
The authors' replies to the referees comments can be found *in italic* below each comment.

REPLIES TO REFEREE # 2:

This manuscript argues that Mt. Pinatubo's 1991 eruption had little to no impact on continental surface temperatures, and hence observed surface warming in midlatitude winters was due to natural variability. The implication is that similar conclusions may hold for other eruptions. Indeed, there has been perhaps excessive focus on explaining and simulating the observed Pinatubo response, without sufficient regard to the inherent role of natural variability.

This manuscript contributes to ongoing discussion by frankly addressing the issue of natural variability, and its novel employment of a coupled atmosphere-ocean-chemistry model for volcanic simulation is a useful addition to the literature, even if having only 13 members for that model leads to difficulties with statistical significance. After tempering the overall claims and investigating a lower-stratospheric pathway as detailed below, this manuscript is suitable for publication.

1. General comments:

1. The text is too quick to dismiss temperature reconstructions. Despite the inherent uncertainties of temperature reconstruction, the key is that averaging over several centuries reveals a statistically significant pattern of winter warming, apparently even stronger for the subsequent winter (Fischer et al., Geophys. Res. Lett. 2007), which would be highly coincidental if volcanic eruptions were unrelated.

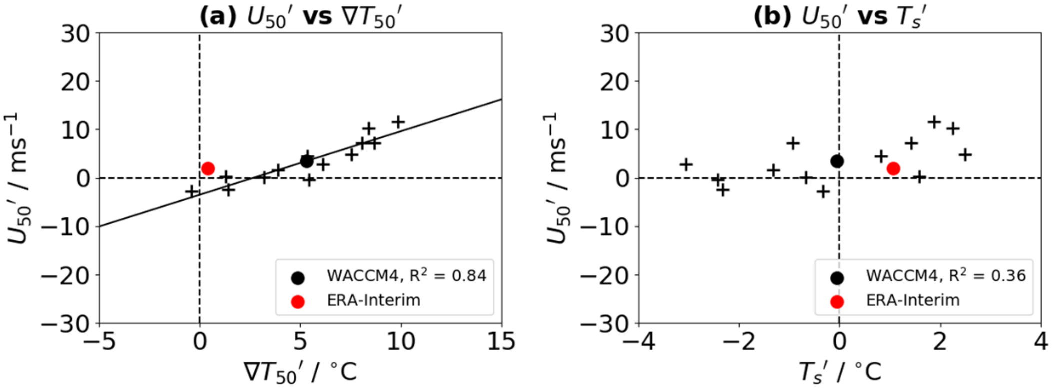
> *Thanks for the suggestion: we now cite Fischer et al (2007), with the appropriate caveats.*

2. P5 L19 and elsewhere describes the ensemble sizes as "large" (13, 42, and 50 members). However, P3 L12 mentions that Bittner et al. (2016) needed 60 members for 95% confidence in the stratospheric response, and other comparable examples are mentioned. Thus "large ensemble" seems incorrect a priori, so the difficulties throughout in achieving significance should not be surprising.

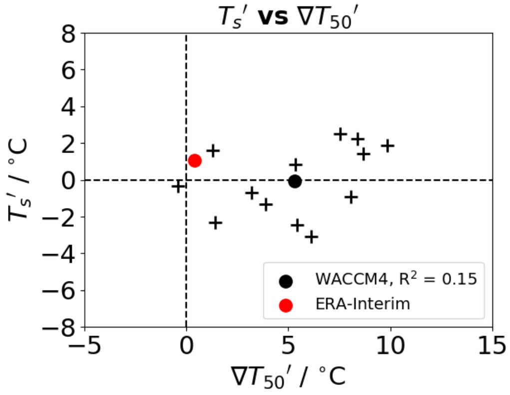
> *The term "large ensemble" is widely used in the literature to refer to the both the CAM5-LE and CanESM2 datasets. We noted in the manuscript that our relatively modest 13-member WACCM ensemble might not be best described as "large": however, it is the largest ensemble of chemistry-climate model integrations studied to date. This is why we use the term here.*
>
> *As for the lack of a surface temperature response: we find none in our 13-member WACCM ensemble, and neither did Bittner et al (2016) in their larger 100-member ensemble (see figure 6.4 of this thesis). And neither did Driscoll et al (2012) using many CMIP5 model runs. So the "difficulties in achieving significance" are very robust result across the literature, and have little to do with the relatively small size of our WACCM ensemble.*

3. To address the underlying mechanism(s), the manuscript uses a 13-member ensemble with an improved stratosphere to argue against a pathway between stratospheric vortex perturbations and surface temperature perturbations. However, the discussion does not address the possibility of a lower-stratospheric pathway, which from Fig. 7 seems plausible. This could be important if the impact on the vortex does not have the same vertical structure as natural variability. Addressing this would be straightforward by repeating the analysis of Fig. 5 at lower levels such as 50 hPa.

> *Following the referee's suggestion, we have recomputed Fig. 5 using U at 50 hPa (see below). It is very similar to the one using U at 10 hPa, which we included in the paper. The mean surface temperature response at the surface is zero (the black dot in the right panel), with 7 members showing cooling and 6 members showing warming.*



> *Furthermore, we combine the two scatter plots into one, to directly show that there is no connection between the temperature gradient at 50 hPa (which is affected by volcanic aerosols) and the Eurasian surface temperature (which is not). This is shown below.*



> *We hope this will suffice to convince the referee that the wind and temperature-gradient anomalies in the stratosphere (whether upper or lower) are not the cause of the surface temperature anomalies. This is the key point of our paper: there is no mechanism to be explained. Internal atmospheric variability suffices to produce the surface anomalies.*

2. Specific comments:

P1 L17: "is likely to be very small" is unclear – it was likely very small based on these model results?

> *Thank you for the suggestion. We have added "in our models" to clarify the sentence.*

P2 L1: "short lived" is relative; P6 L12 cites an e-folding time of 12 months, longer than most natural modes of variability.

> *We are using the expression "short lived" as it refers to a forcing of the climate system. The volcanic forcing is short-lived compared to anthropogenic forcings (e.g. the multi-decadal increase in carbon dioxide) or the solar forcing (whether the 11-year solar cycle or longer fluctuations of the solar constant).*

P2 L16–25: The suggestive tone is biased. In the "widely believed" (P1 L1) viewpoint of winter warming, it should not be "remarkable" that the subsequent winter after Pinatubo "happened to be" warm, nor is it "highly perplexing" and "difficult to reconcile" that historical warming is not exactly correlated with eruption magnitude. Rather, the question is whether eruptions of a given strength can induce a statistically and physically significant winter warming. Missing here is Fischer et al. (Geophys. Res. Lett. 2007), which should be cited here as the most recent (known to this reviewer) post-eruption temperature reconstruction.

> *As mentioned above, we have now cited the Fischer et al (2007) paper, with the appropriate caveats. As for the biased tone: we are simply saying that one expects surface cooling from a strong volcanic eruption, so any surface warming is surprising to us. Perhaps we are too naive, but naivete is not bias: we are simply offering our viewpoint.*

P3 L5: An implicit assumption of this mechanism is that the balanced acceleration lies in the vortex region, which is not necessarily the case. Bittner et al. (J. Geophys. Res. Atmos. 2016) discusses this.

> *That may be the case, but we are simply summarizing the "widely believed" viewpoint as stated, e.g. by Robock (Science, 2002), where one can read:* The polar vortex is strengthened by lower stratosphere warming at low latitudes, which is caused by absorption of solar and terrestrial radiation by the volcanic aerosol cloud.

P3 L13: "tiny" is relative; perhaps relate this to annular mode, standard deviation of lowpass-filtered winds, or similar. See also comment for P9 L13–15.

> *We have rephrased the sentence.*

P3 L19: Stenchikov et al. (2002) had only 4 ensemble members, and argued for a reduction in planetary wave activity, contrary to what Graf et al. (2007) said about a single eruption. Perhaps this paragraph should conclude that the mechanism is not demonstrated by these single-model, small-ensemble studies. Remove “clearly”, “abundantly clear,” etc. for P3 L21, P7 L24, P8 L5, P9 L34, P10 L11, P10 L29, P11 19, P12 L22. The word doesn’t enhance the argument and may come across as proof by intimidation.

> *We have removed most instances of the word “clearly” where suggested by the reviewer.*

P4 L11–14: should also cite Barnes et al. (J. Clim. 2016), which does find a significant response. Thus even when experimental design and intermodel spread are controlled for, the result can still vary!

> *Thank for the suggestion. However, Barnes et al. (J. Clim. 2016) find that the CMIP5 model response of the circulation in the NH projects very poorly on the annular mode (which contradicts with the originally proposed mechanism); also they do not explicitly examine Eurasian temperatures in wintertime.*

P5 L30–P6 L7: a common limitation, here and in many other studies, is that it is unknown whether or not the response is linear. (Perhaps a threshold magnitude of forcing is necessary, or stronger forcing induces feedbacks.) This limitation should be stated, as the conclusions for these Pinatubo-sized eruptions may not hold for smaller or larger eruptions.

> *We agree with the reviewer. We have added a sentence to that effect.*

P7 L23: ensemble averaging reduces, but does not eliminate, internal variability.

> *Thanks for noting this. We have corrected the sentence.*

P7 L25: in F3, the larger ensembles have slight windows of significance. An estimate (even via simple bootstrapping) of the necessary number of ensemble members to achieve continental-scale significance would be very helpful for this discussion and for future studies.

> *The areas of significance are non existent for WACCM and minuscule for the two low-top models. Also, recall that Bittner (2015) reported no significant warming over Eurasia, even with a 100 member ensemble. If hundreds of members are needed to produce a significant warming, the suggested bootstrapping calculation is an academic exercise.*

P7 L35: internal variability is not superimposed to any forced response – it may very well be non-linear (i.e., the higher moments of the underlying probability distribution functions may change).

> *Following Deser et al (2012), we decompose the anomalies in any one realization as a sum of a forced response (defined as the ensemble mean anomaly) and the internal variability (the difference). Hence our use of the word "superimposed". This procedure is standard, we think, in all papers that have analyzed large ensembles of model simulations.*

P7 L11 and F4: rather than a box-and-whisker plot, a plot of the 3 probability distribution functions is preferable here in my opinion, so that the reader can compare the distributions.

> *The whisker plots are PDFs. They show the mean, the percentile ranges and the full extent of each ensemble. Also, in Fig 4 the y-axis is identical: the reader can immediately and quantitatively compare these three distributions.*

P8 L25–26: the other two models may not have as accurate a representation of the stratosphere, but do they give similar results? If so, they should be included. If not, why is the subsequent comparison with Bittner et al. (2016) (which similarly has a non-interactive stratosphere) valid but not with the other two?

> *The CAM-LE and CanESM are "low-top models": as such they do not simulate the observed stratospheric variability, e.g. Stratospheric Sudden Warming events. They are, therefore, inappropriate for examining the stratospheric pathway. This is why we focus our discussion on the WACCM ensemble in the paper.*
>
> *Nonetheless, following the referee's suggestion, we now have added to the supplementary material the equivalent of Fig. 5 for the other two modes. The results are similar, and the very fact these low top models give similar results to WACCM is, of itself, a demonstration that the stratospheric pathway is not needed to explain the surface warming anomalies.*

P9 L13–15: it is not appropriate to compare weekly SSW variability with volcanic forcing as their timescales are well-separated. The appropriate comparison would be some- thing like variability of DJF average, which is approximately 10 m/s at 10 hPa and 6 m/s at 50 hPa, more comparable to the 3.5 m/s reported here.

> *We politely disagree. The point we are making is that even SSWs, which are huge disruptions of the stratospheric circulation, are often incapable of reaching the surface. Hence, it is difficult to imagine how a 1-2 m/s wind anomaly from Mt. Pinatubo would result in strong Eurasian warming. That amplitude is too small, and the accompanying surface signal is swamped by the internal variability.*

P9 L16–17 and F5: It might be helpful to add a third scatter plot of a lower comparison point like ∇T_{50} and T_s, which may correlate better than 10 hPa, if the perturbation is comparatively larger than natural variability.

> *This suggestion has been addressed above, in the "General Comments" section.*

P9 L23–34: examining two individual ensemble members does not offer any insight into the mechanism, especially since the manuscript already argues that natural variability is large. This paragraph and the corresponding F6 should be removed.

> *We politely disagree. There is much recent literature on understanding internal climate variability using large ensembles, and showing two individual members of the ensemble is the simplest and most immediate way to visually convey the importance of internal variability, which often overwhelms the forced response. This was done, for instance, in the papers below, just to cite a few examples.*
>
> - *Deser, C., R. Knutti, S. Solomon, and A. S. Phillips: Communication of the role of natural variability in future North American climate. Nat. Clim. Change (2012)*
> - *Deser, C., et al.: Projecting North American Climate over the next 50 years: Uncertainty due to internal variability. J. Climate (2014)*
> - *Deser, C., J. W. Hurrell and A. S. Phillips: The Role of the North Atlantic Oscillation in European Climate Projections. Clim. Dyn. (2017)*

P10 L7: even if 10 hPa is "canonical," it may be the wrong level for finding the mechanism. Repeating the same analysis at a lower level, such as 50 hPa, would either strengthen the current null-hypothesis argument or provide new insight into the mechanism's vertical extent.

> *This suggestion has been addressed above, in the "General Comments" section.*

P10 L14–16: "quite likely" and "would have emerged" are purely speculative and should be removed, as the null hypothesis was not rejected by the significance test. Instead, a simple bootstrapping estimate of the requisite number of samples to achieve significance may again be helpful.

> *Thanks. Bittner et al (2016) showed that a 100-member ensemble yields a significant vortex response. This is what we are referring to. We have rephrased that sentence.*

P10 L17–24: again, the possibility of a lower stratospheric pathway should, and could easily, be addressed here with the existing methodology.

> *This suggestion has been addressed above, in the "General Comments" section.*

P11 L28–30: are their low model tops thus an indirect argument for a lower stratospheric pathway?

> *To the contrary! Poor vertical resolution in the stratosphere and a low model top suppress stratospheric variability, giving the false impression that the forced response is dominant. As models have added more levels and raised the top over the years, the forced surface warming has disappeared (in the CMPI3 and CMIP5 models).*

P10 L25 to P13 L4: the conclusions should be updated following any relevant changes made as a result of these comments.

> *There have been no relevant changes we are aware of.*

Technical comments:

P8 L6: "stonger" should be "stronger"

> *Fixed. Thank you.*

P8 L11: "Fig. 5" should be "Fig. 4"

> *Fixed. Thank you.*

P9 L12: "need" should be "needed"

> *Fixed. Thank you.*

F4: "nbox" should be "box"

> *Fixed. Thank you.*

F5: "R_2 " in legend of subplot (a) should be "R^2 "

> *Fixed. Thank you.*

REPLIES TO REFEREE # 3:

The paper deals with a longstanding issue of the inability of climate models to reproduce the high-latitude near-surface winter warming following the major low-latitude volcanic eruptions. I appreciate the authors have risen this issue again using a set of the "new generation" models. I believe the reviving this issue is useful but I cannot completely agree with some interpretations and methodology the authors use in this study.

1. General comments:

To test the mechanism based on the troposphere-stratosphere dynamic interaction, the authors conducted the Pinatubo case study focusing on the first winter after the June 1991 volcanic explosion in the Philippines. However, the choice of the case-study is unfortunate as in the winter of 1991/92 the positive AO was not forced by the "stratospheric" mechanism. In observations, the polar vortex was weak and asymmetric with the wave number 2 prevailing. So, it is pointless to analyze this response to prove or disprove the stratosphere/troposphere dynamic interaction mechanism. Stenchikov et al. (2004) indicated that the easterly QBO phase in winter of 1991/92 weakened the polar vortex, and winter of 1992/93 with a westerly QBO phase provides a better case-study to test the "stratospheric" mechanism.

> *As the title of the paper makes clear, our goal is to understand what happened over the NH continents in the winter following the Pinatubo eruption, and to reconcile models and observations. Pinatubo is the largest, most recent, best observed, low-latitude eruption: as such it needs to be understood before any other, much older and poorly observed eruptions. In fact, it is routinely used as the "poster child" for the impact volcanic aerosols on NH continental temperatures, e.g. Robock (Science, 2002).*
>
> *The prevailing narrative has been that the "models are missing something" as they show no surface warming after averaging many runs (of the same or of different models). Our paper argues that this an erroneous interpretation. The model average shows no warming because* there is no significant warming response. *It's that simple.*
>
> *The referee suggests that Pinatubo is a bad choice to test the "stratospheric mechanism" because the positive AO was not forced by that mechanism after that eruption. We agree: is was not not forced by the "stratospheric" mechanism and, in fact, it was not forced by any other mechanism. It was not forced at all. It was just variability. This is what our analysis of the three large ensembles very clearly demonstrates.*
>
> *As for the possible role of the QBO. First, recent studies (Bittner et al 2016, Robock and Zambri 2016) are agreed that only the first winter after the eruption should be averaged: that is when the aerosol presence is largest. Second, the simple fact that the QBO phase can wipe out the volcanic signal confirms our claim: that internal variability (of which the QBO is one aspect) overwhelms any forced response (should one exist). With large ensembles we can actually quantify the forced response, and we find that is is small and confined to the stratosphere. This is what Bittner et al (2016) also found.*

As it is correctly stated in P8, L22-24, the "stratospheric" mechanism involves two steps: strengthening of the stratospheric polar vortex and downward propagation of the signal. The proof of the latter portion of the mechanism did not come directly from "volcanic" studies, as volcanic eruptions are rear and provide insufficient statistics, but from climatological studies of Baldwin and Dunkerton (1999). As mentioned by Stenchikov et al. (2006) the strengthening of the polar vortex caused by the equatorial lower stratospheric warming due to aerosol-induced heating, is robust in the models, but the models fail to reproduce the downward transport. So, to disprove this "stratospheric" mechanism the authors have to deal with the climatological analysis as well.

> *The strengthening of the polar vortex was* **not** *observed the first winter after the Pinatubo eruption, and yet surface warming* **was** *observed. Therefore that surface warming could not have been caused by a stronger polar vortex (see Fig. 8).*

It is not surprising that some of the model ensemble members could produce a "winter warming" pattern. It is more important how frequently this pattern appears and what mechanism causes it. Models have to produce this pattern more frequently to be consistent with the climatological studies that show a statistically significant positive AO pattern after compositing multiple equatorial eruptions. The conclusion that the up-to-date models could perfectly reproduce the winter warming based on the fact that some ensemble members capture it, is not supported.

> *The main finding of our paper is that the models do not produce the winter warming more frequently after the Pinatubo eruption, because the ensemble average is zero. Warming happens half of the time, as a simple consequence of internal variability.*
>
> *We did not conclude that "models could perfectly reproduce the winter warming based on the fact that some ensemble members capture it". Our key finding, after analyzing three large ensembles, is that the observed warming falls well within the distribution of the model members. From this we conclude that the models capture the observations.*

2. Specific comments:

P3, L13-15: A vertical propagation of the planetary waves is a threshold process as suggested by Charney and Drazin (1961), so small change of the wind could qualitatively change the planetary wave reflection coefficient.

> *This is a linear result from highly idealized theoretical studies. Whether and how it may be relevant in practice remains to be demonstrated.*

P4, L26-27: AO response is an atmospheric effect. Why increasing of model complexity should matter to answer the question that the Stratosphere-Troposphere Interaction is real? E.g., if ozone additional radiative effect matters, this has to be specifically shown.

> *The literature on the impact of stratospheric ozone the annular mode is quite large. The referee might wish to consult, e.g. Thompson et al, (Nature Geoscience, 2011).*

P5, L23: The chosen models are inconsistent in reproducing the aerosol forcing. In Figure 2 the aerosol forcing in the models differs by 50%. It would be useful to mention what was the observed forcing to compare with.

> *There is nothing peculiar about the models we have analyzed. They are of the same kind as those analyzed in Driscoll et al (2012) or Robock & Zambri (2016). In Fig. 2 we show the ERA-Interim temperature time series: one can see that WACCM and CAM5-LE simulate excessive warming, a common bias, as noted by Driscoll et al (2012).*

P5, L30-33: The winter of 1991/92 after the Pinatubo eruption is a wrong choice (see Figure 5). A "composite" approach has to be considered to obtain statistically significant anomalies in observations.

> *See our reply to this in the General Comments section above as to why Pinatubo is not the wrong choice. As for adopting a "composite" approach: we have done so, by averaging all members of each ensemble. What we find, in agreement with Driscoll et al (2012) and Bittner (2015), that the surface response is not statistically significant. There is no reason to expect, a priori, that the surface anomaly should be significant, unless one thinks that internal variability is small, which is not the case (see Fig 3).*

P6, L2: This is incorrect. The eruption of Mt Agung of 1963 developed an aerosol equatorial reservoir that caused warming of the equatorial lower stratosphere and enhanced equator-pole temperature gradient in the lower stratosphere. The re-distribution of aerosols between the hemispheres is not directly relevant.

> *We politely disagree: the hemispheric distribution is likely to matter. In any event, we have not analyzed the Mt Agung eruption, so the whole point in nugatory.*

P6, L8: The first winter is a wrong choice.

> *We politely disagree: the aerosols forcing in the stratosphere is largest in the first winter.*

P6, L12: Volcanic aerosols remain in the equatorial reservoir in the second winter after the eruption that is why the effect is seen in the second winter as well.

> *We politely disagree. As already pointed out, others have also concluded that only the first winter should be analyzed (Bittner et al 2019, Robock and Zambri 2016). The amount of volcanic aerosols in the winter 1992-93 was only a small fraction the one in 1991-92: see, for instance, Figure 3 of Stenchikov et al (JGR, 1991). It make no sense to average a large and a small forcing: that simply washes out the signal.*

P6, L18: Driscoll et all. (2012) adopted this methodology from Stenchikov et al. (2006).

> *Thank you for pointing this out. We have now added the Stenchikov et al (2006) paper.*

P7, L8–9: The shortwave (SW) radiative forcing in three chosen models differs by 50%. There is much more differences in SW and Longwave (LW) aerosol absorption.

> *Intermodel differences are not uncommon, and we have noted then.*

P7, L13–15: The models three times overestimate the equatorial lower stratospheric heating caused by volcanic aerosols. This is the main forcing of the stratosphere-troposphere dynamic interaction. There is something wrong here.

> *Yes, we are agreed. This is a well know bias in many of the current-generation models. However, the reviewer will agree that the models we have analyzed are no more biased than the ones in Stenchikov et al (2006), Driscoll et al (2012), Robock & Zambri (2016) and many other studies.*
>
> *More importantly: this model biases makes our argument stronger! Even with an overestimated equatorial lower stratospheric heating caused by volcanic aerosols, our models (and those of the other recent studies) show no statistically significant surface warming. Had the models not overestimated the stratospheric aerosol heating the surface signal would be even smaller. We have noted this in the revised version of the paper (on page 7, lines 17–19).*

P7, L20: "With this IN mind"

> *Thank you. We have corrected this.*

P7, L32–35: I think the correct question to ask is whether models are able to correctly reproduce the probability distributions of the Arctic oscillation (AO) responses to volcanic forcing. But for this one has to extract multiple cases from observations and to construct the observed probability distribution. It is not doable with only one post-eruption season considered.

> *We politely disagree. We believe that one can indeed use "only one post-eruption season" provided on has large esembles of runs available. This is what we did.*

P9, L10: Planetary wave reflection is the threshold process (Charney and Drazin, 1961) when small changes matter.

> *We do not question the Charney and Drazin (1961) result, which is based on small-amplitude linear theory in a highly idealized configuration. What we question is how relevant that particular threshold behavior is to the problem at hand. The claim that a mere 1-2 m/s acceleration of the polar vortex from volcanic aerosol heating has a major impact on planetary wave propagation is purely speculative and, to the best of our knowledge, has yet to be demonstrated. For instance, one would first need to show that the climatological conditions are found to be very near the wave propagation threshold, so that a tiny wind perturbation is able to make the system cross that threshold. Then one would have to show that the linear approximations are actually valid. And so on. We are not aware of studies which have carefully performed such work.*

P9, L14: The wind variability coming from SSW is not relevant to the process. As soon as polar vortex zonal wind weakens below the threshold, planetary waves can propagate upward nonlinearly weakening the polar vortex. So, the amplitude of wind changes below the threshold, no matter how large it is, does not count. Sampling has to focus on the strong vortex cases for this purpose.

> *See the answer to our previous point, and also the answer to the other referee.*

P10, L17–18: Exactly, the winter of 1991/92 is not suitable to study the forced stratosphere-troposphere dynamic interaction, as the positive phase of AO in the troposphere was caused by a different mechanism.

> *This point has been addressed above, in the "General Comments" section.*

P11, L2: You mean surface cooling/warming, not in the lower stratosphere. Please clarify.

> *Yes, at the surface. Thanks for pointing out this ambiguity. We have correct the text.*

P11, L5-8: You have to explain why do we see a positive AO anomaly climatologically after multiple volcanic eruptions. If this would be extremely rear events as in the models, then a positive AO anomaly would not be seen in observations.

> *First: we do not think we need to explain "why we see a positive AO anomaly climatologically after multiple volcanic eruptions". Our paper is about Mt Pinatubo and that anomaly was absent in the winter following the 1991 eruption, demonstrating* ipso facto *that it could not possibly have been responsible for the observed NH surface warming.*
>
> *Second: the evidence for "a positive AO anomaly climatologically after multiple volcanic eruptions" is not terribly robust (see, e.g. Wunderlich & Mitchell, ACP 2017).*

P11, L15–23: The strengthening of the polar vortex caused by the volcanic aerosols heating in the lower equatorial stratosphere is robust. This is the threshold process, so a weak strengthening matters. And it is unfair to apply wind variability in SSW to scale the increase in maximum wind.

> *We have addressed this comment in several locations in the discussion above.*

P11, L25–35: The downward propagation mechanism was proved using climatological analysis (Baldwin and Dunkerton, 1999) and has to be challenged on this basis.

> *We are not sure what the referee means. We are not challenging the fact that SSWs can affect the annular modes and produce surface anomalies. That result is robust. We are questioning the claims of the early papers on the NH warming following volcanic eruptions. Those papers reported significant surface responses because the models used were flawed (they lacked vertical resolution and stratospheric variability). The fact that significant surface responses are not seen in the more recent studies (which are based on much better models) supports our interpretation.*

P12, L25–30: ENSO definitely could affect surface temperatures, although Volcano-ENSO interaction is highly nonlinear. At least contribution of ENSO variability in the volcanic signal has to be removed properly, which was never done in this analysis. Another important mode of variability is QBO that was not considered and reported in this study. QBO plays an important role in stratospheric wave propagation and could directly affect polar vortex and shape the stratosphere-troposphere dynamic interaction.

> *We agree with the referee: ENSO and the QBO do affect stratospheric wave propagation. We also hope the referee agrees with us: that ENSO and the QBO are part of the internal variability of the climate system. Therefore, if their influence needs to be removed in order to detect any putative surface response to the volcanic aerosols (should it be detectable at all) it means that the volcanic surface influence in the NH is clearly masked by natural variability. This is the key point of our paper.*

Northern Hemisphere continental winter warming following the 1991 Mt. Pinatubo eruption: Reconciling models and observations

Lorenzo M. Polvani[1,2], Antara Banerjee[1], and Anja Schmidt[3,4]

[1]Department of Applied Physics and Applied Mathematics and Department of Earth and Environmental Sciences,Columbia University, New York, NY 10027, USA
[2]Lamont Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA
[3]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK
[4]Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, UK

Correspondence to: Lorenzo M. Polvani (LMP@COLUMBIA.EDU)

Abstract. It has been suggested, and is widely believed, that the anomalous surface warming observed over the Northern Hemisphere continents in the winter following the 1991 eruption of Mt. Pinatubo was, in fact, caused by that eruption, via a stratospheric pathway that involves a strengthening of the polar vortex. However, most studies that have examined multiple, state-of-the-art, coupled climate models report that, in the ensemble mean, the models do not show winter warming after the Mt. Pinatubo eruption. This lack of surface warming in the multi-model mean, concomitant with a lack of strengthening of the polar vortex, is often interpreted as a failure of the models to reproduce the observations. In this paper we show that this interpretation is erroneous, as averaging many simulations from different models, or from the same model, is not expected to yield surface anomalies similar to the observed ones, even if the models were highly accurate, owing to the presence of strong internal variability.

We here analyze three large ensembles of state-of-the-art, coupled climate model simulations and show that, in all three, many individual ensemble members are able to produce post-Pinatubo surface warming in winter that is comparable to the observed one. This establishes that current-generation climate models are perfectly capable of reproducing the observed surface post-eruption warming. We also confirm the bulk of previous studies, and show that the surface anomaly is not statistically different from zero when *averaged* across ensembles of simulations, which we interpret as the simple fact that the volcanic impact on continental winter temperatures is tiny compared to internal variability.

We also carefully examine the stratospheric pathway in our models and, again confirming previous work, show that any strengthening of the polar vortex caused by the Mt. Pinatubo eruption is ~~likely to be~~ very small (of the order of a few m/s at best). Such minuscule anomalies of the stratospheric circulation are completely overwhelmed by the tropospheric variability at mid-latitudes, which is known to be very large: this explains the lack of surface winter warming in the ensemble means.

In summary, our analysis and interpretation offers compelling new evidence that the observed warming of the Northern Hemisphere continents in the winter 1991-1992 was very likely unrelated to the 1991 Mt. Pinatubo eruption.

1 Introduction

Large, low-latitude volcanic eruptions produce considerable, albeit short lived, natural perturbations to the radiative forcing of the Earth's climate, and thus offer unique opportunities to probe its dynamics. With an estimated peak aerosol loading of 30 Tg (McCormick and Veiga, 1992), the eruption of Mt. Pinatubo in June 1991 was the largest to occur since the advent satellite observation and, in fact, the second largest over the entire 20th century (after the 1912 Novarupta eruption). Moreover, in terms of dust veil index (Robock, 2000) and stratospheric optical depth (Sato et al., 1993) it stands unrivaled all the way back to the historic eruption of Mt. Krakatau in 1883, and is therefore the premier candidate for understanding how volcanic aerosols affect the climate system.

After the initial cataclysmic eruption of June 14-15 1991, the aerosol cloud from Mt. Pinatubo spread rapidly and encircled the globe in a mere 22 days (Bluth et al., 1992) filling the entire tropical belt, both north and south of the Equator, in a couple of months (McCormick and Veiga, 1992) and then spreading to higher latitudes in subsequent months (Long and Stowe, 1994). Since volcanic aerosols are strong scatterers of incoming solar radiation, they act to cool the troposphere and the Earth's surface. By September 1992, the global lower troposphere had cooled by -0.5°C (Dutton and Christy, 1992), with an even larger cooling of -0.7°C in the Northern Hemisphere (NH). Such large cooling values are comparable to the estimates for the epochal Tambora eruption of 1815 (McCormick et al., 1995).

In the context of such widespread cooling, the surface temperature over the NH continents happened to be anomalously warm in the winter immediately following the Mt. Pinatubo eruption (Robock, 2002). In a series of papers, Groisman (1992), Robock and Mao (1992), and later Robock and Mao (1995) and Kelly et al. (1996), argued that continental winter warming also occurred following several other eruptions since 1850, and suggested that the winter NH warming was actually *caused* by the volcanic eruptions. Further observational evidence was offered by Shindell et al. (2004), who expanded the set to a dozen large, low-latitude eruptions, going back to the year 1600. Their additional evidence, however, includes some perplexing facts. For instance, they show that the continental winter warming following both the 1883 Krakatau and the 1815 Mt. Tambora eruptions is, apparently, much smaller than the one following the 1982 El Chichón eruption (see Figure 1 of Shindell et al., 2004): this is difficult to reconcile with the narrative that volcanoes are the major cause of the NH continental winter warming, since those two earlier eruptions are larger in magnitude than the later one. Finally, after analyzing European climate reconstructions over the last half millennium, Fischer et al. (2007) report that the wintertime surface temperature anomalies caused by low-latitude eruptions appear to be stronger the second post-eruption year than in the first: this puzzling result is clearly at odds with the fact that only a small fraction of the volcanic aerosols are left in the stratosphere in the second winter after an eruption.

Part of the widespread belief in the existence of a causal link between low-latitude volcanic eruptions and winter warming over the NH continents stems from the fact that a mechanism has been proposed to explain that link. Graf et al. (1993), on the basis of highly[1] idealized numerical experiments, followed by the observational studies of Kodera (1994) and Perlwitz and Graf (1995), and further numerical studies by Kirchner et al. (1999), Stenchikov et al. (2002) and many others thereafter, have

---

[1] Their model was run in perpetual January configuration, with prescribed sea surface temperatures and sea ice concentrations, forced with an "externally computed" heating rate, but without interactive aerosols or ozone chemistry modules.

advocated for the existence of what we will refer to as a “stratospheric pathway” causally linking low-latitude eruptions in summer with mid-latitude surface warming the following winter. The starting point for this mechanism is the well known fact that sulfate aerosols of volcanic origin are also strong absorbers of infrared radiation: hence powerful, low-latitude eruptions that are able to penetrate sufficiently high into the atmosphere can cause a strong *warming* of the tropical lower stratosphere, in addition to the tropospheric and surface cooling mentioned above. In the case of Mt. Pinatubo a 2-3°C warming[2] of the tropical lower stratosphere was seen in radiosonde observations (Randel, 2010), in agreement with multiple reanalyses (Fujiwara et al., 2015). Such a perturbation increases the equator-to-pole temperature gradient in the stratosphere, notably in winter, and induces a strengthening of the stratospheric polar vortex. The stronger polar vortex, it is claimed, then causes a positive phase of the North Atlantic Oscillation (or the Northern Annual Mode), which finally results in warmer surface temperatures over the NH continents, notably over Eurasia.

In spite of its simplicity, this proposed mechanism remains unconvincing because it has yet to be properly quantified. For instance, one could ask: *how large* is the polar vortex acceleration caused by an eruption comparable to the one of Mt. Pinatubo in 1991? A recent study (Bittner et al., 2016), using a very large ensemble of runs with a well-tested stratosphere-resolving model, suggests a polar vortex acceleration possibly as large as 2 m/s at 10 hPa around 60N, but also reports that even 100 model runs are insufficient to establish that fact at the 99% confidence level and if one lowers the level to 95% more than 60 runs are needed for a statistically significant 2 m/s acceleration of the polar vortex (see their Figure 2a). Moreover, the large internal variability associated with the North Atlantic Oscillation can easily overwhelm the surface effects of such a small stratospheric perturbation, as it even confounds the forced signal from increasing greenhouse gases over an entire 50-year period (see, for instance, Deser et al., 2017).

In fact, the original stratospheric pathway mechanism has been called into question, even by its original proponents. Stenchikov et al. (2002) suggested that the stratospheric pathway may be part of a more complex mechanism and, on the basis of results from a single model, proposed that an additional tropospheric pathway may be equally important. In addition, Graf et al. (2007) reported that observations actually show *increased* planetary wave activity in the winter following the Mt. Pinatubo eruption, which is ~~clearly~~ at odds with the claim of a stronger polar vortex that winter causing the NH surface warming, and completely invalidates the original mechanism. Thus, they suggest “that the climate effects of volcanic eruptions are *not* being explained by the excitation of inherent zonal mean variability modes such as Strong Polar Vortex or Northern Annular Mode, but rather is another mode that possibly reflects upon the North Atlantic Oscillation” (Graf et al., 2007).

Furthermore, one can find in the literature many modeling studies whose findings are often diametrically opposite to each other. We will not exhaustively cite all previous papers, but simply limit ourselves to highlighting a few key studies to illustrate the contradictory claims that can be found in the peer-reviewed literature. Let us start by summarizing the findings of Driscoll et al. (2012), who analyzed 13 models from the Climate Model Intercomparison Project, Phase 5 (CMIP5). These models were specifically selected so as to have at least two ensemble members available. Comparing the average across all the models, as well as the averages across all the members of the each model, they concluded that “none of the models manage to simulate a sufficiently strong dynamical response,” given the absence of NH continental warming following the Mt. Pinatubo eruption in

---

[2] At levels close to 20 km, taking the one-year mean after the eruption minus the mean of the preceding three years.

the model averages. Their study confirms the earlier conclusion reached with the CMIP3 models (Stenchikov et al., 2006), and many other studies (e.g. Thomas et al., 2009; Marshall et al., 2009; Bittner, 2015; Wunderlich and Mitchell, 2017).

Against this body of evidence, analyzing two version of the NASA/GISS model, Shindell et al. (2004) have claimed that "driven by solar heating induced by the stratospheric aerosols, these models produce enhanced westerlies from the lower stratosphere all the way to the surface" and a significant wintertime warming over the NH continents, in agreement with Graf et al. (1993) and Kirchner et al. (1999), who also claimed that climate models are able to simulate the continental winter warming following the Mt. Pinatubo eruption via the stratospheric pathway. In fact, Shindell et al. (2004) concluded that their results "provide a further strong indication of the critical role of the stratosphere in the dynamic response to external forcing," with a suggestion that a well resolved stratosphere is crucial for capturing the NH winter warming that would be caused by volcanic eruptions. That suggestion, however, would seem soundly refuted by the evidence presented in Charlton-Perez et al. (2013), who separately analyzed models with and without a well-resolved stratosphere, and showed no difference between the two sets in the forced response of the polar vortex in the winter following volcanic eruptions.

And lastly, Zambri and Robock (2016) reanalyzed the CMIP5 models using a different methodology. Averaging only the largest eruptions, and only the first winter after those eruptions, they concluded that "most models do produce a winter warming signal, with warmer temperatures over NH continents and a stronger polar vortex in the lower stratosphere," directly contradicting Driscoll et al. (2012).

It is in the context of such multiple inconsistent claims, that our paper aims to answer two questions:

1. Are current-generation climate models able to simulate the continental winter warming in the NH following the 1991 Mt. Pinatubo eruption?

2. If so, does the stratospheric pathway proposed by Robock and Mao (1992) and Graf et al. (1993) play any role in simulating that warming?

Analyzing large ensembles of model integrations from three different state-of-the-art coupled climate models over the historical period, we show below that (1) models *are* perfectly capable of simulating NH continental warming in the winter following the Mt. Pinatubo eruption, but (2) the stratospheric pathway – and, more importantly, the Mt. Pinatubo eruption itself – very likely played *no significant role* in the occurrence of that warming.

2 Methods

2.1 The models

Three large ensembles of integrations with state-of-the-art, comprehensive climate models are analyzed in here. All our models include atmosphere, land, ocean and sea-ice components, fully coupled [3] to accurately simulate the climate system response to the Mt. Pinatubo eruption. Here are, in brief, the specifications of our three models: WACCM4, CAM5-LE, CanESM2

---

[3] Note that was mostly not the case in the earlier studies. Neither Graf et al. (1993), not Kirchner et al. (1999), nor Stenchikov et al. (2002), nor Shindell et al. (2004) used fully coupled climate models.

- WACCM4 is the Whole Atmosphere Community Climate Model, Version 4, developed by the Community Earth System Model (CESM) Project. WACCM4 is a high-top model, with a lid at 140 km and 66 vertical levels, and a horizontal resolution of $\sim 2°$. Its climate over the 20th century has been thoroughly evaluated by Marsh et al. (2013), where further details about this model may be found. We emphasize that WACCM4 also includes interactive stratosphere ozone chemistry and, therefore, has the most realistic representation of stratospheric dynamics and chemistry of the three models analyzed here.

- CAM5-LE was also developed under the CESM project, with ocean and sea ice components similar to those of WACCM4. However, the atmospheric component of CAM5-LE is very different: it is a low-top model with only 30 vertical levels but with a higher horizontal resolution ($\sim 1°$) and, most importantly, employs very different physical parameterizations than those in WACCM4 (Neale et al., 2010) and, in fact, has a considerably different climate sensitivity (Gettelman et al., 2013). CAM5-LE has been at the heart of the CESM Large Ensemble Project (see Kay et al., 2015, for details) and its performance, therefore, has been thoroughly tested in dozens of studies which have analyzed its output.

- CanESM2 is the second-generation Canadian Earth System Model, developed at the Canadian Centre for Climate Modeling and Analysis (CCCma). The atmospheric component of CanESM2 is a spectral model with an approximate horizontal resolution of 2.8°and with 35 unevenly spaced vertical levels and a model top near 0.1 hPa. For more details the reader may consult von Salzen et al. (2013). Again, this is a well-tested model which has contributed a whole suite of runs to the CMIP5 project, and it has been widely used in many climate studies (e.g. Swart et al., 2015).

We note that all three models were previously used to study the climatic effects of volcanic eruptions (English et al., 2013; Lehner et al., 2016; Gagné et al., 2017). More importantly, for all three models we have available a *large ensemble* of integrations which cover the second half of the 20th century. For these integrations, the models include all known natural and anthropogenic forcings, as per the so-called “historical” specifications of the CMIP5 protocol (Taylor et al., 2012). Specifically, we have analyzed 13 runs with WACCM4, 42 runs with CAM5-LE, and 50 runs with CanESM2. We stress that the model forcings are *identical for all members* of the same ensemble. The differences among members of the same ensemble arise uniquely from minuscule perturbations imposed on the models’ atmospheric initial conditions: the differences allow us to explore the internal variability of the system which, in many cases, can be much larger than the response to an external forcing, be it natural or anthropogenic. The reader is referred to Deser et al. (2012) for the seminal exposition of this methodology.

2.2 The analysis

We here discuss three key methodological choices we made in designing the best strategy to determine whether current-generation climate models are able to capture the wintertime NH continental warming following volcanic eruptions.

1. *Choice of eruption.* Although the model runs available to us cover the 1963 Agung and 1982 El Chichón eruption, we will here focus solely on the 1991 eruption of Mt. Pinatubo, in view of the following. First, as already noted, that eruption is the best observed of all known eruptions, and thus offers the best opportunity to contrasting models and observations.

Second, one can easily argue that every eruption is unique: for instance, while the aerosol cloud from Pinatubo spread out in both hemispheres, the one of Mt. Agung spread primarily into the Southern Hemisphere (Viebrock and Flowers, 1968). So, combining these seems inappropriate. Third, and most importantly: since we are seeking to isolate and quantify the *forced response* to volcanic eruptions, it make no sense to average eruptions of different magnitudes. This would be tantamount to trying to estimate the Earth's climate sensitivity by averaging together 2×CO_2 and 4×CO_2 model runs. And we do not know whether the forced response varies linearly with the magnitude of an eruption. Other recent studies have also argued against averaging stronger and weaker eruptions when seeking to isolate their climatic impacts (Bittner et al., 2016; Zambri and Robock, 2016).

2. *Choice of winters.* We will here analyze only the *first* winter following the June 1991 eruption, i.e. the three month period from December 1991 to February 1992. Many (if not most) of the earlier studies assumed that the effect of volcanic eruptions can be felt for several years, and averaged together the first and second winters after each eruption. We see no cogent reason for doing so: the Mt. Pinatubo volcanic aerosols were removed from the atmosphere with an e-folding timescale of about 12-months (Barnes and Hofmann, 1997), so that the aerosol optical depth in January 1993 is much smaller than in January 1992 (see also Long and Stowe, 1994). Furthermore, if indeed the stratospheric pathway is crucial to carrying the response down to the surface at higher latitudes, it is difficult to imagine what memory the stratosphere would posses to remember in the winter of 1993 an eruption that occurred in June 1991. The recent study of Zambri and Robock (2016) also argues that only the first winter should be used, since "averaging the first two winters after each eruption may have had a damping effect."

3. *Choice of reference period.* ~~For this,~~ Here we follow the methodology of ~~Driscoll et al. (2012)~~ Stenchikov et al. (2006) and Driscoll et al. (2012), and define the winter-time anomalies after the Mt. Pinatubo eruption as the difference between 1991/1992 winter and the mean of the winters in the 1985-1990 reference period. While this need not be the best way to quantify the post-eruption anomalies, we nonetheless adopt it in order to be consistent with recent studies who analyzed models similar to ours (Bittner et al., 2016; Zambri and Robock, 2016). As we will show below, our conclusions differ significantly from those of previous studies, and we want to make it clear that the choice of reference period is not at the root of those differences.

In summary then: for all quantities in all figures below (except Fig. 2) we will be showing and discussing anomalies defined as the difference between the first winter following the June 1991 Mt. Pinatubo eruption and the reference period defined in Driscoll et al. (2012). We will refer to these as the "post-Pinatubo anomalies," or just "the anomalies" for short and, for simplicity, denote them with a prime (e.g. T_s' for the surface temperature anomalies).

3 Can climate models simulate the observed NH continental warming following the 1991 Mt. Pinatubo eruption?

It is useful to start by recalling what the observed wintertime, post-Pinatubo, surface temperature anomalies over the NH continents actually look like. They are shown in Fig. 1, from four different datasets: two observational ones, GISSTEMP

(Hansen et al., 2010) and HadCRUT4 (Morice et al., 2012), and two reanalyses, NCEP/NCAR (Kalnay et al., 1996) and ERA-Interim (Dee et al., 2011). Note the excellent agreement between these four data products, which all show warming over both North America and the Eurasian continent. The fact that *both* continental masses were anomalously warm, is of relevance for the stratospheric pathway mechanism to be discussed in the next section. These anomalies are also in excellent agreement with the lower tropospheric temperature anomalies from the Microwave Sounding Unit, Channel 2 (MSU2) satellite observations shown by Robock (2002), albeit for a slightly different reference period.

We now turn to analyzing the models. Before showing the simulated surface temperatures, however, we wish to illustrate the models' response to the Mt. Pinatubo eruption in the stratosphere, as the warming of the tropical stratosphere is an essential component of the stratospheric pathway mechanism. The global top-of-the-atmosphere (TOA) net outgoing shortwave radiation anomalies are shown in the top row of Fig. 2, for WACCM4, CAM5-LE, and CanESM2, from left to right. These panels may be contrasted directly with Fig. 2 of Driscoll et al. (2012), as they demonstrate that our three models are comparable to most CMIP5 models.

The resulting warming of the tropical lower stratosphere (30S-30N) is shown in the bottom row of Fig. 2. The ERA-Interim reanalyses are also shown for comparison (black curves in each panel). While the CanESM2 model appears to be in good agreement with the observations, both WACCM4 and CAM5-LE greatly overestimate the post-eruption warming in the lower stratosphere. Reanalyses show an anomaly of roughly 2°C , but those models show ensemble mean anomalies closer to 6 and 9 °C , respectively. This is not exceptional, as Driscoll et al. (2012) reports that most CMIP5 models simulate a much stronger anomaly than was observed (see their Fig. 3). The interesting point, however, is that ~~we will be turning~~ this model bias can be turned to our advantage: as will become clear below, the fact that ~~the WACCM4 model, in particular, simulates a stronger than observed~~ models simulate an unrealistically strong warming of the tropical lower stratosphere after the Mt. Pinatubo eruption ~~will greatly strengthen~~ greatly strengthens our interpretation and conclusion.

With this ~~is~~ in mind, we now proceed to examine the surface temperature anomalies simulated by our three models following the Mt. Pinatubo eruption, shown in Fig. 3. It is important to keep in mind that for each ensemble member the post-Pinatubo anomalies arise from two distinct sources: the external forcing and internal variability. The former is computed by averaging together all the members of each ensemble, as that procedure nearly eliminates the internal variability. For WACCM4, CAM5-LE and CanESM2, the left column of Fig. 3 shows that forced response. ~~It is abundantly clear that in~~ In the winter following the Mt. Pinatubo eruption, ~~the~~ all three models show *no statistically significant response* in NH continental surface temperatures.

We stress that this result is in agreement with most of the literature on this subject, notably the multi-model studies with the CMIP3 and CMIP5 models (Stenchikov et al., 2002; Driscoll et al., 2012; Wunderlich and Mitchell, 2017), which have shown that the forced post-Pinatubo anomalies in CMIP-class models are not statistically significant. Moreover, it has been validated with an even larger ensemble size: Bittner (2015), employing a fully-coupled stratosphere resolving model, concluded that after Mt. Pinatubo "the continental winter warming over Northern Europe and Siberia is not significantly different from zero even with 100 ensemble members" (as shown in Fig. 6.4 of that doctoral dissertation).

However, and this is perhaps the key point of our paper: from the fact that the ensemble mean (i.e. the forced) anomalies are not significant, it is *erroneous* to conclude that the models are unable to simulate the NH continental winter warming following the eruption. Recall that the observed anomalies are not expected to resemble the ensemble mean of any set of simulation, as internal variability is superimposed to any forced response in the observations. The correct question to ask is: do any individual simulations resemble the observations? Or, more precisely: do the observed anomalies fall within the range, over the ensemble, of the simulated anomalies? The answer to that question is a resounding yes, as we show next.

Since that answer crucially depends on the range of anomalies that any one model is able to simulate, we start by illustrating that range. In the middle column of Fig. 3 we show the extreme members, i.e. the members with the largest warming anomalies, for each of the three models we have analyzed. Noting that the color-bar is identical to the one in Fig. 1, it is clear that the models are able to simulate much ~~stonger~~ stronger warming anomalies than the observed ones. Even more: different ensemble members of the same models, with *an identical volcanic forcing*, are able to simulate equally strong *cooling* over the northern continents, as shown in the right column of Fig. 3, where the coldest members can be seen. The point of this figure is to illustrate how large the internal variability is (in these models), and how tiny the forced response is in comparison. For completeness, the surface temperature anomalies for each member of each ensemble are shown in supplementary Figs. ~~S1-S3~~S1-S5.

We quantify the relative magnitude of the forced response and the internal variability in Fig. ~~5~~ 4 with box and whisker plots for the quantity T_s', defined as the surface temperature anomaly averaged over the landmasses in the region (40-70N, 0-150W), roughly corresponding to the Eurasian continent. First note that the mean of each ensemble is very near zero (a few tens of degrees at most, and not statistically significant), confirming the results of many previous studies that the forced response in the NH midlatitudes in the winter following the Mt. Pinatubo eruption is basically non-existent in the models. Second, the models are in reasonably good agreement about the internal variability, showing a warming/cooling range of 2 to 4°C on each side of zero, which is much larger than the forced response. Third, and most importantly: the reanalysis (red dot) falls well within the simulated range, indicating that the models are perfectly capable of capturing the post-Pinatubo winter anomalies in the NH.

4 Does the stratospheric pathway play a role in simulating the NH winter warming following the Pinatubo eruption?

Having established that our three models are able to simulate the observed NH continental warming after the Mt. Pinatubo eruption, we now turn to examining the stratospheric pathway mechanism proposed by Robock and Mao (1992), Graf et al. (1993) and others. In a nutshell, that mechanism involves two steps: (1) a strengthening of the stratospheric polar vortex caused by the enhanced equator-to-pole lower stratospheric temperature gradient following the Mt. Pinatubo eruption which, in turn, causes (2) an anomalous atmospheric circulation resulting in a warming anomaly over the Eurasian continent.

To carefully investigate the existence of a possible stratospheric pathway, we will limit ourselves to the WACCM4 model, as the other two do not have an accurate representation of the stratosphere and, more importantly, of its variability. We recognize that 13 members may perhaps not qualify as a “large” ensemble but, as we will show, the results presented below are in excellent agreement with those of Bittner et al. (2016) who used a much larger[4] 100-member ensemble.

Now, to quantify the strength of the polar vortex we compute the quantity U'_{10}, defined as the anomaly in the zonal mean, zonal wind at 10 hPa and 60N. This quantity is widely used for the detection of stratospheric sudden warmings (see, e.g., Charlton and Polvani, 2007; Butler and Gerber, 2018). To quantify the meridional lower stratospheric temperature gradient we compute the quantity $\nabla T'_{50}$, defined as the difference in zonal mean temperature between the tropics (30S-30N) and the polar cap (60-90N) at 50 hPa: that level is chosen so as to capture the maximum amplitude of the stratospheric warming from Mt. Pinatubo at low-latitudes. The relationship between the U'_{10} and $\nabla T'_{50}$ is shown in Fig. 5a: their correlation is exceedingly high (with an r^2 value of 0.89). From the ensemble mean value (black dot) one can see that, indeed, a warming of the tropical lower stratosphere by a potent low-latitude eruption does indeed result in a stronger[5]from wintertime polar vortex in our model.

The key question, however, is: how much stronger? In the case of the Mt. Pinatubo eruption, this is given by the black circles in Fig. 5a, which indicate the ensemble mean value of 3.5 m/s for our WACCM4 simulation. This is in excellent agreement with the findings of Bittner et al. (2016), who also reported 1-2 m/s acceleration of the polar vortex following large low-latitude eruptions, and emphasized that 50-100 of ensemble members are need to establish this result in a statistically significant way. One cannot overemphasize how minuscule this forced response is when contrasted with the unforced, internal variability of the wintertime polar vortex, whose strength can vary by many tens of meters per second over a period as short as a week (e.g. during a stratospheric sudden warming event, which occur roughly every other year, see Charlton and Polvani, 2007).

With this in mind, we now proceed to examining the second step of the proposed mechanism, the relationship between the polar vortex anomaly U'_{10} and the Eurasian surface temperature anomaly T'_s. We find no meaningful correlation between the two, as evident from Fig. 5b (the r^2 value is 0.06), the ensemble mean temperature anomaly is indistinguishable from zero. It is widely appreciated that the variability of the midlatitude tropospheric circulation is very large, so that it can easily overwhelm polar vortex anomalies of tens of meters per second. In fact, even stratospheric sudden warmings – which correspond to massive perturbations of the stratospheric polar vortex ~~which~~ and results in a ~~complete wind reversal~~ wind reversal at 10 hPa, from westerlies to easterlies – are not always able to produce a significant surface signal (see the Sudden Warming Compendium, Butler et al., 2017).

Another way of illustrating the weakness of the connection between polar vortex strength and Eurasian surface temperature anomalies is to contrast two WACCM4 ensemble members – specifically #2 and #12 – for which T'_s is shown in the top row of Fig. 6. We have chosen these two particular members as they simulate very similar Eurasian surface warming anomalies, not unlike the ones in the observations. In spite of those surface similarities, the corresponding stratospheric temperature gradients

---

[4] The WACCM4 simulations analyzed here are a lot more computationally expensive those in (Bittner et al., 2016), as they involve interactive ozone chemistry. In fact, we are aware of no other study with a coupled atmosphere-ocean-chemistry model which has analyzed ensembles with more than a handful of members. Just to cite a few recent studies: McLandress et al. (2011) analyze 3 members, Solomon et al. (2015) 6 members, Li et al. (2018) 4 members. So, we submit that a 13-member ensemble with interactive chemistry, and coupled ocean and sea-ice components, represents a susbtantial step forward.

[5] Although we do not believe that it is appropriate to analyze the CAM5-LE and CanESM models for possible evidence of a stratospheric pathway – as those are low-top models with a poorly resolved stratosphere and thus unrealistic stratospheric variability – we nonetheless include in supplementary Fig. S6 the same scatter plots as in Fig. 5, to satisfy the request of one anonymous referee. The reader can see that those two other models confirm the WACCM results.

are completely different (see the middle row of Fig. 6). The tropical lower stratosphere is anomalously warm in both members, owing to the direct radiative effect of the volcanic aerosols, which is robust. In contrast the polar stratosphere is anomalously warm for one case (#2) but cold for the other (#12). The corresponding temperature gradients $\nabla T'_{50}$ are thus of opposite sign and, predictably, the polar vortex is anomalously weak for the former and strong for the latter member, as seen in the bottom row of Fig. 6, where we show the zonal mean zonal wind at 10 hPa. Note that these opposite-signed polar vortex anomalies have an amplitude of about 10 m/s, which is three times larger than the forced response documented above. In spite of such large and opposite-signed polar vortex anomalies, both members exhibit very similar surface temperature anomalies over Eurasia, as seen in the top row: this ~~clearly~~ demonstrates that polar vortex anomalies do not *necessarily* determine the surface anomalies.

For completeness, the full vertical structure of the ensemble mean temperature anomalies for the WACCM4 model is shown in Fig. 7a. The only statistically significant signal is found in the tropics, where WACCM4 greatly overestimates the post-Pinatubo warming, yielding a temperature gradient in the lower stratosphere that is considerably larger than the observed one: as seen in Fig. 5a, the ensemble-mean simulated value of $\nabla T'_{50}$ is 5.3°C , whereas the observed value is 0.4°C. In spite of a much larger temperature gradient anomaly than the observed one, we find little statistically significant response in the polar stratospheric winds, as seen in Fig. 7b. There is an overall acceleration of the polar vortex, as one might expect, but the area of significance is quite small, and the grid point at 10 hPa and 60N (the canonical metric for the polar vortex strength) is not statistically significant.

This conclusion does not contradict the findings of Bittner et al. (2016), who reported a statistically significant response of the stratospheric polar vortex after the Mt. Pinatubo eruption in their model. We have only 13 members at our disposal here, and this is why we are unable to establish clear significance with WACCM4. To appreciate how difficult it is to obtain a statistically significant response in the polar vortex, we show the U'_{10} anomalies for each of the 13 members in Fig. 8: there is a wide scatter across the 13 members, yielding an ensemble mean which is much smaller than most individual members. Nonetheless, the fact that only 4 members show a vortex weakening and the remaining 9 show a vortex strengthening is suggestive ~~: it is quite likely that had we had~~ of polar vortex acceleration. But, as reported in Bittner et al. (2016), as many as 50 ~~or~~ to 100 members ~~in our ensemble,~~ may be needed to obtain a statistically significant strengthening of the polar vortex~~would have emerged~~.

More important, however, is the red line in Fig. 8, showing the ERA-Interim anomalies: it indicates that the polar vortex was, actually, anomalously weak in the winter following the Mt. Pinatubo eruption. For clarity, we show the entire latitude/pressure profiles of the ERA-Interim temperature and wind anomalies in the bottom row of Fig. 7. Amazingly[6] enough, the polar stratosphere was anomalously warm (not cold) after the eruption (panel c), and the polar vortex was anomalous weak (not strong): note, in panel d, the negative zonal wind between 10 and 1 hPa, and between 50N and 60N, where the climatological polar vortex is located. So we conclude by asking: How can the stratospheric pathway mechanism be invoked as an explanation

---

[6] This crucial fact seems to have gone largely unnoticed in the literature. It is reported in the doctoral dissertation of Thomas (2008, see her Figures 4.16 and 4.17), and tangentially noted by Mitchell et al. (2011, see their Figure 8, and the accompanying text), who employed so-called "elliptical" diagnostics for the polar vortex. It is also briefly discussed in Toohey et al. (2014, see their Figure 1), who argue that wintertime stratospheric state in the first winter after Mt. Pinatubo may be not be representative of the "pure response" to the volcanic aerosols owing to confouding factors (e.g. the Quasi-Biennial Oscillation).

for the observed warming over the NH continents, if the polar stratosphere was actually *warmer* and the polar vortex was actually *weaker* in the winter that followed the 1991 eruption of Mt. Pinatubo?

5 Summary and Discussion

The aim of this paper has been to understand the cause of the warm anomalies that were observed over the NH continents following the Mt. Pinatubo eruption in June 1991. More specifically, referring back to the introduction, we have addresss two related but distinct questions: the ability of the models to simulate the observations and the importance of the stratospheric pathway.

First, we have ~~clearly~~ demonstrated that the current generation of coupled climate models is eminently capable of simulating such anomalies. Unlike previous studies, our conclusion follows from comparing the observed anomalies to *individual* model simulations, not to the *average* of multiple simulations. We have shown that climate models, when forced with an identical volcanic perturbation, can actually simulate a much larger surface warming than observed and, in fact, an equally large cooling. Furthermore, confirming many previous studies, we have shown that averaging across model simulations results in statistically insignificant surface temperature anomalies in the NH following the eruption. Taken together, and assuming climate models are not fundamentally flawed, these facts are here interpreted as follows: the internal variability of the climate system in the NH in wintertime is much larger than any impact from the Mt. Pinatubo eruption. As a consequence, it is hard to imagine that any substantial fraction of the observed warming anomalies in the NH during the 1991-1992 winter were caused by that volcanic eruption.

Second, we have examined in detail the potential role of an often invoked stratospheric pathway mechanism, which would allegedly mediate the signal from a low-latitude eruption to the higher-latitude continents by accelerating the polar vortex, and subsequently causing a positive phase of North Atlantic Oscillation (or the annular mode). Analyzing the WACCM4 model, which is a stratosphere-resolving model with interactive stratospheric ozone chemistry, we find the polar vortex acceleration accompanying the increased lower stratospheric temperature gradient after the Mt. Pinatubo eruption to be no larger than a few meters per second at best. And, we wish to emphasize, the WACCM4 model (like most others) produces an unrealistically large warming of the tropical lower stratosphere (see Figs.2d and 7a,b), which implies an unrealistically strong acceleration of the polar vortex. Even so, that acceleration is actually *not* statistically significant in our 13-member WACCM4 ensemble. This is in total agreement with the recent study of Bittner et al. (2016), who show that 50-100 members are needed to detect a significant acceleration of the polar vortex in the winter following a large-magnitude low-latitude eruption. This, in and of itself, is clear evidence that the forced polar vortex response is very small compared to the internal stratospheric variability in wintertime, where wind perturbations of many tens of meters per second are not unusual. And ultimately, in terms of affecting the tropospheric circulation and surface temperature, such small polar vortex anomalies are completely dwarfed by the internal tropospheric variability; this is why no statistically significant anomalies are found when averaging over many model simulations.

One might now ask how such evidence can be reconciled with several influential early studies, which have argued for the key role of the stratospheric pathway in causing the NH continental surface warming in the winter following the Mt. Pinatubo eruption. We suggest the following: those early models simply lacked a good representation of the stratosphere and, more crucially, of its variability, and this resulted in an overestimate of the forced response to the volcanic eruption. For instance, the model employed in Graf et al. (1993) had a mere 19 vertical levels in the vertical direction, with the model top at only 10 hPa. The same applies to the study of Kirchner et al. (1999), who improved the horizontal resolution but retained the same deficient vertical structure of their model. A severe lack of vertical resolution is also evident in the AMIP models analyzed in Mao and Robock (1998), all of which (with only one exception) have between 10 and 20 vertical levels (see Table 2 of Gates, 1992). Ditto for the study of Collins (2003): 19 vertical levels. As for Shindell et al. (2004), the two models used in that study have only 20 and 23 vertical levels, and the latter has a very coarse horizontal resolution as well (8°latitude × 10°longitude): that model was, in fact, evaluated for its ability to simulate stratospheric sudden warmings, and found to greatly underestimate their frequency (see Fig. 3c of Charlton et al., 2007, under the item GISS23). The reader may want to contrast that model with the WACCM4 model used here, with 66 vertical levels, a model top at ∼140km, and an excellent simulation of the frequency of stratospheric sudden warmings (see Fig. 3a of Marsh et al., 2013).

A note is also in order regarding the recent study Zambri and Robock (2016). They reanalyzed a larger set of CMIP5 models than those in Driscoll et al. (2012), and considered only the anomalies in the first winter after the eruptions. From the multi-model average anomalies following the two largest eruptions since the pre-industrial era they conclude that “the observed surface temperature anomalies are related to changes in the winter circulation *caused* by the volcanic eruptions” (emphasis added), a claim obviously at odds with much of the previous literature, and with the results presented here. However, as their conclusion was drawn by averaging anomalies from the 1883 Krakatau eruption with those from the 1991 Mt. Pinatubo eruption, it is not immediately obvious how to disentangle the forced response to Mt. Pinatubo alone, which is the sole subject of the present study. We plan to carefully examine other volcanic eruptions in an upcoming paper.

Nonetheless, we have briefly analyzed other recent[7] eruptions simulated by the three models described in Section 2.1. Of particular interest is the 1982 eruption of El Chichón (Robock, 1983), which was also followed by anomalous wintertime warming over the Nothern Hemisphere continent (as shown in supplementary Fig. ~~S6~~S7). As for the 1991 Mt. Pinatubo eruption, all three models produce (1) a statistically insignificant forced response and (2) both warm and cold anomalies with identical volcanic forcing (see supplementary Fig. ~~S7~~S8), indicating that the observed continental winter warming following the 1982 El Chichón eruption was also, very likely, a simple manifestation of internal variability. Of course, the validity of our interpretation is dependent on the models’ ability to accurately simulated the internal variability of the climate system.

Still, leaving models – and their possible biases – aside, one could nonetheless argue that several studies have “demonstrated”, on the basis of various temperature reconstructions, that many low-latitude volcanic eruptions have been followed by NH continental warming in wintertime. Whether those demonstrations are truly convincing depends, crucially, on the quality of the surface temperature reconstructions and on the soundness of the methodology employed. Just to give an example: the early

---

[7] After the 1963 eruption of Mt. Agung, the volcanic aerosol cloud spread primarily into the Southern Hemisphere (Viebrock and Flowers, 1968): that eruption is thus not the best candidate for exploring the causal link between low-latitude eruptions and anomalies over the Northern Hemisphere continents.

claim of Robock and Mao (1992) was based on the analysis of a single temperature dataset for – literally – one dozen eruptions, half of which occurred at latitudes outside 30S-30N, averaging together larger and smaller events, including a mixture of first and second winter anomalies (depending on the eruptions). For the reasons stated in Section 2.2 above, we very much agree with the recent suggestion of Zambri and Robock (2016) that (1) only the first winter after each eruption should be considered, (2) eruptions of different magnitudes should not be averaged together : if these two procedural choices are important, many studies in the literature would need to be reconsidered.

In any case, going back to Mt. Pinatubo, the fact remains that from December 1991 to February 1992, the observed surface temperatures were anomalously warm over North America and Eurasia, and that fact may deserve an explanation. Our analysis ~~clearly shows~~ indicates that the continental warming that occurred in the first winter following the 1991 eruption was most likely a simple manifestation of internal atmospheric variability, and was completely unrelated to the eruption itself. So, the next question is: what might be the source of variability that resulted in the NH continental warming? An obvious candidate would be the El Niño-Southern Oscillation (ENSO) phenomenon, since it is well known that the eruption of Mt. Pinatubo corresponded with an El Niño event (see, e.g., Lehner et al., 2016), which is believed to influence the North Atlantic and Eurasia in winter (Brönnimann, 2007; Rodríguez-Fonseca et al., 2016). Unfortunately, El Niño conditions are typically associated with a contraction of the tropical belt (Lu et al., 2008) and a negative phase of the North Atlantic Oscillation (Li and Lau, 2012), which is typically accompanied by cold anomalies over Eurasia. It is, therefore, difficult to argue that the observed post-Pinatubo continental warming was caused by El Niño. In fact, there is some good modeling evidence confirming this. First, Thomas et al. (2009) reported a “very strong” response to El Niño in their model, that “can mask the effects due to volcanic warming”. Second, analyzing so-called pacemaker[8] simulations with the CAM5-LE model, McGraw et al. (2016) show a large forced signal in the tropospheric circulation from El Niño in the Northern Hemisphere, which greatly resembles a negative annular mode (see their Fig. 11f); they don’t show surface temperatures, but one would easily expect cold anomalies over the NH continents in those simulations. If, then, El Niño needs to be ruled out, we may just have to admit that the intrinsic variability of the high latitude tropospheric circulation, which is known to be very large (Shepherd, 2014), might have to suffice as an explanation.

Acknowledgements. LMP and AB are grateful for the support of the US National Science Foundation (USNSF), and wish to express their gratitude to Dr. Ryan Neely for performing and making available some of the WACCM4 model integrations. Computing resources for the WACCM4 and CAM5-LE were provided by the Computational and Information Systems Laboratory (CISL) at the National Center for Atmospheric Research (NCAR). We also acknowledge the National Oceanic and Atmospheric Administration’s (NOAA) Research and Development High Performance Computing Program, for providing computing and storage resources for some of the WACCM4 runs, and wish to thank Henry LeRoy Miller for his assistance with NOAA’s high performance computing facilities. The CESM Large Ensemble Project and supercomputing resources were provided by the USNSF and NCAR/CISL. We also acknowledge Environment and Climate

---

[8] In these simulations a fully coupled atmosphere-ocean is employed, but SST anomalies in the eastern tropical Pacific are nudged to observations, so as to faithfully simulate El Niño events.

Change Canada's Canadian Centre for Climate Modelling and Analysis for executing and making available the CanESM2 Large Ensemble simulations used in this study, and the Canadian Sea Ice and Snow Evolution Network for proposing the simulations.

References

Barnes, J. E. and Hofmann, D. J.: Lidar measurements of stratospheric aerosol over Mauna Loa Observatory, Geophysical Research Letters, 24, 1923–1926, 1997.

Bittner, M.: On the discrepancy between observed and simulated dynamical responses of Northern Hemisphere winter climate to large tropical volcanic eruptions, Ph.D. thesis, Univerisity of Hamburg, Reports on Earth System Science, no. 173, 2015.

Bittner, M., Schmidt, H., Timmreck, C., and Sienz, F.: 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, 2016.

Bluth, G. J., Doiron, S. D., Schnetzler, C. C., Krueger, A. J., and Walter, L. S.: Global tracking of the SO2 clouds from the June, 1991 Mount Pinatubo eruptions, Geophysical Research Letters, 19, 151–154, 1992.

Brönnimann, S.: Impact of El Niño–Southern Oscillation on European climate, Reviews of Geophysics, 45, 1–28, 2007.

Butler, A. H. and Gerber, E. P.: Optimizing the definition of a sudden stratospheric warming, Journal of Climate, 31, 2337–2344, 2018.

Butler, A. H., Sjoberg, J. P., Seidel, D. J., and Rosenlof, K. H.: A sudden stratospheric warming compendium, Earth System Science Data, 9, 63–76, 2017.

Charlton, A. J. and Polvani, L. M.: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks, Journal of Climate, 20, 449–469, 2007.

Charlton, A. J., Polvani, L. M., Perlwitz, J., Sassi, F., Manzini, E., Shibata, K., Pawson, S., Nielsen, J. E., and Rind, D.: A new look at stratospheric sudden warmings. Part II: Evaluation of numerical model simulations, Journal of climate, 20, 470–488, 2007.

Charlton-Perez, A. J., Baldwin, M. P., Birner, T., Black, R. X., Butler, A. H., Calvo, N., Davis, N. A., Gerber, E. P., Gillett, N., Hardiman, S., et al.: On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models, Journal of Geophysical Research: Atmospheres, 118, 2494–2505, 2013.

Collins, M.: Predictions of climate following volcanic eruptions, in: Volcanism and the Earth's Atmosphere, vol. 139 of *Geophysical Monograph Book Series*, pp. 123–135, American Geophysical Union, Washington, DC, USA, 2003.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, 2011.

Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability, Climate dynamics, 38, 527–546, 2012.

Deser, C., Hurrell, J. W., and Phillips, A. S.: The role of the North Atlantic Oscillation in European climate projections, Climate Dynamics, 49, 3141–3157, 2017.

Driscoll, S., Bozzo, A., Gray, L. J., Robock, A., and Stenchikov, G.: Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions, J. Geophys. Res., 117, D17 105, doi:10.1029/2012JD017607, 2012.

Dutton, E. G. and Christy, J. R.: Solar radiative forcing at selected locations and evidence for global lower tropospheric cooling following the eruptions of El Chichón and Pinatubo, Geophysical Research Letters, 19, 2313–2316, 1992.

English, J. M., Toon, O. B., and Mills, M. J.: Microphysical simulations of large volcanic eruptions: Pinatubo and Toba, Journal of Geophysical Research: Atmospheres, 118, 1880–1895, 2013.

Fischer, E. M., Luterbacher, J., Zorita, E., Tett, S., Casty, C., and Wanner, H.: European climate response to tropical volcanic eruptions over the last half millennium, Geophysical Research Letters, 34, 2007.

Fujiwara, M., Hibino, T., Mehta, S., Gray, L., Mitchell, D., and Anstey, J.: Global temperature response to the major volcanic eruptions in multiple reanalysis data sets, Atmospheric chemistry and physics, 15, 13 507–13 518, 2015.

Gagné, M.-È., Kirchmeier-Young, M., Gillett, N., and Fyfe, J.: Arctic sea ice response to the eruptions of Agung, El Chichón and Pinatubo, Journal of Geophysical Research: Atmospheres, 2017.

Gates, W. L.: AMIP: The atmospheric model intercomparison project, Bulletin of the American Meteorological Society, 73, 1962–1970, 1992.

Gettelman, A., Kay, J., and Fasullo, J.: Spatial decomposition of climate feedbacks in the Community Earth System Model, Journal of Climate, 26, 3544–3561, 2013.

Graf, H., Kirchner, I., Robock, A., and Schult, I.: Pinatubo eruption winter climate effects: Model versus observations, Climate Dynamics, 9, 81–93, 1993.

Graf, H.-F., Li, Q., and Giorgetta, M.: Volcanic effects on climate: revisiting the mechanisms, Atmospheric Chemistry and Physics, 7, 4503–4511, 2007.

Groisman, P. Y.: Possible regional climate consequences of the Pinatubo eruption: An empirical approach, Geophysical Research Letters, 19, 1603–1606, 1992.

Hansen, J., Ruedy, R., Sato, M., and Lo, K.: global surface temperature change, Reviews of Geophysics, 48, n/a–n/a, 2010.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bulletin of the American Meteorological Society, 77, 437–472, 1996.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The Community Earth System Model (CESM) Large Ensemble Project, Bulletin of the American Meteorological Society, 96, 1333–1349, 2015.

Kelly, P. M., Jia, P., and Jones, P.: The spatial response of the climate system to explosive volcanic eruptions, International Journal of Climatology, 16, 537–550, 1996.

Kirchner, I., Stenchikov, G. L., Graf, H.-F., Robock, A., and Antuña, J. C.: Climate model simulation of winter warming and summer cooling following the 1991 Mount Pinatubo volcanic eruption, Journal of Geophysical Research: Atmospheres, 104, 19 039–19 055, 1999.

Kodera, K.: Influence of volcanic eruptions on the troposphere through stratospheric dynamical processes in the northern hemisphere winter, Journal of Geophysical Research: Atmospheres, 99, 1273–1282, 1994.

Lehner, F., Schurer, A. P., Hegerl, G. C., Deser, C., and Frölicher, T. L.: The importance of ENSO phase during volcanic eruptions for detection and attribution, Geophysical Research Letters, 43, 2851–2858, 2016.

Li, F., Newman, P., Pawson, S., and Perlwitz, J.: Effects of Greenhouse Gas Increase and Stratospheric Ozone Depletion on Stratospheric Mean Age of Air in 1960-2010, Journal of Geophysical Research: Atmospheres, 2018.

Li, Y. and Lau, N.-C.: Impact of ENSO on the atmospheric variability over the North Atlantic in late winter—Role of transient eddies, Journal of Climate, 25, 320–342, 2012.

Long, C. S. and Stowe, L. L.: Using the NOAA/AVHRR to study stratospheric aerosol optical thicknesses following the Mt. Pinatubo eruption, Geophysical research letters, 21, 2215–2218, 1994.

Lu, J., Chen, G., and Frierson, D. M.: Response of the zonal mean atmospheric circulation to El Niño versus global warming, Journal of Climate, 21, 5835–5851, 2008.

Mao, J. and Robock, A.: Surface air temperature simulations by AMIP general circulation models: Volcanic and ENSO signals and systematic errors, Journal of climate, 11, 1538–1552, 1998.

Marsh, D. R., Mills, M. J., Kinnison, D. E., Lamarque, J.-F., Calvo, N., and Polvani, L. M.: Climate change from 1850 to 2005 simulated in CESM1 (WACCM), Journal of Climate, 26, 7372–7391, 2013.

Marshall, A., Scaife, A., and Ineson, S.: Enhanced seasonal prediction of European winter warming following volcanic eruptions, Journal of climate, 22, 6168–6180, 2009.

McCormick, M. and Veiga, R.: SAGE II measurements of early Pinatubo aerosols, Geophysical Research Letters, 19, 155–158, 1992.

McCormick, M. P., Thomason, L. W., and Trepte, C. R.: Atmospheric effects of the Mt Pinatubo eruption, Nature, 373, 399, 1995.

McGraw, M. C., Barnes, E. A., and Deser, C.: Reconciling the observed and modeled Southern Hemisphere circulation response to volcanic eruptions, Geophysical Research Letters, 43, 7259–7266, 2016.

McLandress, C., Shepherd, T. G., Scinocca, J. F., Plummer, D. A., Sigmond, M., Jonsson, A. I., and Reader, M. C.: Separating the dynamical effects of climate change and ozone depletion. Part II: Southern Hemisphere troposphere, Journal of Climate, 24, 1850–1868, 2011.

Mitchell, D., Gray, L., and Charlton-Perez, A.: The structure and evolution of the stratospheric vortex in response to natural forcings, Journal of Geophysical Research: Atmospheres, 116, 2011.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, d08101, 2012.

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., Conley, A. J., Garcia, R., Kinnison, D., Lamarque, J.-F., et al.: Description of the NCAR community atmosphere model (CAM 5.0), NCAR Tech. Note NCAR/TN-486+ STR, 1, 1–12, 2010.

Perlwitz, J. and Graf, H.-F.: The statistical connection between tropospheric and stratospheric circulation of the Northern Hemisphere in winter, Journal of Climate, 8, 2281–2295, 1995.

Randel, W. J.: Variability and Trends in Stratospheric Temperature and Water Vapor, in: The Stratosphere: dynamics, transport, and chemistry, edited by Polvani, LM and Sobel, AH and Waugh, DW, vol. 190 of *Geophysical Monograph Book Series*, pp. 123–135, American Geophysical Union, Washington, DC, USA, 2010.

Robock, A.: The dust cloud of the century, Nature, 301, 373, 1983.

Robock, A.: Volcanic eruptions and climate, Reviews of Geophysics, 38, 191–219, 2000.

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

Robock, A. and Mao, J.: Winter warming from large volcanic eruptions, Geophysical Research Letters, 19, 2405–2408, 1992.

Robock, A. and Mao, J.: The volcanic signal in surface temperature observations, Journal of Climate, 8, 1086–1103, 1995.

Rodríguez-Fonseca, B., Suárez-Moreno, R., Ayarzagüena, B., López-Parages, J., Gómara, I., Villamayor, J., Mohino, E., Losada, T., and Castaño-Tierno, A.: A review of ENSO influence on the North Atlantic. A non-stationary signal, Atmosphere, 7, 87, 2016.

Sato, M., Hansen, J. E., McCormick, M. P., and Pollack, J. B.: Stratospheric aerosol optical depths, 1850–1990, Journal of Geophysical Research: Atmospheres, 98, 22 987–22 994, 1993.

Shepherd, T. G.: Atmospheric circulation as a source of uncertainty in climate change projections, Nature Geoscience, 7, 703, 2014.

Shindell, D. T., Schmidt, G. A., Mann, M. E., and Faluvegi, G.: Dynamic winter climate response to large tropical volcanic eruptions since 1600, Journal of Geophysical Research: Atmospheres, 109, 2004.

Solomon, A., Polvani, L. M., Smith, K., and Abernathey, R.: The impact of ozone depleting substances on the circulation, temperature, and salinity of the Southern Ocean: An attribution study with CESM1 (WACCM), Geophysical Research Letters, 42, 5547–5555, 2015.

Stenchikov, G., Robock, A., Ramaswamy, V., Schwarzkopf, M. D., Hamilton, K., and Ramachandran, S.: Arctic Oscillation response to the 1991 Mount Pinatubo eruption: Effects of volcanic aerosols and ozone depletion, Journal of Geophysical Research: Atmospheres, 107, 2002.

Stenchikov, G., Hamilton, K., Stouffer, R. J., Robock, A., Ramaswamy, V., Santer, B., and Graf, H.-F.: Arctic Oscillation response to volcanic eruptions in the IPCC AR4 climate models, Journal of Geophysical Research: Atmospheres, 111, 2006.

Swart, N. C., Fyfe, J. C., Hawkins, E., Kay, J. E., and Jahn, A.: Influence of internal variability on Arctic sea-ice trends, Nature Climate Change, 5, 86, 2015.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, 2012.

Thomas, M., Timmreck, C., Giorgetta, M., Graf, H.-F., and Stenchikov, G.: Simulation of the climate impact of Mt. Pinatubo eruption using ECHAM5–Part 1: Sensitivity to the modes of atmospheric circulation and boundary conditions, Atmospheric Chemistry and Physics, 9, 757–769, 2009.

Thomas, M. A.: Simulation of the climate impact of Mt. Pinatubo eruption using ECHAM5, Ph.D. thesis, Univerisity of Hamburg, Reports on Earth System Science, no. 52, 2008.

Toohey, M., Krüger, K., Bittner, M., Timmreck, C., and Schmidt, H.: The impact of volcanic aerosol on the Northern Hemisphere stratospheric polar vortex: mechanisms and sensitivity to forcing structure, Atmospheric Chemistry and Physics, 14, 13 063–13 079, 2014.

Viebrock, H. J. and Flowers, E. C.: Comments on the recent decrease in solar radiation at the South Pole, Tellus, 20, 400–411, 1968.

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J., Cole, J. N., Plummer, D., Verseghy, D., Reader, M. C., Ma, X., Lazare, M., et al.: The Canadian fourth generation atmospheric global climate model (CanAM4). Part I: representation of physical processes, Atmosphere-Ocean, 51, 104–125, 2013.

Wunderlich, F. and Mitchell, D. M.: Revisiting the observed surface climate response to large volcanic eruptions, Atmospheric Chemistry and Physics, 17, 485–499, 2017.

Zambri, B. and Robock, A.: Winter warming and summer monsoon reduction after volcanic eruptions in Coupled Model Intercomparison Project 5 (CMIP5) simulations, Geophysical Research Letters, 43, 2016.

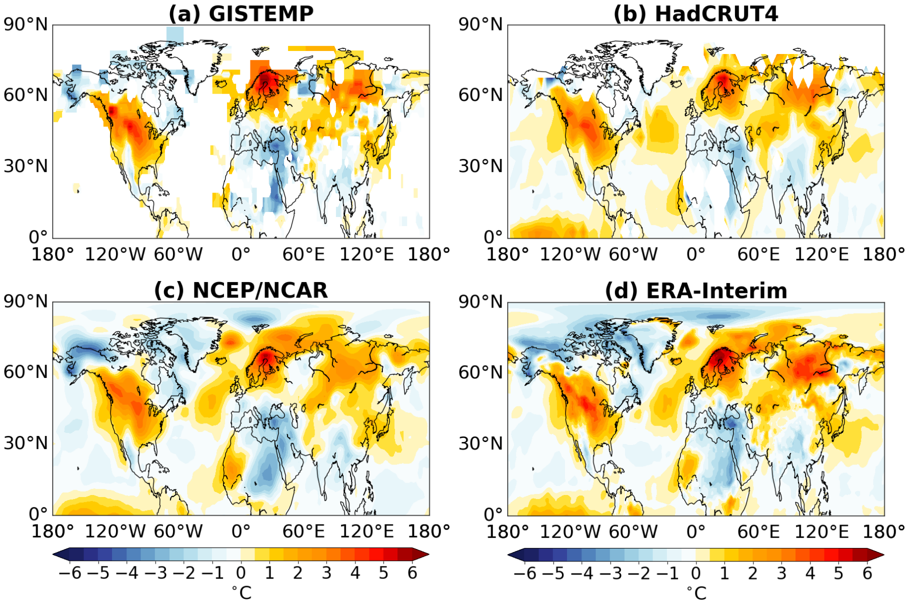


Figure 1. Surface air temperature anomalies (in °C) for the post-Pinatubo winter of 1991-92 relative to the reference period (1985-1990) in observations (a) GISTEMP and (b) HadCRUT4, and in reanalyses (c) NCEP/NCAR and (d) ERA-Interim.

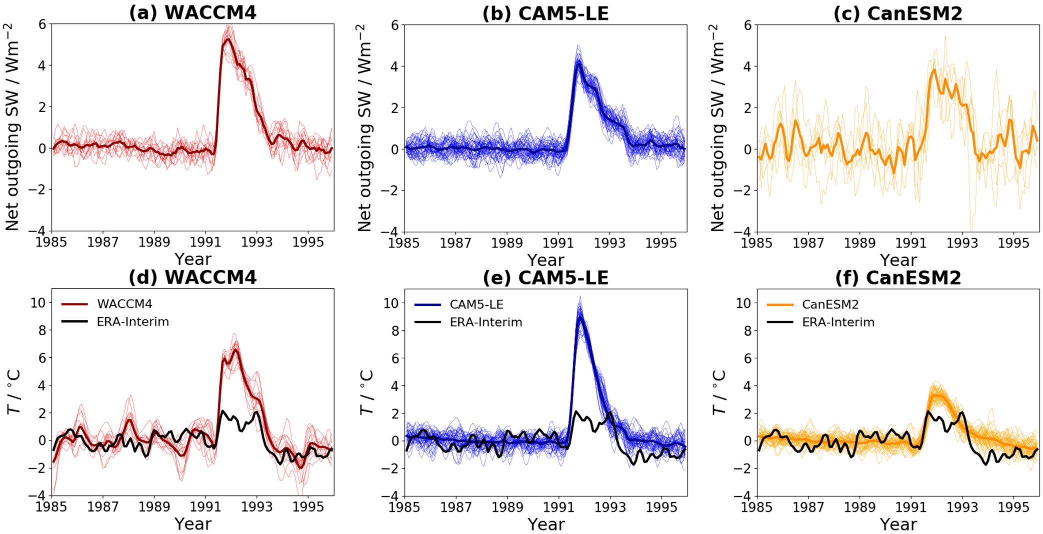
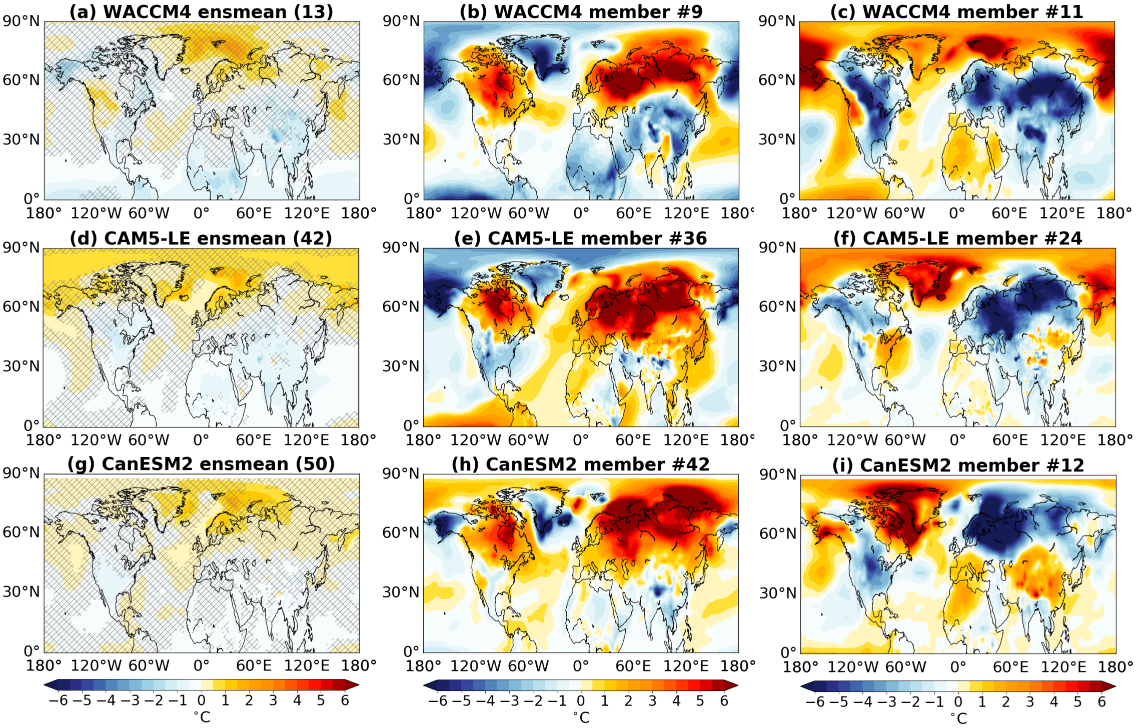


Figure 2. Top row: globally averaged, de-seasonalized, net, outgoing SW radiation at the top of the atmosphere (in W/m^2). Bottom row: tropically averaged (30S-30N), deseasonalized temperature (in °C) at 50 hPa. Left column: WACCM4 (red lines). Middle column: CAM5-LE-LE (blue lines). Right column: CanESM2 (yellow lines). In each panel, the time series for each ensemble member (thin lines) and for the ensemble mean (bold line) are shown. In the bottom row, ERA-Interim values are also shown for comparison (black). All time series are anomalies from the 1985-1990 mean, and are smoothed with a 3-month running average, for direct comparison with Figs. 2 and 3 of Driscoll et al. (2012).


(a) WACCM4 ensmean (13)

(b) WACCM4 member #9

(c) WACCM4 member #11

(d) CAM5-LE ensmean (42)

(e) CAM5-LE member #36

(f) CAM5-LE member #24

(g) CanESM2 ensmean (50)

(h) CanESM2 member #42

(i) CanESM2 member #12

Figure 3. Wintertime surface air temperature anomalies (in °C) as simulated by WACCM4 (top row), CAM5-LE (middle row) and CanESM2 (bottom row) following the 1991 Mt. Pinatubo eruption. Left column: the ensemble mean for each model (with the number of ensemble members in parentheses), and hatching over areas where the anomalies not significant at the 95% confidence level. Middle column: individual members exhibiting extreme warming over the NH continents for each model. Right column: individual members exhibiting extreme cooling.

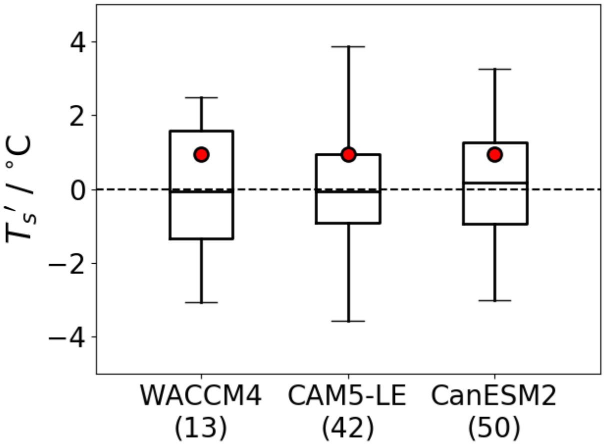


Figure 4. Box and whisker plots of simulated surface temperature anomaly (in °C) over Eurasia (40-70N, 0-150W) in the first post-Pinatubo winter (1991-92) relative to the reference period (1985-1990). The horizontal line inside each box denotes the ensemble mean; the lower and upper limits of each ~~nbox~~ box denote the 25th and 75th percentile values, respectively; the whiskers span the full range of the ensemble members. For comparison, the red circles denote the value calculated from the ERA-Interim reanalyses.

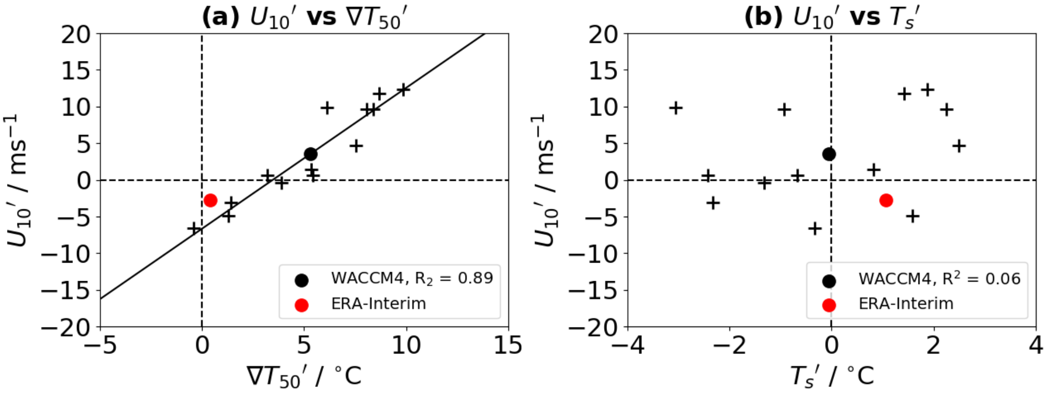


Figure 5. Scatter plots showing the relationship between U_{10}', the anomalies in the zonal mean zonal wind at 10hPa and 60N (in m/s) and the anomalies in (a) the NH meridional temperature gradient $\nabla T_{50}'$ between the tropics (30S-30N) and the pole (60-90N) (in °C), and (b) the Eurasian surface air temperature T_s' (also in °C). Crosses show individual ensemble members, and the black dot shows the ensemble mean value. The red dot shows the ERA-Interim reanalysis.

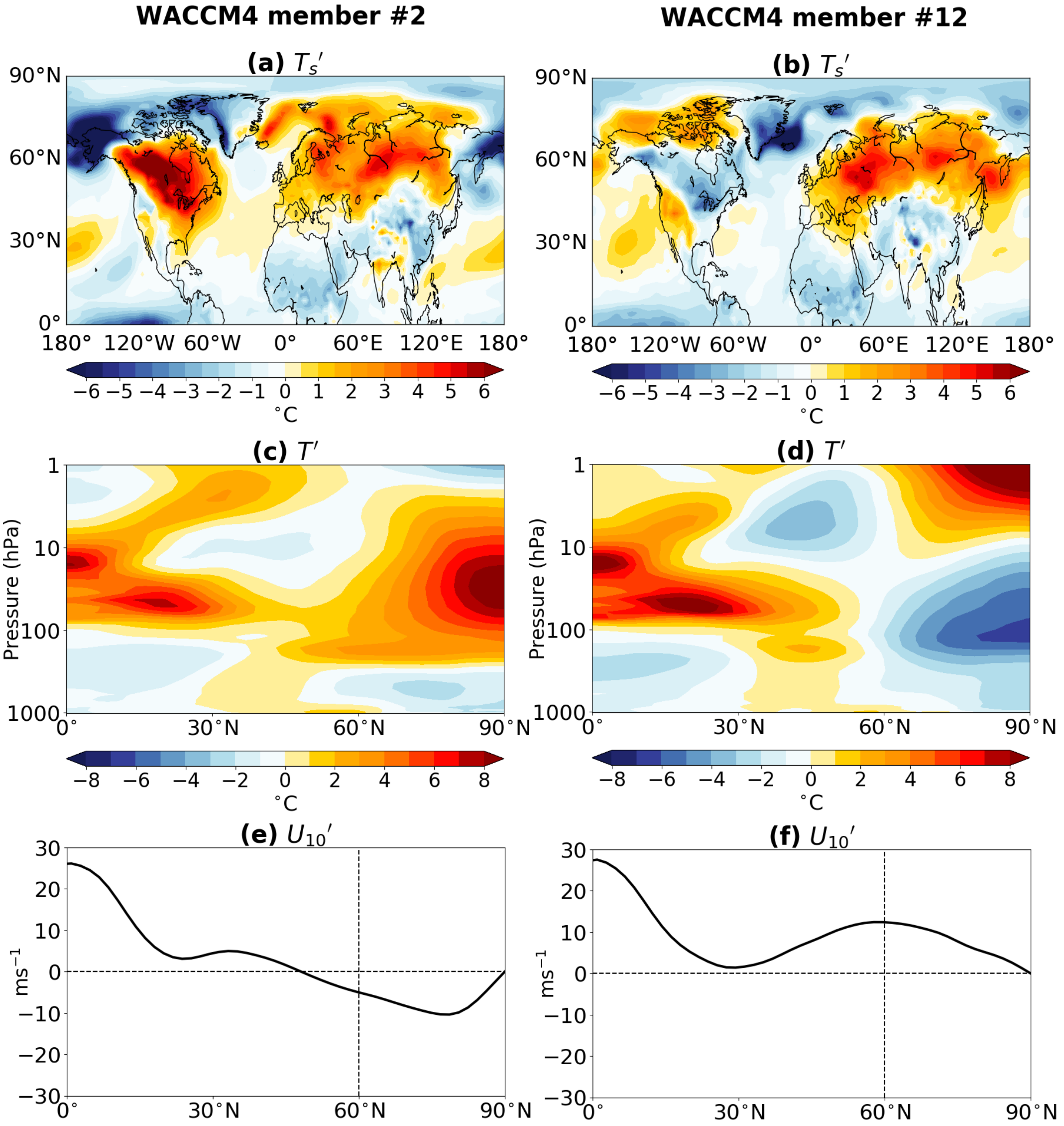


Figure 6. The surface temperature T_s' (top), the zonal mean temperature T' (middle) and 10h hPa zonal mean zonal wind U_{10}' anomalies for WACCM4 member #2 (left) and member #12 (right)

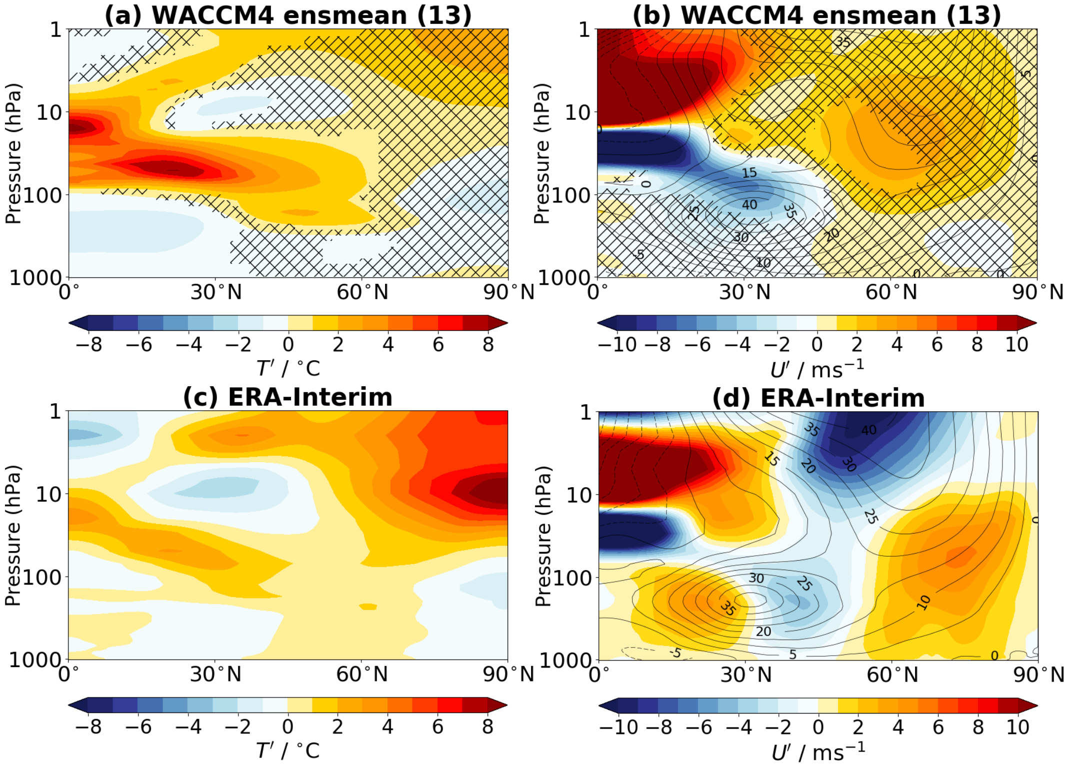


Figure 7. Latitude/pressure anomalies for the winter following the 1991 Mt. Pinatubo eruption. Left: zonal mean temperature (T'). Right: zonal mean zonal wind U', with the climatology in black contours. Top: the ensemble mean of the WACCM simulations, with hatching for values that are not significant at the 95% confidence level. Bottom: corresponding anomalies in the ERA-Interim reanalysis.

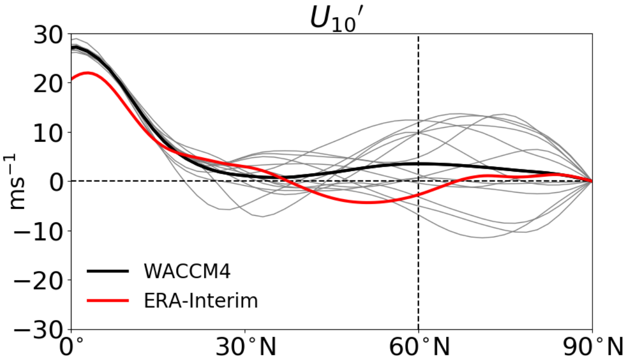


Figure 8. Zonal mean zonal wind anomalies at 10 hPa (U'_{10}) vs. latitude, for the individual WACCM4 simulations (gray), for the ensemble mean (black), and for the ERA-Interim reanalysis (red).