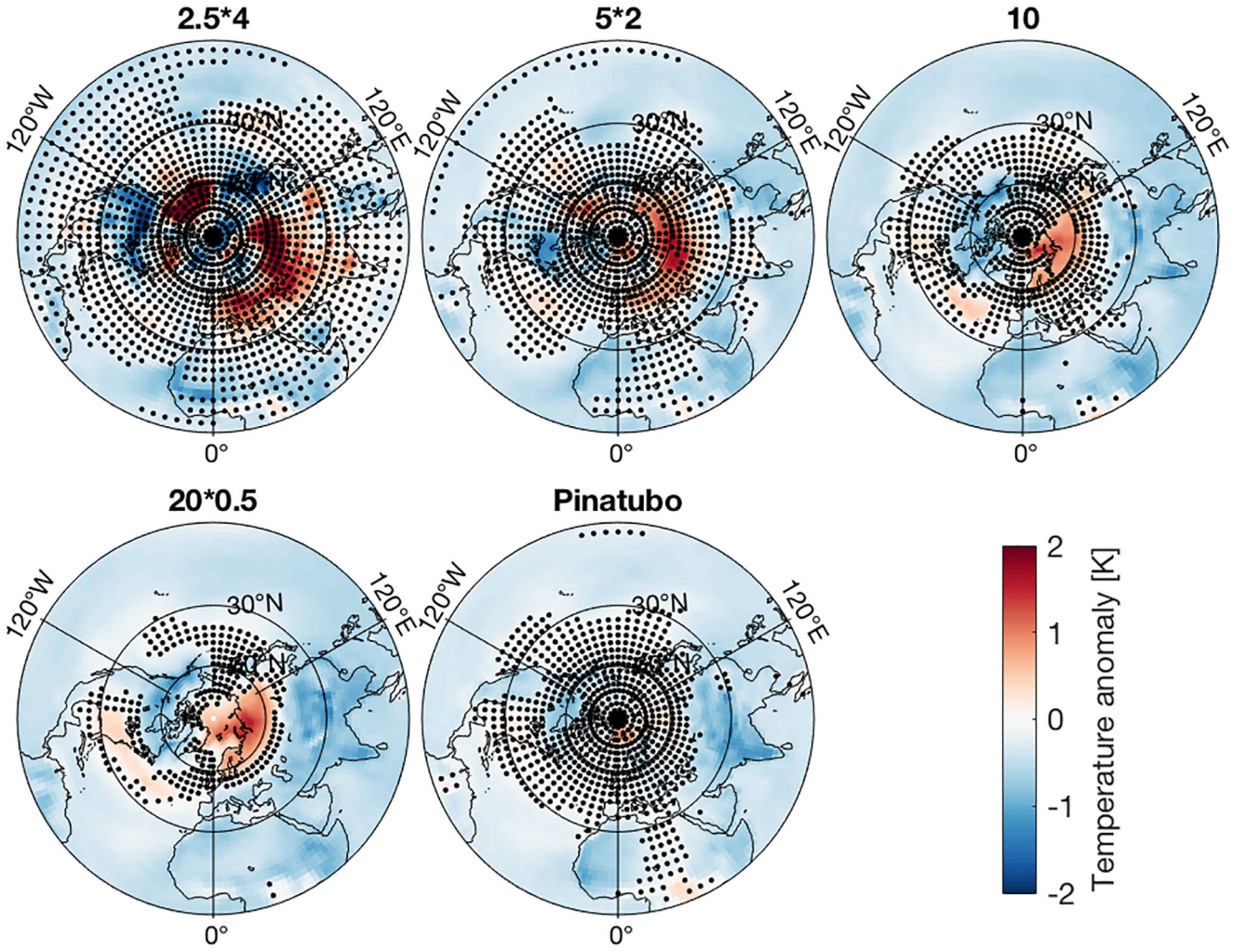

Figure 9. North polar stereographic view of ensemble mean near-surface (2m) temperature (K) anomalies simulated for the first posteruption winters (DJF). From top left to bottom right panels show the anomalies from the experiments with 2.5, 5, 10, and 20 Tg(S) injections, scaled by 10/*n* where *n* is the amount of sulfur injected, and the unscaled anomalies from the first post-Pinatubo winter from the MPI-GE. In unhatched regions, the anomalies are statistically significantly different from zero at the 95% level according to a *t* test. DJF, December–February; MPI-GE, Max Planck Institute Grand Ensemble.

3.3. Surface Temperature Anomalies

As said in Section 1, in several studies, it was argued that the positive surface temperature anomalies observed over North America and in particular Northern Eurasia in the winters after the Pinatubo eruption were related or even caused by this event via the stratospheric pathway. We will concentrate here on the more frequently discussed Northern Eurasia winter warming which was also identified from proxies for other large historical eruptions. Polvani et al. (2019), however, argued that "the internal variability of the climate system in the NH in wintertime is much larger than any impact from the Pinatubo eruption." This was supported by results from ensemble simulations (13–50 members) with the state-of-the-art WACCM4, CAM5-LE, and CAN-ESM2 climate models. Ensemble mean surface temperature anomalies in all three models are small and statistically not significant in the first post-Pinatubo winter over Northern Eurasia. The variability of all models is, however, so large that the observed anomaly is well covered by the ensembles. Polvani and Camargo (2020) came to similar conclusions based on large ensemble simulations of the Krakatoa eruption in 1883.

Figure 9 shows that in our case the ensemble mean DJF anomalies for the 10 and 20 Tg(S) experiments are statistically significantly positive in parts of the high northern latitudes over Eurasia (and negative over parts of northern North America). The two eruptions with weaker injections show even larger positive

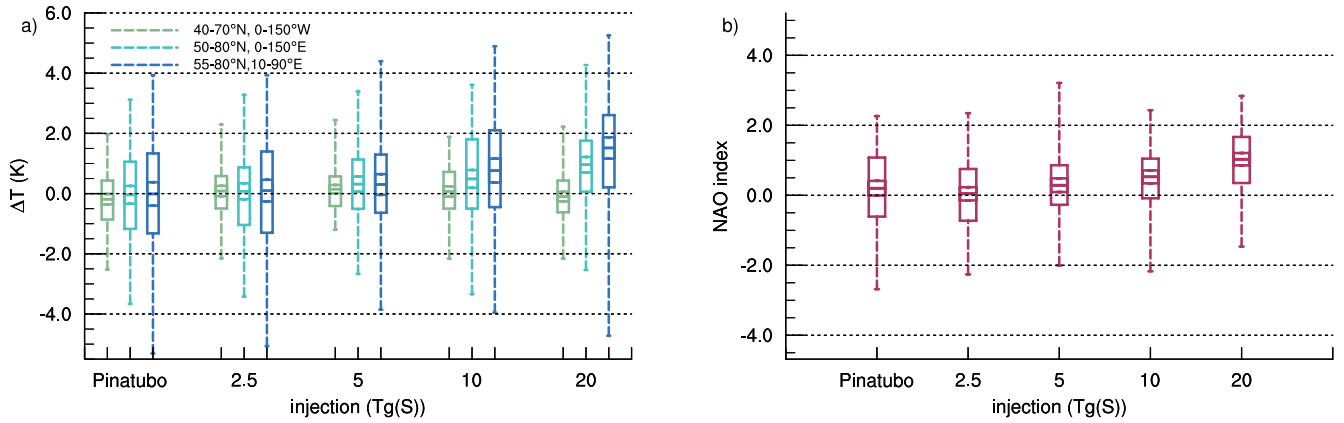

Figure 10. Box-and-whisker plots of simulated anomalies in first posteruption winters (DJF) of the experiments for Pinatubo and injections from 2.5, 5, 10, and 20 Tg(S). The whiskers show the ensemble members with the minimum and maximum anomalies, the lower and upper limits of the boxes show the 25th and 75th percentiles, the central line inside the box indicates the ensemble mean, and the lines adjacent to the mean mark twice the standard error. (a) Near-surface (2 m) air temperature anomalies (K) over different regions of Northern Eurasia (olive: 40°N–70°N, 0°W–150°W; turquoise: 50°N–80°N, 0°E–150°E; blue: 55°N–80°N, 10°E–90°E). (b) NAO index anomalies calculated as the first principle component of an EOF analysis of North Atlantic sea level pressure as described by Banerjee et al. (2021). DJF, December–February; NAO, North Atlantic Oscillation.

anomaly peaks over Northern Eurasia which are, however, not statistically significant at the 95% level. The upscaling of the anomalies, for example, by a factor of 4 in the case of the 2.5 Tg(S) injection, may lead to a prominent appearance of anomalies not necessarily attributable to the forcing. Analysis of individual winter months (not shown) indicates very similar patterns for all winter months for the 10 and 20 Tg(S) experiments, which we interpret as a further indication of the robustness of the signal for relatively large eruptions. Intrawinter variability is much higher for the anomalies from the two experiments with smaller injections. The ensemble mean anomaly simulated for the first post-Pinatubo winter is small and not significant all over Northern Eurasia, which is in line with the simulations presented by Polvani et al. (2019).

Figure 10a indicates the intraensemble variability of the posteruption surface temperature anomalies for Northern Eurasian regions of different size by box-whisker plots. The largest region (40°N–70°N, 0°W–150°W; marked in olive) is chosen as in the analysis from Polvani et al. (2019) (their Figure 4) and the results for this region are very similar for all injection magnitudes and very similar to those of the three models presented by Polvani et al. (2019). Ensemble mean anomalies are close to zero and similar numbers of ensemble members have positive and negative anomalies. Shifting the region northward and reducing it in longitudinal extent (to better capture the spatial pattern of warming shown in Figure 9) leads to an increase in variability but also a shift of the distribution to more positive anomalies. For the cases of 10 and 20 Tg(S), and both smaller regions (50°N–80°N, 0°E–150°E, and 55°N–80°N, 10°E–90°E), the mean of the temperature anomalies is significantly positive, indicated by the range of twice the standard error around the mean not covering zero. In the 5 Tg(S) case, the lower limit of this range is very close to zero for all regions.

This dependence of the signal on the choice of the region is in contrast to the simulation of the eruptions of Krakatoa (Polvani & Camargo, 2020) and Pinatubo (included, e.g., in our Figure 10) in the MPI-ESM Grand Ensemble where the insignificance of the temperature response is not sensitive to the choice of the region. As indicated already in the maps of Figure 9, our model has a tendency toward a winter warming in Eurasia north of about 50°N for EVA forcing caused by injections of about 5 Tg(S), that is, less than assumed for Pinatubo, and higher. However, due to the high internal variability of winter temperatures in this region, in a nonnegligible number of cases even very large eruptions could be followed by colder than average winters. Different to what is reported by Polvani et al. (2019) for the larger region and the WACCM simulations of the Pinatubo eruption, our simulations show a positive correlation between anomalies of stratospheric vortex strength and Northern Eurasian surface temperature as shown in Figure 11. The correlation calculated over all simulations is small ($R^2 \approx 0.1$) but highly significant due to the large number of experiments, indicating the possibility of a top-down influence, but not proving it.

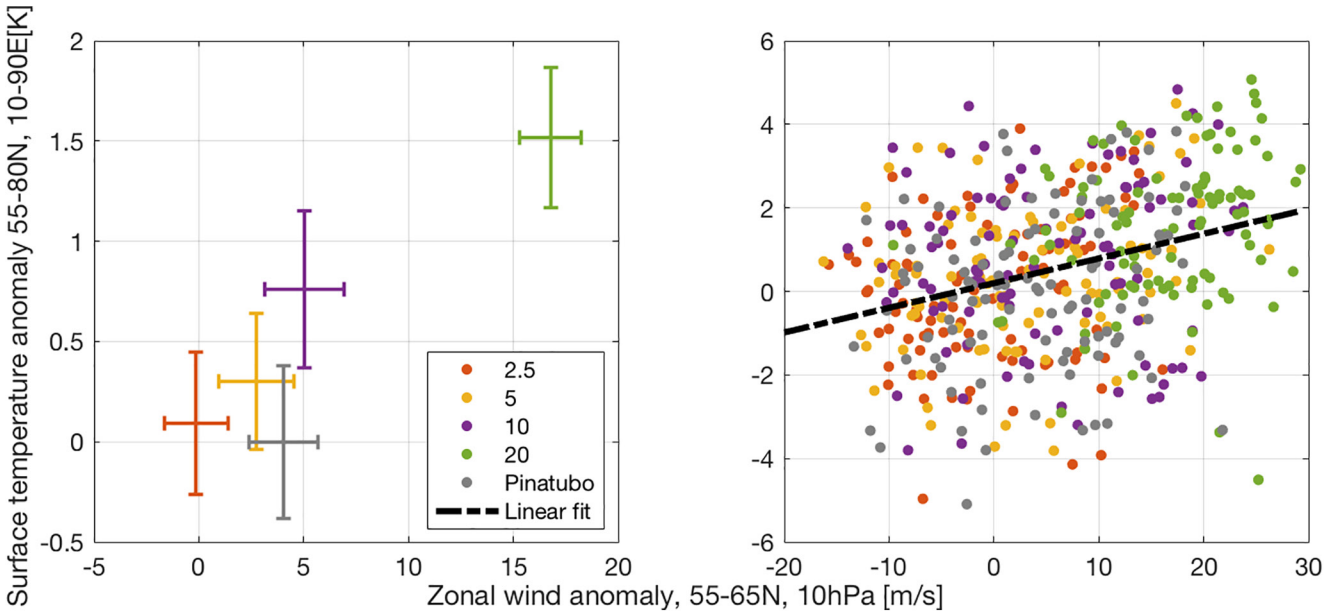

Figure 11. Simulated near-surface (2 m) air temperature anomalies averaged over 55°N–80°N and 10°E–90°E versus polar vortex zonal mean zonal wind anomalies at 10 hPa averaged over 55°N–65°N in first posteruption winters (DJF). Right panel: individual ensemble members of the EVA-ENS experiments with injections of 2.5 (red), 5 (yellow), 10 (purple), and 20 (green) Tg(S) and the Pinatubo experiment (gray). The dashed line is the result of a linear regression including all shown experiments. Left: Ensemble mean values with error bars indicating twice the standard error. DJF, December–February.

All figures in this section show a surface response to the Pinatubo forcing which is indistinguishable from zero, in contrast to the significantly positive responses in the high-injection EVA ensembles. The small difference in vortex anomaly of about 4 m/s for the Pinatubo and about 5 m/s for the 10 Tg(S) experiment makes it difficult to understand why the 2 standard error ranges of the surface responses are only very weakly overlapping for these two experiments. It is usually assumed that the Eurasian winter warming is linked to a shift toward a more positive phase of the North Atlantic Oscillation (NAO). Banerjee et al. (2021) discuss this, for example, for the occurrence of vortex strengthening and winter warming in sulfate geoengineering experiments. Figure 10b shows that also in our experiments with injections between 5 and 20 Tg(S) a statistically significant positive anomaly of the NAO index is simulated. In the 2.5 Tg(S) case, the average NAO index anomaly is close to zero, consistent with the negligible vortex anomaly. For the Pinatubo case which has a positive vortex anomaly statistically significant at more than 95% a positive NAO index anomaly is simulated but significant only at the 90%, not 95% level. Still, there is zero Eurasian winter warming, as discussed before.

Unfortunately, our analysis cannot reveal why the positive vortex anomaly simulated for the Pinatubo reconstruction is propagating downward less strongly to influence the NAO than, for example, shown in the 5 Tg(S) case, and the signal is completely lost in terms of Eurasian winter temperatures. A hypothesis to explain the different responses between Pinatubo and idealized experiments is that the differences of the aerosol distributions as discussed in Section 2.4 could cause the differences in the propagation of the dynamical signals. An importance of the specifics of an aerosol distribution was already suggested by Toohey et al. (2014).

4. Summary and Conclusions

We have analyzed large (100-member) ensembles of climate model simulations of tropical volcanic eruptions with stratospheric sulfur injections between 2.5 and 20 Tg(S). The idealized prescribed stratospheric aerosol optical properties have been constructed with the EVA forcing generator. The main motivation for this approach was that the analysis of simulations using reconstructed aerosol distributions from historical eruptions does not allow to identify if differences between responses are related to differences in strength of the eruption or differences in the distribution of the aerosol cloud that is in particular influenced by location

and time of an eruption. Our EVA-generated aerosol distributions have the important advantage that even for very different eruption strengths they are differing mostly in terms of magnitude of extinction rates, but not in the spatial pattern and time evolution.

A central result of our analysis is that for injection amounts of 5–20 Tg(S) the MPI-ESM simulates statistically significant ensemble mean, that is, forced, boreal winter polar vortex strengthenings (Figures 4a and 6). For the small injection of 2.5 Tg(S), the vortex response is indistinguishable from zero. Individual ensemble members can also show weak vortex anomalies for all eruption strengths.

Moreover, we have shown that the vortex response for the 2.5 Tg(S) eruption is significantly weaker than for larger eruptions, even if linearly scaled to the larger eruption (Figure 7), which suggests the existence of a threshold for the forced response of the Arctic stratospheric polar vortex, implying a nonlinearity of this response. This corroborates an earlier finding of Bittner, Schmidt, et al. (2016) gained from the simulation of historical eruptions that the inclusion of small eruptions in an analysis of circulation responses to volcanic aerosol distribution makes the identification of signals more difficult. Adding to this result, we have now shown that at least for the MPI-ESM this is clearly influenced by the size of the eruptions, even if location and timing of them were identical. However, our results suggest that a statistically significant vortex strengthening can be simulated already for an eruption of about half the strength of the Pinatubo eruption.

The exact cause for the existence of a threshold remains somewhat unclear. Moreover, we cannot say if the threshold limits a range of a zero forced response or if it just represents a detection threshold for the given ensemble size. Part of the nonlinearity can be explained by the stronger than linear increase of infrared AOD with the injection amount. However, also the weakest injection of our simulations (2.5 Tg[S]) causes a statistically significant warming in the tropical lower stratosphere, and a statistically significant westerly anomaly of zonal winds in the subtropical lower stratosphere (Figures 4d and 6). Bittner, Timmreck, et al. (2016) argued that this early winter zonal wind change at low latitudes would allow a more equatorward planetary wave propagation and subsequently a vortex strengthening over the course of the winter. Our analysis shows a poleward-directed EP-flux anomaly in the lower midlatitude stratosphere for all injection amounts from late autumn to early spring. However, in the 2.5 Tg(S) ensemble mean, such EP-flux anomalies are statistically significant only in a very small altitude range (Figure 8). Furthermore, the occurrence of westerly anomalies in the vortex region happens later the smaller an injection is. We conclude that for an injection as small as 2.5 Tg(S), the wave propagation is not altered sufficiently to noticeably affect the polar vortex.

A polar vortex strengthening could be considered of mainly academic interest. However, earlier studies have argued that the stratospheric Arctic polar vortex can influence surface weather on a variety of time scales through dynamical downward coupling (e.g., Kidston et al., 2015). More specifically, anomalously warm winters in Northern Eurasia as observed after the Pinatubo eruption have been linked to polar vortex strengthenings induced by volcanic stratospheric sulfate aerosol. Polvani et al. (2019) questioned this, arguing that the internal variability at this time and region is too large for an individual event to be unambiguously attributed to a volcanic eruption, and showed that in ensemble simulations with three comprehensive climate models the mean winter temperature anomaly spatially averaged over 40°N–70°N and 0°E–150°W is close to zero. While this is true also for our simulations, the results are different for smaller, more northerly regions. Near-surface air temperature averaged over 55°N–80°N and 10°E–90°E has, for example, a warm anomaly of about 0.8 K in the ensemble mean for an injection of 10 Tg(S), that is, only slightly larger than estimates for the Pinatubo eruption. While many individual ensemble members show negative anomalies even for this small region, the probability of warmer than normal winter temperatures clearly increases with the strength of the injection. It is interesting to note that the same model produces no statistically significant winter warming for a simulation of the Pinatubo eruption using an aerosol distribution prescribed following the Stenchikov climatology. This suggests that not only the amount of injected sulfur but also specifics of the aerosol pattern may be relevant for the downward propagation of a stratospheric anomaly which we think mandates further research to be better understood.

A central disadvantage of studies like this is that results are model dependent. The polar vortex response to greenhouse gas forcing, for example, has been shown to differ among models (Manzini et al., 2014). While this particular model is in many respects relatively typical for state-of-the-art CMIP-type climate models,

it is known to neglect some processes, for instance it uses prescribed instead of interactive aerosols and chemistry. In particular, interactive ozone has been shown to influence circulation responses to stratospheric events (Haase & Matthes, 2019; Lin et al., 2017; Romanowsky et al., 2019) and to alter the temperature response to volcanic aerosols (Muthers et al., 2015). To our understanding, however, it remains to be shown if interactive ozone is crucial to realistically represent the circulation response to volcanic eruptions. Additionally, the increase of model resolution could improve model dynamics. Here, we are thinking specifically of a higher vertical resolution that would enable the simulation of a QBO which is known to impact the polar vortex and to interact with the aerosol distribution (Niemeier & Schmidt, 2017). Due to computing time limitations, the chosen model configuration can be seen as a compromise necessary to produce a large ensemble. The next big volcanic eruption, as well observed as it may be, will not answer the question of what the forced circulation response to an eruption is. Natural variability of both features, the Arctic stratospheric polar vortex and near-surface air temperatures in Northern Eurasia, is large, and it is likely that the next posteruption winter will remain inside the observed year-to-year variability. Further proxy studies for historical eruptions may help to improve the identification of surface temperature responses although they include large uncertainties themselves. We argue that large ensemble simulations with other and improved climate models are another option to make progress. The CMIP6 Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP, Zanchettin et al., 2016) could provide an important step in this direction.

Data Availability Statement

For obtaining the code of the MPI-ESM (we have used version 1.1.00p2), please follow the instructions at https://mpimet.mpg.de/en/science/modeling-with-icon/code-availability. Information on how to access the MPI-GE output (used here for the Pinatubo eruptions) are available: https://mpimet.mpg.de/en/grand-ensemble. Further data and scripts used in the analysis and other supplementary information that may be useful in reproducing this work are archived by the Max Planck Institute for Meteorology under http://hdl.handle.net/21.11116/0000-0007-8B38-E.

Acknowledgments
The authors thank M. Toohey for support using the EVA forcing generator. The research has been supported by the Deutsche Forschungsgemeinschaft Research Unit VolImpact (FOR2820 Grant 398006378) within the projects VolDyn and VolClim. Computations were performed at the Deutsches Klimarechenzentrum (DKRZ).

References

Andrews, D. G., Leovy, C. B., & Holton, J. R. (1987). *Middle atmosphere dynamics*. San Diego, CA: Academic Press.

Banerjee, A., Butler, A. H., Polvani, L. M., Robock, A., Simpson, I. R., & Sun, L. (2021). Robust winter warming over Eurasia under stratospheric sulfate geoengineering—The role of stratospheric dynamics. *Atmospheric Chemistry and Physics Discussions*, *21*, 6985–6997.

Bittner, M., Schmidt, H., Timmreck, C., & Sienz, F. (2016). Using a large ensemble of simulations to assess the northern hemisphere stratospheric dynamical response to tropical volcanic eruptions and its uncertainty. *Geophysical Research Letters*, *43*, 9324–9332. https://doi.org/10.1002/2016GL070587

Bittner, M., Timmreck, C., Schmidt, H., Toohey, M., & Krüger, K. (2016). The impact of wave-mean flow interaction on the northern hemisphere polar vortex after tropical volcanic eruptions. *Journal of Geophysical Research: Atmospheres*, *121*, 5281–5297. https://doi.org/10.1002/2015JD024603

Charlton-Perez, A. J., Baldwin, M. P., Birner, T., Black, R. X., Butler, A. H., Calvo, N., et al. (2013). On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. *Journal of Geophysical Research: Atmospheres*, *118*, 2494–2505. https://doi.org/10.1002/jgrd.50125

Cionni, I., Eyring, V., Lamarque, J. F., Randel, W. J., Stevenson, D. S., Wu, F., et al. (2011). Ozone database in support of CMIP5 simulations: Results and corresponding radiative forcing. *Atmospheric Chemistry and Physics*, *11*(21), 11267–11292.

Crowley, T. J., & Unterman, M. B. (2013). Technical details concerning development of a 1200 yr proxy index for global volcanism. *Earth System Science Data*, *5*(1), 187–197.

Driscoll, S., Bozzo, A., Gray, L. J., Robock, A., & Stenchikov, G. (2012). Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions. *Journal of Geophysical Research*, *117*, D17105. https://doi.org/10.1029/2012JD017607

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., et al. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, *5*, 572–597. https://doi.org/10.1002/jame.20038

Graf, H.-F., Li, Q., & Giorgetta, M. A. (2007). Volcanic effects on climate: Revisiting the mechanisms. *Atmospheric Chemistry and Physics*, *7*(17), 4503–4511.

Guo, S., Bluth, G. J., Rose, W. I., Watson, I. M., & Prata, A. (2004). Re-evaluation of SO_2 release of the 15 June 1991 Pinatubo eruption using ultraviolet and infrared satellite sensors. *Geochemistry, Geophysics, Geosystems*, *5*, Q04001. https://doi.org/10.1029/2003GC000654

Haase, S., & Matthes, K. (2019). The importance of interactive chemistry for stratosphere-troposphere coupling. *Atmospheric Chemistry and Physics*, *19*(5), 3417–3432.

Holton, J. R., & Tan, H.-C. (1980). The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb. *Journal of the Atmospheric Sciences*, *37*(10), 2200–2208.

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

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., et al. (2013). Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*, 422–446. https://doi.org/10.1002/jame.20023

Kidston, J., Scaife, A. A., Hardiman, S. C., Mitchell, D. M., Butchart, N., Baldwin, M. P., & Gray, L. J. (2015). Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nature Geoscience*, *8*(6), 433–440.

Kodera, K., & Kuroda, Y. (2002). Dynamical response to the solar cycle. *Journal of Geophysical Research*, *107*(D24), 4749. https://doi.org/10.1029/2002JD002224

Kremser, S., Thomason, L. W., von Hobe, M., Hermann, M., Deshler, T., Timmreck, C., et al. (2016). Stratospheric aerosol—Observations, processes, and impact on climate. *Reviews of Geophysics*, *54*, 278–335. https://doi.org/10.1002/2015RG000511

Labitzke, K., & McCormick, M. P. (1992). Stratospheric temperature increases due to Pinatubo aerosols. *Geophysical Research Letters*, *19*(2), 207–210.

Lin, P., Paynter, D., Polvani, L., Correa, G. J. P., Ming, Y., & Ramaswamy, V. (2017). Dependence of model-simulated response to ozone depletion on stratospheric polar vortex climatology. *Geophysical Research Letters*, *44*, 6391–6398. https://doi.org/10.1002/2017GL073862

Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., et al. (2019). The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *Journal of Advances in Modeling Earth Systems*, *11*, 2050–2069. https://doi.org/10.1029/2019MS001639

Manzini, E., Giorgetta, M. A., Esch, M., Kornblueh, L., & Roeckner, E. (2006). The influence of sea surface temperatures on the northern winter stratosphere: Ensemble simulations with the MAECHAM5 model. *Journal of Climate*, *19*(16), 3863–3881.

Manzini, E., Karpechko, A. Y., Anstey, J., Baldwin, M. P., Black, R. X., Cagnazzo, C., et al. (2014). Northern winter climate change: Assessment of uncertainty in CMIP5 projections related to stratosphere-troposphere coupling. *Journal of Geophysical Research: Atmospheres*, *119*, 7979–7998. https://doi.org/10.1002/2013JD021403

Marshall, A. G., Scaife, A. A., & Ineson, S. (2009). Enhanced seasonal prediction of European winter warming following volcanic eruptions. *Journal of Climate*, *22*(23), 6168–6180.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2019). Developments in the MPI-M earth system model version 1.2 (MPI-ESM1.2) and its response to increasing CO_2. *Journal of Advances in Modeling Earth Systems*, *11*, 998–1038. https://doi.org/10.1029/2018MS001400

Muthers, S., Arfeuille, F., Raible, C. C., & Rozanov, E. (2015). The impacts of volcanic aerosol on stratospheric ozone and the northern hemisphere polar vortex: Separating radiative-dynamical changes from direct effects due to enhanced aerosol heterogeneous chemistry. *Atmospheric Chemistry and Physics*, *15*(20), 11461–11476.

Niemeier, U., & Schmidt, H. (2017). Changing transport processes in the stratosphere by radiative heating of sulfate aerosols. *Atmospheric Chemistry and Physics*, *17*(24), 14871–14886.

Polvani, L. M., Banerjee, A., & Schmidt, A. (2019). Northern hemisphere continental winter warming following the 1991 Mt. Pinatubo eruption: Reconciling models and observations. *Atmospheric Chemistry and Physics*, *19*(9), 6351–6366. https://doi.org/10.5194/acp-19-6351-2019

Polvani, L. M., & Camargo, S. J. (2020). Scant evidence for a volcanically forced winter warming over Eurasia following the Krakatau eruption of August 1883. *Atmospheric Chemistry and Physics*, *20*(22), 13687–13700.

Reick, C. H., Raddatz, T., Brovkin, V., & Gayler, V. (2013). Representation of natural and anthropogenic land cover change in MPI-ESM. *Journal of Advances in Modeling Earth Systems*, *5*, 459–482. https://doi.org/10.1002/jame.20022

Robock, A. (2000). Volcanic eruptions and climate. *Reviews of Geophysics*, *38*(2), 191–219.

Robock, A. (2002). Pinatubo eruption: The climatic aftermath. *Science*, *295*(5558), 1242–1244.

Robock, A., & Mao, J. (1992). Winter warming from large volcanic eruptions. *Geophysical Research Letters*, *19*(24), 2405–2408.

Romanowsky, E., Handorf, D., Jaiser, R., Wohltmann, I., Dorn, W., Ukita, J., & Rex, M. (2019). The role of stratospheric ozone for Arctic-midlatitude linkages. *Scientific Reports*, *9*(1), 7962.

Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., et al. (2011). Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0). *Geoscientific Model Development*, *4*(1), 33–45.

Schmidt, H., Rast, S., Bunzel, F., Esch, M., Giorgetta, M., Kinne, S., et al. (2013). Response of the middle atmosphere to anthropogenic and natural forcings in the CMIP5 simulations with the Max Planck Institute Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*, 98–116. https://doi.org/10.1002/jame.20014

Shindell, D. T., Schmidt, G. A., Mann, M. E., & Faluvegi, G. (2004). Dynamic winter climate response to large tropical volcanic eruptions since 1600. *Journal of Geophysical Research*, *109*, D05104. https://doi.org/10.1029/2003JD004151

Stenchikov, G., Hamilton, K., Stouffer, R. J., Robock, A., Ramaswamy, V., Santer, B., & Graf, H.-F. (2006). Arctic oscillation response to volcanic eruptions in the IPCC AR4 climate models. *Journal of Geophysical Research*, *111*, D07107. https://doi.org/10.1029/2005JD006286

Stenchikov, G., Robock, A., Ramaswamy, V., Schwarzkopf, M. D., Hamilton, K., & Ramachandran, S. (2002). Arctic oscillation response to the 1991 Mount Pinatubo eruption: Effects of volcanic aerosols and ozone depletion. *Journal of Geophysical Research*, *107*(D24), 4803. https://doi.org/10.1029/2002JD002090

Stenchikov, G. L., Kirchner, I., Robock, A., Graf, H.-F., Antuña, J. C., Grainger, R. G., et al. (1998). Radiative forcing from the 1991 Mount Pinatubo volcanic eruption. *Journal of Geophysical Research*, *103*(D12), 13837–13857.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., et al. (2013). Atmospheric component of the MPI-M Earth system model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, *5*, 146–172. https://doi.org/10.1002/jame.20015

Timmreck, C. (2012). Modeling the climatic effects of large explosive volcanic eruptions. *Wiley Interdisciplinary Reviews: Climate Change*, *3*(6), 545–564.

Toohey, M., Krüger, K., Bittner, M., Timmreck, C., & Schmidt, H. (2014). The impact of volcanic aerosol on the northern hemisphere stratospheric polar vortex: Mechanisms and sensitivity to forcing structure. *Atmospheric Chemistry and Physics*, *14*(23), 13063–13079.

Toohey, M., Stevens, B., Schmidt, H., & Timmreck, C. (2016). Easy volcanic aerosol (EVA v1.0): An idealized forcing generator for climate simulations. *Geoscientific Model Development*, *9*(11), 4049–4070.

Zambri, B., & Robock, A. (2016). Winter warming and summer monsoon reduction after volcanic eruptions in Coupled Model Intercomparison Project 5 (CMIP5) simulations. *Geophysical Research Letters*, *43*, 10920–10928. https://doi.org/10.1002/2016GL070460

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., et al. (2016). The model intercomparison project on the climatic response to volcanic forcing (VolMIP): Experimental design and forcing input data for CMIP6. *Geoscientific Model Development*, *9*(8), 2701–2719.