We thank the reviewer for their thorough and thoughtful review. We have addressed your comments below and have added our own comments in RED:

The authors present the development of a stratospheric prognostic aerosol (SPA) capability for the Energy Exascale Model, version 2 (E3SMv2) to simulate the stratospheric aerosol formation in the aftermath of large explosive volcanic eruptions. Their implementation includes changes to the 4-mode Modal Aerosol Module microphysics to allow for larger particle growth and more accurate stratospheric aerosol lifetime following the Mt. Pinatubo eruption. Hunter et al. tested their model for the Post Pinatubo period with remote sensing and in situ observations and the interactive chemistry-climate model, CESM2-WACCM. On the global scale, E3SMv2-SPA performs well compared to observational datasets and has similar behavior to CESM2-WACCM. They found that the modeled aerosol effective radius for both versions is consistently lower than satellite and in-situ measurements (max differences of ~30%). Compared to observations, the models also show a higher diffuse radiation at the surface and a larger cooling and an underestimation in stratospheric heating in the models.

Although the manuscript type is declared as a development and technical paper, the content should be placed in the general context of global stratospheric aerosol modelling, ***otherwise it should be published as a specific technical institute report. The introduction and discussion sections therefore need some substantial improvements. The motivation of the paper could be more clearly stated, and some of the results could be discussed in a broader context. I therefore recommend publication after major revisions, see below.

General comments

In the introduction important literature is missing. Several global stratospheric aerosol modelling studies have been published in the last year. An overview of the development and current state of stratospheric aerosol modelling can be found for example in Kremser et al. (2016) and in Timmreck et al. (2018). In addition, a number of global comparative aerosol modelling studies have been published in recent years, e.g. for background aerosol (Brodowsky et al., 2024), volcanic events (Marshall et al., 2018; Clyne et al., 2021; Quaglia et al., 2023) and artificial sulphur injections (Franke et al., 2021; Weisenstein et al., 2022). I was more than surprised that these studies were completely ignored by the authors. In particular, the results of the Pinatubo study by Quaglia et al. (2023) should be mentioned and discussed in the paper and not just used as a reference to observational data.

Thank you for pointing this out. We have addressed his comment by adding a paragraph to the introduction specific to stratospheric aerosol, touching on background aerosol, the role of volcanic injections, and stratospheric aerosol injection. We also added three paragraphs that address the significance of microphysical representation, injection height, and sulfate chemistry for modeling volcanic eruptions from the context of the above and related references.

I wonder how model specific your results are? How valuable are they to other stratospheric aerosol modellers? I am missing in the discussion section a dedicated paragraph on the strengths and weaknesses of the applied global aerosol models with respect to other global stratospheric aerosol models. Recent intercomparison studies of global aerosol models reveal several difficulties that the current generation of global aerosol models has to deal with. For example, the study by Qualia et al. (2023) comparing the different model results with satellite observations after the eruption of Mt. Pinatubo shows a stronger transport towards the NH extra-tropics, suggesting a much weaker subtropical barrier in all models. How does the spatial aerosol distribution in your model look like? It should be much better as you nudge the winds, so discrepancies can be traced back to other sources. This could be more elaborated with respect to free running models. Nevertheless, it would be nice to see a global distribution of your sulfate burden/AOD also in the paper or in the supplements.

We agree that the larger context of these results was lacking. To address this we have added a zonally averaged AOD plot to the paper to help clarify global transport in these models and allow for comparison to models in Quaglia et al. (2023). We have also added additional discussion of where the AOD and R_{eff} compare to other modeling work. We also include discussion on results from fully coupled historical simulations (1850-2014) run with E3SMv2-SPA to support the use of this model for historical eruptions.

The motivation of the paper could be stress out a bit more. It is also not really clear to me how different your SPA version is from the MAM4 version in WACCAM, except the model reversal and simplified precursor chemistry.

We have added more to the introduction and discussion regarding the purpose of this model. The key difference is the interactive chemistry versus simplified precursor chemistry. There are also slightly different assumptions for sulfate in the stratosphere (i.e., inclusion of a reversible coarse to accumulation mode transfer in the stratosphere). We present arguments for these choices in our paper and in some of your comments below.

The applied methodology is not sufficiently explained in the manuscript. I miss for example a detailed description how you calculate a spatially averaged aerosol size distribution or effective radii which is not straightforward. A subsection "Methodology" for section 2 would be helpful with more details in the appendix.

We have added more details regarding calculation of stratospheric Reff and size distributions to the end of Appendix A, now titled, "Effective radius and aerosol size distribution calculations"

Specific comments

- Line 2: "…using *observations* after the MT. Pinatubo eruption"
  - Fixed title to reflect this change
- Line 45: "Mt. Pinatubo" sometimes you use "Mt. Pinatubo" sometimes "Pinatubo" only, please be consistent
  - We have removed occurrences of 'Mt.' as well as 'Cerro'
- Lines 49-51 The fact does a model produce similar results like another model does not make it per se to a viable tool

- We acknowledge that saying this minimizes some of the uncertainties between the two models. We also recognize that this neglects the most important comparisons, which are to the observational data. Still, we think the comparison with CESM2-WACCM is important given that it is a well-validated and widely used model for these comparisons. We have changed this sentence to read: "The overall agreement of E3SMv2-SPA compared to observations and its similar performance to the well-validated CESM2-WACCM makes E3SMv2-SPA a viable alternative to simulating climate impacts from stratospheric sulfate aerosols."

- Lines 95 ff: Concerning the advantages of sectional aerosol models there is a recent paper by Tilmes et al. (2023) in GMD where they are describing a sectional aerosol microphysical model in CESM2 and compare it with the CESM2 standard version with the Modal Aerosol Model MAM4 for the Pinatubo episode. This paper should be cited and briefly discussed here as well.
  - We have included a reference to this paper. The sentence now reads: "…Pinatubo and larger magnitude eruptions. **More recently, Tilmes et al. (2023) showed that coupling CARMA to WACCM6 better represents the largest aerosol sizes following Pinatubo than a parallel running modal aerosol model.** The modal aerosol approach – representing aerosol size distributions by multiple, evolving lognormal functions – strikes a balance between bulk simplicity and sectional cost. …"
- Line 148: What about sedimentation?
  - We included a mention of sedimentation.
- Line s172-174: Any reason why you did not take this process into account in your model.
  - Including this process would certainly be a more complete representation, but our goal with this work was to have a simpler model that performed well in representing stratospheric sulfate formation. We left this reaction out under the assumption that the effects on coarse mode aerosol in the stratosphere would be minimal given the temperature and relative humidity dependent nature of the reversible process. We have added a statement to this effect: "In CESM2-WACCM, this transfer is reversible in the stratosphere, with an aqueous sulfuric acid (H_2SO_4) equilibrium pressure that depends on temperature and relative humidity. **We left this out of our implementation under the assumption that, at the low relative humidities and low temperatures characteristic of the stratosphere, the effect from this process would be minimal**."
- Lines 213 -214: This is not what Kremser et al (2016) wrote *"Recent modeling studies support lower stratospheric sulfur levels than those inferred from the TOMS and TOVS observations [Dhomse et al.,2014; Mills et al., 2016]. The difference between the initial and the persistent sulfur levels is important and generally supports a more complex development process following a major eruption than has been considered in the past. (Kremser et al., 2016, page 12),* Please cite them properly
  - I am also confused as to why the Kremser reference was used here, and thank you for pointing this out. We have changed the lines to the following: "In the case of Pinatubo, while 18-19 Tg of SO2 erupted in the atmosphere, only ~10 Tg remained in the stratosphere 7-9 days after the eruption (Guo et al., 2004b). This rapid reduction in SO2 corresponds to >99% removal of volcanic ash mass (Guo

et al., 2004a). Therefore, 10 Tg of SO2 is emitted in this dataset for further chemical and microphysical evolution (Mills et al., 2016)."

- Line 451 ff: Solar flux changes: you can also compare your model results to observations by the Earth Radiation Budget Experiment (ERBE) (Barkstrom, 1984; Barkstrom and Smith, 1986). This would an additional approach

We have added an additional TOA flux figure and new section (4.3 TOA radiative flux) to address your concern. We also include clear-sky and aerosol-only model diagnostics to identify biases arising from our subtraction of year 1990 conditions:

3.3 Top-of-atmosphere radiative flux

3.3.1 ERBS

The TOA global radiative flux at a 1°x1°resolution is used from version 2 of the Diagnosing Earth's Energy Pathways in the Climate project (DEEP-C) merged data product drawing from the Earth Radiation Budget Satellite (ERBS) near-global (60°S-60°N) non-scanning instrument and other reanalysis and observational datasets (Allan et al., 2014). The ERBS instrument measures reflected shortwave radiation and total outgoing radiation, allowing for the separation of longwave radiative flux by subtraction (Minnis et al., 1993).

4.3 TOA radiation flux

Figure 4 compares the TOA radiative flux from model simulations to the all-sky ERBS observations over the 1991-1993 period, subtracting out corresponding monthly means from the pre-Pinatubo year 1990. Model TOA flux is shown for all-sky (solid lines), clear-sky (faint dashed lines), and aerosol impact only (faint dotted line) conditions. The radiative flux is reported as absorbed shortwave radiation (ASR, positive downward flux; Fig. 4a), outgoing longwave radiation (OLR, positive upward flux; Fig. 4b), and net radiative flux (NET, positive downward flux; Fig. 4c). In Fig. 4a, ASR shows the clearest model separation 3-4 months after Pinatubo corresponding with peak AOD (Fig. 2). There is close agreement between E3SMv2-SPA, E3SMv2-presc, and CESM2-WACCM during the year 1992 which corresponds to the largest sulfate particles during the Pinatubo plume evolution (see Sections 4.5 and 4.6). The all-sky signal exhibits noise due to differences in atmospheric conditions (i.e., cloud cover, tropospheric aerosol) and surface albedo between the period of interest and our control year (1990). There is a clear seasonal increase in ASR in 1991/1992 and 1992/1993 Northern Hemisphere winters relative to Northern Hemisphere summer. When clear-sky (no influence from clouds) is compared to all-sky conditions in the models the seasonality disappears, implying that the seasonality is cloud-related and cloud albedo was greater in Northern Hemisphere winter 1990 than Northern Hemisphere winter 1991/1992 and 1992/1993. Even with noise introduced by non-Pinatubo factors, there is a distinct all-sky ASR signal in E3SMv2-SPA, CESM2-WACCM, and E3SMv2-presc that is improved compared to ERBS.

The all-sky OLR (Fig. 4b), which is affected both by aerosol absorption of infrared emissions from the earth's surface and the cooling of the troposphere and surface by the scattering of solar radiation, has a weaker response across these models than ASR. This is due in part to a less efficient absorption of outgoing longwave radiation than scattering of incoming solar radiation, leading to a lower sensitivity of OLR to aerosol growth and evolution (see Section 4.8). The largest spread in model simulations occurs during 1992 when aerosols are at their largest (i.e., highest absorption efficiency of longwave radiation; Section 4.8) and the highest reduction in surface temperatures were observed (Parker et al., 1996). All-sky E3SMv2-SPA has the greatest reduction in OLR from April 1992 to the end of 1993, and overestimates the longwave flux reduction compared to ERBS. This corresponds with E3SMv2-SPA overestimation of global AOD values compared to AVHRR over this period (Fig. 2). During this same period, CESM2-WACCM has slightly better agreement with ERBS, which may be related to the temperature nudging in this simulation which will modulate CESM2-WACCM surface temperature reduction and stratospheric temperature. When clear-sky OLR fluxes are compared, there is a weaker reduction in OLR for E3SMv2-PA, E3SMv2-SPA, and CESM2-WACCM, and nearly no change in E3SMv2-presc during 1992. Due to the lack of stratospheric aerosol in E3SMv2-presc, this appears to be evidence of volcanic influence on high altitude clouds which act to reduce OLR further supporting conclusions from Liu and Penner (2002) and Wylie et al. (1994). Lastly, the aerosol-only model simulations remove the 1991/1992 and 1992/1993 wintertime peaks in the OLR signal, indicating similar or smaller OLR in 1990 than our period of interest due to cooler surface conditions.

The improvements in all-sky NET (Fig. 4c; solid lines) with volcanic parameterizations are less apparent across the models than in ASR (Fig. 4a), but do show improvement during the first 6 months after the eruption and during 1992. Differences in cloud cover and surface conditions between our period of interest and 1990 introduce substantial noise to this comparison, but the removal of clouds (clear-sky) and the isolation of aerosol TOA forcing (aerosol only) show a clear separation of volcanic parameterizing models and E3SMv2-PA.

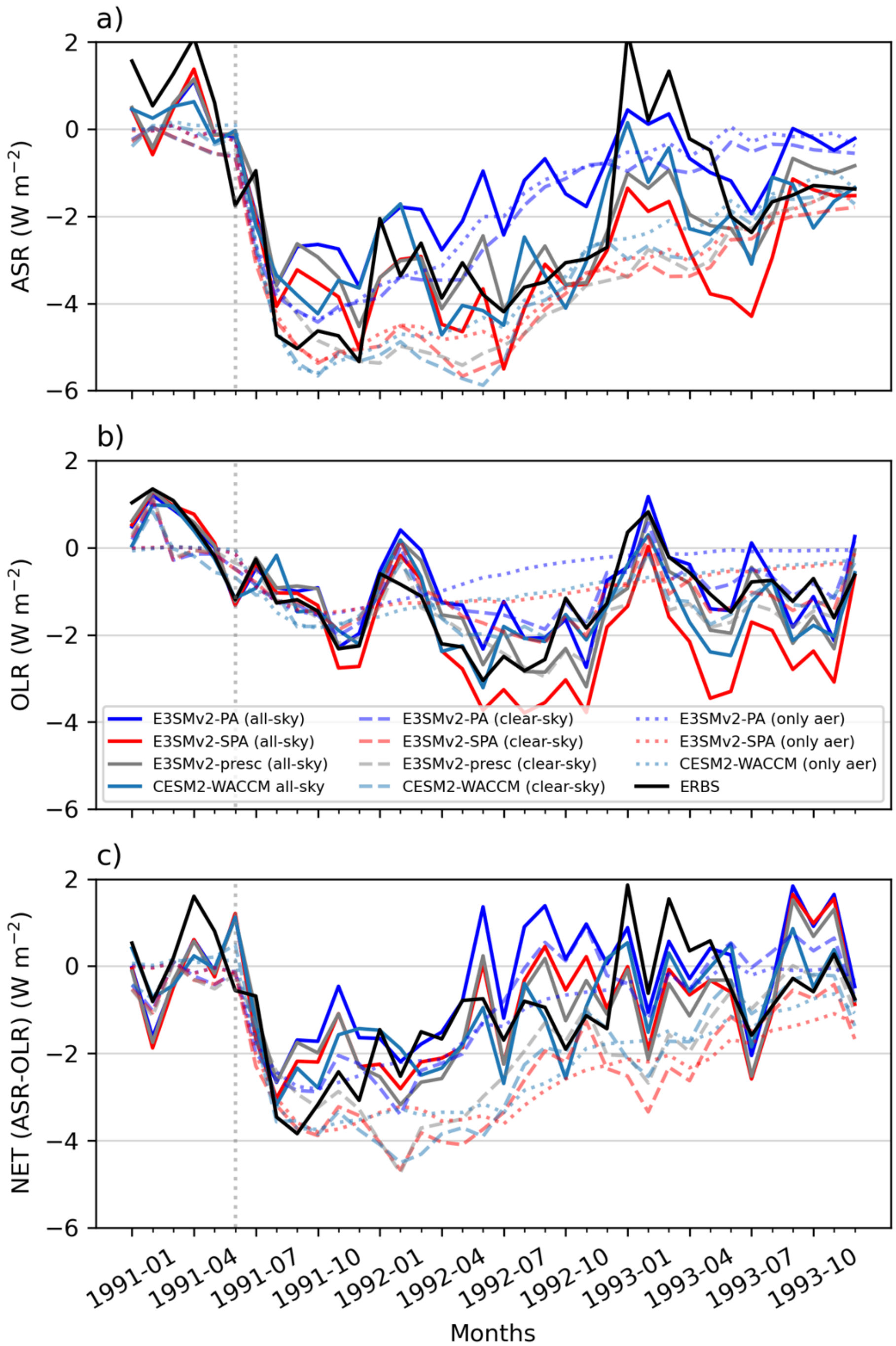

Figure 4: Top-of-atmosphere, radiative flux from model simulations and ERBS observations (Allan et al., 2014; Liu et al., 2015). The panels describe: (a) absorbed solar radiation (ASR; positive downward flux); (b) outgoing longwave radiation (OLR; positive upward flux); and (c) net radiative flux (NET=ASR-OLR; positive downward flux). Monthly mean data is normalized to the pre-Pinatubo conditions by subtracting respective monthly means from the year 1990. ERBS TOA flux is under all-sky conditions, while model TOA flux is shown under all-sky (solid line), clear-sky conditions (faint dashed line), and aerosol only (faint dotted line) conditions.

- Line 494: Date of the Lascar eruption is not correct
  - Fixed.

- Line 571: I am wondering why you choose the following latitudes band and not (also) the location of Mauna Loa where you have some data for a direct comparison.
    - We didn't include a direct comparison between this dataset and the models because it was not clear how to create an equivalent comparison between the coarser temporally resolved model data and the Mauna Loa dataset, the latter of which was reported for clear-sky morning conditions at a fixed solar zenith angle.
- Lines 639-647: Does an integrated longwave heating rate really make sense here. Would not it be more useful to compare stratospheric temperature profiles here where are at least some observations are available, e.g. Free and Angell (2002) and Free and Lazante (2009).
    - We chose longwave heating rate due to its clearer signal (at least given our aerosol specific diagnostic methods) and its physical relationship to AOD and absorption efficiency. We agree that temperature profiles would give a better idea of model performance in the stratosphere. We have included a temperature comparison (Section 4.4) earlier in the paper following the TOA flux plot.

3.4 Atmospheric temperature profiles

3.4.1 MERRA-2

The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) is a reanalysis product that assimilates satellite, radiosonde, radar, ship, buoy, and aircraft observations into version 5.12.4 of the Goddard Earth Observing System (GEOS) atmospheric general circulation model (Rienecker et al., 2011; Gelaro et al., 2017). This data is produced on a 0.5°x0.625°grid with 72 vertical levels from the surface to 0.01 hPa. MERRA-2 observations include atmospheric state (temperature, pressure, humidity), dynamics, precipitation, radiation, and ozone, with updated aerosol observations from AVHRR over the period 1979-2002 (Gelaro et al., 2017).

3.4.2 RICH-obs

Version 1.5.1 of the Radiosonde Inovation Composite Homogenization (RICH-obs) software package is a compiled global radiosonde dataset that is merged with the help of reanalysis climatologies and neighboring data temperature records dating back to 1958 (Haimberger et al., 2012, 2008). The data gaps in station data are identified by divergence from 40-year climatology in the European Centere for Medium-Range Weather Forecasts Reanalysis (ERA-40) and the interpolation of these gaps are estimated from time series of neighboring radiosonde measurements, making RICH-obs less affected by satellite observations in the reanalysis but potentially biased in remote regions due to interpolation errors. This dataset also consists of 32-ensemble members that span a variety of sensitivity parameters and thresholds for interpolating to nearby radiosonde time series (Haimberger et al., 2012).

4.4 Atmospheric temperature profiles

The radiation interactions described in Section 4.3 will lead to changes in atmospheric temperature. Namely, a warming of the stratosphere due to aerosol absorption of outgoing longwave radiation and a cooling of the surface due to reflection and scattering of incoming solar radiation by the aerosol plume. Figure 5 shows the 1992 annual mean atmospheric temperature anomalies (subtracting the 1990 annual mean) in the models (Fig. 5a-d), MERRA-2 reanalysis data (Fig. 5e), and the RICH-obs radiosonde product (Fig. 5f). The year 1992 was chosen given the highest model spread in TOA flux (Fig. 4), peak modeled reduction in ASR (Fig. 4a) and reduction in OLR (Fig. 4b), and peak surface cooling (Parker et al., 1996) over this period. Models and observations share similar anomaly spatial patterns, with the exception of RICH-obs in the 60S-90S upper troposphere and near the tropical tropopause. Differences in RICH-obs may be related to temperature interpolation errors introduced in these remote regions due to fewer radiosonde datasets (Haimberger et al., 2012; Free and Lanzante, 2009). There is greater stratospheric warming in E3SMv2-SPA (Fig. 5b), E3SMv2-presc (Fig. 5c), and CESM2-WACCM (Fig. 5d) compared to E3SMv2-PA (Fig. 5a). Furthermore, there is an improvement in midlatitude warming at higher altitudes (i.e., 50 hPa and above) over E3SMv2-PA when comparing Fig. 5a-d to observations (Fig. 5f), reflecting the higher plume heights in these models (Fig. S6). CESM2-WACCM and MERRA-2 have very similar temperature magnitude and distribution, which is due to temperature nudging of CESM2-WACCM to the latter reanalysis dataset. There is not as obvious a surface cooling difference between E3SMv2-PA and other models and observations. All datasets show a large cooling signal in the northern troposphere that roughly corresponds with early-1992 max AOD between 30°N and 50°N (Fig. 3), but this cooling signal could be influenced by internal variability in the normalization year of 1990 (Section 4.3).

Table 2 shows 50 hPa and 850 hPa pressure level averages from Fig. 5. These comparisons represent stratospheric and near-surface changes in temperature, with the 850 hPa level chosen to accommodate the lowest pressure level in the RICH-obs data. These latitude-weighted averages range from 65°S-65°N to avoid missing data in the upper atmosphere and surface RICH-obs data (Fig. 5f). This comparison shows stratospheric warming that is overestimated E3SMv2-SPA (1.57°K) and underestimated in CESM2-WACCM (0.9 °K) compared to MERRA-2 reanalysis (0.89 °K) and previously reported estimates of ~1°K (Ramachandran et al., 2000). RICH-obs struggles to represent lower stratospheric warming either due to aforementioned sparcity of data and/or its high horizontal resolution (5°) compared to models (1°) and MERRA-2 (0.5°). E3SMv2-presc shows a more than three times the stratospheric warming of MERRA-2, which is likely due to a known error converting CLAES infrared extinction to the SAGE-II and GloSSAC V1 reported 1020 nm extinction coefficient, resulting in an exaggeration of peak aerosol extinction (Kovilakam et al., 2020). The 850 hPa cooling in CESM2-WACCM (-0.33 °K) agrees best with MERRA-2 (-0.36 °K) and RICH-obs (-0.29±0.007 °K) anomalies, due in part to nudging of CESM2-WACCM temperatures to MERRA-2. There is small improvement in E3SMv2-SPA (-0.23 °K) and E3SMv2-presc (-0.26 °K) compared to E3SMv2-PA (-0.22 °K), but it is unclear how much internal variability is influencing these values.

This comparison gives an all-sky snapshot of surface and stratospheric 1992 temperature anomalies due to Pinatubo. The 50 hPa show a clearer improvement in simulated temperature anomaly in E3SMv2-SPA and CESM2-WACCM than 850 hPa height due to the influence of interannual differences in internal variability (Section 4.3) and internal modes of variability (e.g., ENSO; Santer et al., 2014) in the troposphere. The model trends in stratospheric and near-surface temperature changes are consistent with changes in OLR and ASR (Fig. 4), respectively. Temperature trends also tend to agree better with observations and reanalysis with stratospheric volcanic parameterizations (E3SMv2-SPA, CESM2-WACCM) and prescribed volcanic aerosol (E3SMv2-presc). The next sections explore the microphysical representation within the models and how this influences lifetime, AOD, TOA flux, and temperature.

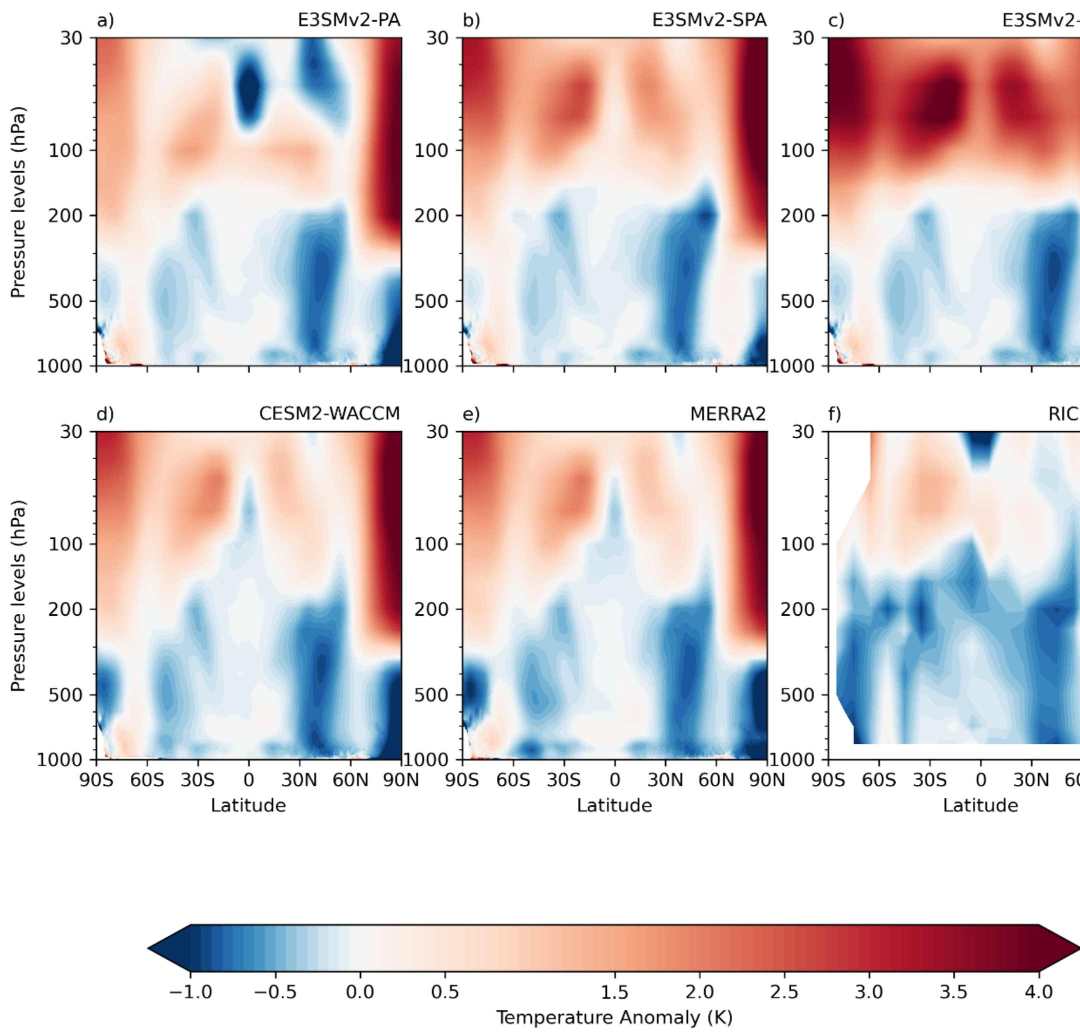

Figure 5: Annual mean change in atmospheric temperatures (˚K) for the year 1992, shown for (a-d) model simulations,€) MERRA-2 reanalysis, and (f) RICH-obs radiosonde dataproduct. Anomalies are calculated by taking the difference between year 1992 and 1990 annual means. Model data is remapped from hybrid vertical coordinate to MERRA-2 pressure levels.

Table 2: Annual mean temperature anomalies at 50 hPa and 850 hPa levels, shown from model simulations, MERRA-2 reanalysis, and RICH-obs radiosonde data product. As for Fig. 5, anomalies are calculated as the difference between the year 1992 and 1990 annual means. Data is averaged over 65S-65N to avoid missing data in RICH-obs in the Antarctic. Included in the RICH-obs is one standard deviation about the 32-member ensemble spread.

	Pressure level	**E3SMv2-PA**	**E3SMv2-SPA**	**E3SMv2-presc**	**CESM2-WACCM**	**MERRA-2**	**RICH-obs**
Temperature anomaly (K)	50 hPa	-0.17	1.57	3.03	0.9	0.89	0.4±0.015
	850 hPa	-0.22	-0.23	-0.26	-0.33	-0.36	-0.29±0.007

- Line 667-668: Here you can also refer to work of Clyne et al (2021) and Quaglia et al 2023
    - We have added references to AOD results from these publications.

Figures:

- Figure 1: The figure caption is very short, lacks information and is difficult to understand, e.g. What does the grey shaded region indicate? What does for "mode sensitivity tests" mean? Are you referring to the global sulfate burden?

  Thank you for pointing this out. We have included more information in the figure caption, including some information that was previously included in the beginning of section 4.1. It now reads: **"Figure 1: Stratospheric sulfate burden – reported in Tg of the sulfur mass contribution – for model simulations, as well as HIRS and SAGE-3λ remote sensing observations. The model data is processed to match the HIRS and SAGE-3λ data coverage of 80˚N – 80˚S above the model lapse rate tropopause height. The sulfur component is determined by scaling modeled sulfate mass by the ratio of sulfur and sulfate mass weights (MW) such that $Tg\ S = Tg\ SO_4 * \frac{32.066\ g\ mole^{-1}}{MW\ Sulfate}$. In MAM4 of E3SM, sulfate is assumed to be ammonium bisulfate ($(NH_4)HSO_4$; MW= 115.11 g mole^{-1}) (Liu et al., 2012, 2016). In MAM4 of CESM2-WACCM, sulfate is assumed to be sulfuric acid (H_2SO_4; MW= 98.08 g mole^{-1}) (Mills et al., 2016). Gray shading around the HIRS data represents systematic error of ~10% (±1.4 Tg aerosol) and the minimum and maximum aerosol composition bounds (59%–77% H_2SO_4)."**

- Figure 2: Again, the gray shading?

  We have added the following sentence to the end of the Figure 2 caption: "The gray shading indicates ±11.3% uncertainty in AVHRR AOD."

- Figure 3: Citation of Quaglia et (2023) is misleading here, as it is a model intercomparison paper which uses the observational data for comparison and validation.

  We have changed this to the citation to the source datasets from the Quaglia et al., 2023 paper: Quaglia, I., Niemeier, U., Visioni, D., Pitari, G., Brühl, C., Dhomse, S. S., Franke, H., Laakso, A., Mann, G. W., Rozanov, E., and Sukhodolov, T.: Data from: Interactive Stratospheric Aerosol moels resonse to different amount and altitude of SO2 injections during the 1991 Pinatubo eruption, https://doi.org/10.7298/mm1s-ae98, 2022. This maintains the source of the postprocessed data for reproducibility. We make sure to properly cite the instrument publications in section 3.

Table:

- Table 1: you can get rid of the third column and include the text in the table caption

  We have removed the nudging column and added this information to the table caption: "Table 1: Model details for the simulations used within this study. All simulations are run for 5 years (1989-1993) with 1989 discarded for aerosol spinup. All E3SMv2 simlations are run with U + V winds nudged to MERRA2 reanalysis data; CESM2-WACCM has U + V winds and temperature nudged to MERRA2 reanalysis."

Literature

- Barkstrom, B. R.: The Earth Radiation Budget Experiment (ERBE), Am. Meteorol. Soc., 65, 1170–1185, 1984.
- Barkstrom, B. R. and Smith, G. L.: The Earth Radiation Budget Experiment: Science and implementation, Rev. Geophys., 24, doi:10.1029/RG024i002p00379, 1986.
- Brodowsky, C., Analysis of the global atmospheric background sulfur budget in a multi-model framework, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-1655, 2023.
- Clyne, M., et al..: Model physics and chemistry causing intermodel disagreement within the VolMIP-Tambora Interactive Stratospheric Aerosol ensemble, Atmos. Chem. Phys., 21, 3317–3343, https://doi.org/10.5194/acp-21-3317-2021, 2021.
- Free, M., and J. K. Angell, 2002: Effect of volcanoes on the vertical temperature profile in radiosonde data. J. Geophys. Res., 107,4101, doi:10.1029/2001JD001128
- Free, M., and J. Lanzante, 2009: Effect of Volcanic Eruptions on the Vertical Temperature Profile in Radiosonde Data and Climate Models. *Climate*, **22**, 2925–2939, https://doi.org/10.1175/2008JCLI2562.1.
- Kremser, S., et al.: Stratospheric aerosol – Observations, processes, and impact on climate, Rev. Geophys., 54, 1–58,https://doi.org/10.1002/2015RG000511, 2016.
- Marshall, L., et al..: Multi-model comparison of the volcanic sulfate deposition from the 1815 eruption of Mt. Tambora, Atmos. Chem. Phys., 18, 2307–2328, https://doi.org/10.5194/acp-18-2307-2018, 2018.
- Quaglia, I.et al.: Interactive stratospheric aerosol models' response to different amounts and altitudes of SO2 injection during the 1991 Pinatubo eruption, Atmos. Chem. Phys., 23, 921–948, https://doi.org/10.5194/acp-23-921-2023, 2023.
- Tilmes, S.et al.: Description and performance of a sectional aerosol microphysical model in the Community Earth System Model (CESM2), Geosci. Model Dev., 16, 6087–6125, https://doi.org/10.5194/gmd-16-6087-2023, 2023.
- Timmreck, C. et al.: The Interactive Stratospheric Aerosol Model Intercomparison Project (ISA-MIP): Motivation and experimental design, Geosci. Model Dev., 11, 2581-2608, doi.org/10.5194/gmd-11-2581-2018, 2018.

- Toohey, M., Krüger, K., Niemeier, U., and Timmreck, C.: The influence of eruption season on the global aerosol evolution and radiative impact of tropical volcanic eruptions, Atmos. Chem. Phys.,11, 12351–12367, doi:10.5194/acp-11-12351-2011, 2011.
- Weisenstein, D. K., Visioni, D., Franke, H., Niemeier, U., Vattioni, S., Chiodo, G., Peter, T., and Keith, D. W.: An interactive stratospheric aerosol model intercomparison of solar geoengineering by stratospheric injection of SO2 or accumulation-mode sulfuric acid aerosols,Atmos. Chem. Phys., 22, 2955–2973, https://doi.org/10.5194/acp-22-2955-2022, 2022.