We thank Thomas Aubry for making very useful suggestions to improve the paper. Our point-by-point responses to the reviewers' comments and corresponding changes are detailed below in blue text, and the changes are shown in the version of the manuscript with track changes.

Comment on acp-2021-654
Thomas Aubry (Referee)

Referee comment on "Radiative forcing by volcanic eruptions since 1990, calculated with a chemistry-climate model and a new emission inventory based on vertically resolved satellite measurements" by Jennifer Schallock et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-654-RC2, 2021

This study investigates the impact of explosive volcanic eruptions on the stratospheric aerosol burden, optical depth and Earth's radiative balance using a chemistry-climatemodel and a new inventory of volcanic SO2 emissions. Comparison with satellite observations are presented and with the exception of the Pinatubo period, the simulations are shown to be in good agreement.

Although the overall methodology of this paper is not new, there exists very few studies of the historical stratospheric aerosol forcing that use chemistry-climate models and SO2 inventories and this paper is thus a useful contribution. Furthermore, the study brings two novel elements:

i) the use of a new volcanic SO2 emission inventory, argued to be more comprehensive and better compared to other inventories;

ii) ii) the strategy used to inject volcanic SO2 into the model, consisting in injecting 3D SO2 plumes instead of the standard "point source" injection. Unfortunately, I find these two points to be not sufficently motivated and explained (in the case of the second one), and analyses conducted do not enable to assess whether the new inventory and injection strategy result in improved volcanic forcing, which undermines the key contributions of this study. Furthermore, there is little to no comparison with previous work (e.g. different emission inventories, or different emission strategies). Many important references are lacking. To sum-up, I think this manuscript has the potential to become a really valuable paper for the community, but that further analyses as well as an improvement of the discussion section are still required.

Major comments:

1) The first novel aspect of the paper is the way in which volcanic SO2 is injected in the model. Previous studies have used a "point-source" approach with SO2 injected in one model column over a range of altitudes, with a few studies also injecting over a range of latitude for Pinatubo. However, in this study, the authors instead inject a "plume" consistent with spatially-resolved satellite observations. First, I think that this novel aspect is not highlighted enough in the introduction section and throughout the text, and it could be one of the key point of the manuscript.
-> Response: These points are now more highlighted in the introduction:
"For the ENVISAT (European Environmental Satellite) period 2002-2012 a first version of a new volcanic SO2 inventory with improved temporal and spatial resolution was developed within the framework of ISA-MIP (Timmreck et al. 2018, Brühl et al. 2018). The corresponding data base (link) contains 3D-SO2-perturbations derived from satellite data as well as integrated injected SO2 masses. In this work the data base is expanded to the period 1990-2019 and considerably improved for the period 1998-2001. The simultaneous measurements from up to four instruments from 2002 to 2011 enabled us to develop a novel procedure for conversion of aerosol extinction to SO2 needed for the period before and after ENVISAT.
Our method circumvents problems and uncertainties related to the classical point source approach like dependence on the box size and location, the time interval during which the mass is injected, and

effects of microphysical and chemical interactions of SO2 and sulfate with injected volcanic ash and water in the early phase (Zhu et al., 2020)."

I also find your new method to be poorly explained and justified, in particular in section 5. On line 264, you say that the total amount of SO2 is calculated by integrating the SO2 profile but then mentioned that you add a 3-dimensionnal perturbation to the model which confused me. In section 5, you also don't clearly state how these 3D plumes are obtained. My understanding from sections 3/4/5 is that:
-> Response: Parts of section 4 are rewritten to clarify the method. For illustrative purposes, a detailed description of case studies is added in Appendix B (see reply to referee 1).

For each eruption, 3D SO2 plumes are obtained from time-averaged SO2 observations between the 8th and 17th day following each eruption?
-> Response: Yes, 3D SO2 plumes are obtained from the satellite observations. The temporal resolution of MIPAS, GOMOS and OSIRIS data is 5-days. The chosen time interval for the integration of the emitted SO2 amount is case dependent for every single eruption, depending on data availability (data gaps, volcanic ash plumes, duration of the eruptions, etc.).
MIPAS SO2 data are used immediately after the eruption, extinction data with a time lag of about a week. This is now explained in section 4 in more detail for each data source.

The 3D plumes, obtained from measurement 8-17 days after the eruption, are injected at the time of the eruption. The 3D plumes are injected at latitude consistent with measurement taken but centered on the longitude of the volcano Did I get this right? It all need to be crystal-clear and more detailed in the text as this is key to your method and a very unusual approach?
-> Response: As mentioned in the text, it should be noted as well that the date of the volcanic eruption can differ by a few days from the date of injection in the model simulation, because the temporal resolution of the data sets is about five days at least (or weeks in the SAGE period).

You need to justify these choices better and show sensitivity tests for a large and small eruption (or ideally a full 1990-2019 simulation) showing how this differ from a standard "point" injection at the volcano location/plume height with a mass of SO2 corresponding to the initial total SO2 (not the SO2 after 8-17 days). Such tests seems really critical to demonstrate that your proposed method is better than standard methods, otherwise any related claim is unfounded.

One of the main justification you provide to justify your injection strategy is that it removes any tropospheric SO2 that is not climatically relevant but:

i) you already only consider SO2 above a threshold height (which is not justified; e.g. why 14km at the tropics instead of the tropopause height? If it's because of radiative heating and lofting where does the threshold come from?) so why do you need further processing to remove potential "shortlived" SO2?;
-> Response: The EMAC simulations include comprehensive tropospheric chemistry and various inventories of tropospheric sulfur emissions (SO2 from outgassing volcanoes, anthropogenic emissions, etc., as well as other sulfur-containing species such as DMS from ocean fluxes). Therefore, double counting of tropospheric volcanic emissions should be avoided.The lower limit of 12 km altitude at high latitudes was chosen based on the signal-to-noise ratio, uncertainties for low altitudes, and clouds in the volume mixing ratio profiles obtained by MIPAS and the other used satellite instruments. In the tropics, we set the lower limit at 14 km to account for transport processes in the UTLS layer, especially during the Asian summer monsoon and over the Andes. Here, we also have to consider high altitude clouds.

ii) The SO2 e-fold time is on the order of days-weeks (Carn et al. 2016, Fig 14); Even for stratospheric SO2 one would expect a significant amount of SO2 to be already converted to aerosol by the end of your 8-17 day time window, in particular for lower stratospheric injections. So would your method not result in large underestimation of SO2 amounts injected? I can see reasons why your method could make sense, e.g. fast SO2 scavenging

by ash during the first days-weeks, but I think it is still not justified enough in the paper. More importantly, you need to show comparison between your approach vs standard point injection with the full SO2 mass to be able to really discuss the strengths and weaknesses of your strategy.
-> Response: MIPAS directly provides SO2 data, while for the other 3 instruments we have to convert the aerosol extinction. Here we consider a time lag of some days to few weeks from the conversion of SO2 to aerosol. During the MIPAS period direct SO2 observation could be used for development and validation of a conversion formula for the 750 nm extinction seen by GOMOS and OSIRIS, which works also approximately for SAGE (see for detailed description the case studies in Appendix B).

2) Overall, your paper really lacks comparison with existing work – including that from Bruhl et al 2015 – and a lot of key references are missing. As an example, on line 245-247, you suggest that your SO2 mass estimates will be very different from those in the dataset by Carn et al. (2016). Why not show a figure, at least in SI, comparing SO2 masses and heights for all events in common? This would be really informative.
Regarding your simulations, you do not mention at all the work by Schmidt et al. (2018) which conducted exactly the same type of simulations, albeit with a different SO2 inventory and model. Citing it seems critical, and some of their time series (SAOD, radiative forcing) are likely available and could be compared to your model which would really improve the discussion.
Also, it would have been nice to see a comparison of your new simulations with the previous model version/inventory used by some of the co-authors (Bruhl et al 2015) to get a sense of whether there is improved agreement with observations.
-> Response: Pink lines with the results from Brühl et al. (2015) are added to Fig. 9-11. Schmidt et al. (2018) is available for global AOD at 550 nm and added as black line in Fig. 9, AOD at 750 nm (Fig. 10) is not shown by Schmidt et al. (2018). In Fig. 11 the data for volcanic effective radiative forcing from Schmidt et al. (2018) is added by a black line.
A comparison with Carn et al. (2016) for annual sums is added in Appendix C.

Last, you compare your simulations with observations from multiple satellite instruments which is welcome, but I was under the impression that the GloSSAC dataset – built using some of the data you use – is now the reference for the community (at least for CMIP6 forcing). Could you add a comparison to GloSSAC?
-> Response: GLOSSAC (Thomason et al. 2018), a time dependent aerosol climatology sometimes used for climate simulations, has a coarse temporal resolution, and many discontinuities in space and time due to change of instruments or gaps and excludes important satellite data (e.g. MIPAS). It does not provide SO2 (sulfur dioxide) needed as input for chemistry climate models directly but only extinction and highly derived quantities like estimates for aerosol surface area density and mode radius assuming unrealistic monomodal size distributions. (Added in introduction)
A line for GLOSSAC is added in Fig.9 for convenience.

Minor comments

Title: I think the title does not convey clearly enough the novelty of the new injection method; consider replacing “vertically-resolved satellite measurements” by something else? Maybe “Reconstructing volcanic forcing since 1990 using a comprehensive volcanic emission inventory and spatially resolved sulfur injection in a chemistry-climate model”?
-> New title: “Reconstructing volcanic radiative forcing since 1990, using a comprehensive emission inventory and spatially resolved sulfur injections from satellite data in a chemistry-climate model”

Your 3D plume are not just vertically resolved?
-> Response: No, we are working with 3-dimensional plumes.

Abstract: the long list of satellite instruments and their acronym is not needed in an abstract?
-> Response: Names and acronyms of satellites are removed.

I find that the abstract does not highlight enough the novel and extensive character of the SO2 emission inventory nor the 3D plume injection method.
-> Response: Parts of the abstract are rewritten to highlight the novel character of the paper:
"This paper presents model simulations of stratospheric aerosols with a focus on explosive volcanic eruptions. Using various (occultation and limb-based) satellite instruments, with vertical profiles of sulfur dioxide (SO2) and vertical profiles of aerosol extinction, we characterized the influence of volcanic aerosols for the period between 1990 and 2019.
We established an improved and extended volcanic sulfur emission inventory that includes more than 500 explosive volcanic eruptions reaching the upper troposphere and the stratosphere. Each perturbation identified was derived directly from the satellite data and incorporated as a three-dimensional SO2 plume into a chemistry-climate model. The simultaneous measurements of SO2 and aerosol extinction by up to four instruments enabled us to develop a reliable method to convert extinction measurements into injected SO2. In the chemistry climate model, the SO2 from each individual plume is converted into aerosol particles and their optical properties are determined. Furthermore, the Aerosol Optical Depth (AOD) and the instantaneous climate radiative forcing are calculated online. Combined with model improvements, the simulations are consistent with the observations of the various satellites.
Slight deviations between the observations and model simulations were found for the large volcanic eruption of Pinatubo in 1991 and cases where simultaneous satellite observations were not unique or too sparse. Weak- and medium-strength volcanic eruptions captured in satellite data and the Smithsonian database typically inject about 10 kt to 50 kt SO2 directly into the upper troposphere/lower stratosphere (UTLS) region or transport it indirectly via convection and advection. Our results confirm that these relatively smaller eruptions, which occur quite frequently, can nevertheless contribute to the stratospheric aerosol layer and are relevant for the Earth's radiation budget. These eruptions cause a global radiative forcing of the order of −0.1 Wm^{-2} at the tropopause (compared to a background aerosol forcing of about −0.04 Wm^{-2})."

Abstract, lines 17-20: you say that your results "show" and that eruption "are found to"; I would instead say that your "confirm" these results as this has been shown by Schmidt et al. (2018)?
-> Corrected: "confirm"

Introduction: Also see major comment 1: the two main novelties of your study are overall not motivated in your intro (i.e. new injection strategy and improved SO2 inventory).
-> Response: These points are now more strongly highlighted in the introduction:
"For the ENVISAT (European Environmental Satellite) period 2002-2012 a first version of a new volcanic SO2 inventory with improved temporal and spatial resolution was developed within the framework of ISA-MIP (Timmreck et al., 2018; Brühl et al., 2018). The corresponding data base (https://doi.org/10.1594/WDCC/SSIRC_1) contains 3D-SO2-perturbations derived from satellite data as well as integrated injected SO2 masses. In this work the data base is expanded to the period 1990-2019 and considerably improved for the period 1998-2001. The simultaneous measurements from up to four instruments from 2002 to 2011 enabled us to develop a novel procedure for conversion of aerosol extinction to SO2 needed for the period before and after ENVISAT. Our method circumvents problems and uncertainties related to the classical point source approach like dependence on the box size and location, the time interval during which the mass is injected, and effects of microphysical and chemical interactions of SO2 and sulfate with injected volcanic ash and water in the early phase (Zhu et al., 2020)."

Introduction: I think the work of Mills et al (2016) and Schmidt et al 2018 (not cited) need to be discussed more given strong similarities with your study. Also you don't mention ISAMIP at all (Timmreck et al 2018) whereas your simulations are obviously relevant to this MIP?

-> Added citations: Timmreck et al. 2018: “For the ENVISAT (European Environmental Satellite) period 2002-2012 a first version of a new volcanic SO2 inventory with improved temporal and spatial resolution was developed within the framework of ISA-MIP (Timmreck et al., 2018; Brühl et al., 2018).”

Line 36: is it important to specify at which level it affects Earth radiative balance? If so also mention surface level in addition to TOA and tropopause.
-> Response: The forcing at that the surface cannot be retrieved by satellites, i.e. this would show model only.

Line 38: Multiple papers discuss how climate-volcano feedback could modulate future volcanic forcing though, and it may be a good place to mention it? See e.g. Swindles et al. (2017) (deglaciation effect on eruption frequency), Fasullo et al. (2018) (modulation of volcanic influence on surface temperature by changes in ocean stratification), Aubry et al. (2021) (impact of climate change on the volcanic stratospheric sulfate aerosol cycle).
-> Suggested references added.

Line 40: unless I misunderstand I guess you are talking about (mostly CMIP5) simulations that did not account for this forcing? Many model studies have accounted for this forcing since then, including CMIP6 historical simulations that use GloSSAC or e.g. Mills et al. (2016) and Schmidt et al. (2018)?
->These references and an additional one and more text added, see also reply to major comments.

Line 43/44: please add references
-> Response: See above.

Line 46: do you mean “overlooked” instead of “underestimated”? If not what was underestimated? Their radiative forcing? But does it not contradict the previous sentence?
-> Response: Text rewritten here, now related to GloSSAC and ISA-MIP.

Line 50: The SO2 emission and time-averaged volcanic forcing of degassing volcanoes and small eruptions is one order of magnitude larger than that of eruptions associated with stratospheric SO2 injections (e.g. Schmidt et al. 2012, Carn et al 2016). So clarify what you mean by “smaller natural source of aerosols” as this seems wrong as written.
-> Corrected: “~~a smaller~~” natural source -> “another” natural source

Section 2: could you group satellite instruments in terms of those used to constrain SO2 inputs in your model vs those used to evaluate the output of the model simulations? This would add a lot of clarity to this section. Also why not using GloSSAC (Thomason et al. 2018, 2020)?
-> Response: Data from all four satellite instruments described in this section are used as input to the simulations. MIPAS directly provides SO2 mixing rations, while for GOMOS, OSIRIS and SAGE II the SO2 input data are derived from aerosol extinction (for details see Sec. 4 and Appendix B). As explained in the introduction GloSSAC does not provide what we need.

Line 119-120: as said in my major comment I think you need to discuss the strength and limitations of choosing such a time window, and in particular how it compares to the SO2 e-folding time and the fact that choosing this time window may result in neglecting a large portion of SO2 converted to aerosols (even though I understand the argument that an earlier time window could account for SO2 estimates accounting for SO2 that will be rapidly scavenged by co-injected ash or hydrometeors; but this all needs to be discussed carefully). Sensitivity tests for this time window and understanding its impact on your SO2 estimates would be welcome.
-> Response: The satellite data (for MIPAS; GOMOS and OSIRIS) are provided in 5-day time intervals. These processes are important if point sources are used since the occur in the very first days. Use of a period of about at least 10 days of satellite date circumvents this, see introduction and Appendix B. In our institute studies on the early phase processes are in progress, also we included 2 references on this (Zhu et al. 2020 and Clyne et al. 2021).

Line 137: again this time window needs to be justified better. Also I'm not at all a remote sensing expert but I think it's the first time I see SO2 estimated from extinction coefficients in visible wavelength? Is that a standard method? How is the effect of SO2 on radiation properties isolated from other species, in particular sulfate aerosols? It may be standard techniques that I'm not aware of about but it would be good to clarify.
-> Response: The time lag is needed to allow for production of particles from oxidation of the injected SO2. Our method is a novel alternative also addressed in the introduction now. The part you refer to has been moved to section 4 and includes more details.

Line 139: My understanding here is that you are saying that if there is a data gap during the peak perturbation, you scale up by an arbitrary factor to recover a reasonable peak value? How is that factor chosen? There is absolutely no explanation nor reference and it may deserve dedicated SI plots?
-> Response: To estimate the factor, we iterated calculated extinctions to agree with OSIRIS and also used observations and assumptions by Vernier et al. (2016). A detailed description with 3 case studies is added to the Appendix:
"If the time lag of data is several weeks a correction factor >1 has to be applied to account for removal processes, if another event occurs relatively close in time, the factor has to be <1 to remove the influence of the previous event." See section 4 now.

Line 139 and 170: about data gaps and how to treat them, I'm just wondering why not using GloSSAC where the same problem had to be addressed and which is the reference dataset for the community? I understand you can't use it for SO2 but surely for aerosol properties it would make sense? The fact that major initiatives such as GloSSAC or ISAMIP are not mentioned is a bit surprising.
-> Response: GloSSAC does not solve problems with data gaps in some cases. If there is no other data set available it just smears out the gap (see also remark in introduction). However, for Pinatubo, SAGE data could be improved by CLAES data. In ISAMIP also an earlier version of our 5-day-dataset (link in Brühl et al. 2018 or Timmreck et al. 2018) is included.

Line 236: no apostrophe needed for Global Volcanism Program
-> Corrected.

Line 241: The tropopause altitude varies between ca. 8-9 and 16-17km depending on latitude and season, why not using the model diagnostic tropopause instead of the three thresholds used? Justify rigorously why you consider a threshold way below the tropopause height in the tropics but potentially way above at high latitudes. Also why do you need to mask tropospheric SO2? Would your model not account for the fact that tropospheric aerosol would have minimal impact on climate? I get that you don't want an overlap between the tropospheric and stratospheric volcanic SO2 inventory, but does the tropospheric SO2 inventory really account for emissions as high as 12-14 km or is it only passively degassing volcanoes?
-> Response: The EMAC simulations include comprehensive tropospheric chemistry and various inventories of tropospheric sulfur emissions (SO2 from outgassing volcanoes, anthropogenic emissions, etc., as well as other sulfur-containing species such as DMS from ocean fluxes). Therefore, double counting of tropospheric volcanic emissions should be avoided.
The lower limit of 12 km altitude at high latitudes was chosen based on the signal-to-noise ratio, uncertainties for low altitudes, and clouds in the volume mixing ratio profiles obtained by MIPAS and the other used satellite instruments. In the tropics, we set the lower limit at 14 km to account for transport processes in the UTLS layer, especially during the Asian summer monsoon. Here, we also have to consider high reaching clouds.

Lines 243-247: see my major comment #1
-> See response to major comment #1.

Table 2: this table really must be made available as a csv file or something that researchers can download and read in scientific programming software. Remove the table from the body of text as it is way too big.

-> Response: As an essential part of the novelty of this paper, the table should remain in the text because it is a comprehensive reference that cannot be represented by a single visualization. Additionally, the table is available for data processing at this link: https://doi.org/10.26050/WDCC/SSIRC_3

Lines 257-264: see my major comment #1. While I think this is at the moment poorly explained and that you have to show analyses demonstrating the advantages and challenges with this injection method, I do think that it is one of the most novel and important aspect of the paper (combined with your inventory) and that it should be highlighted and motivated a lot more.
-> Response: We added a detailed explanation with case studies in Appendix B (see reply to referee 1).

Line 275-276: you either need a reference backing this claim or data analysis to support it (e.g. does the GVP database have a comparable number/frequency of VEI 3-5 eruptions during 1991-2002 relative to 2002-present day? Or was it really a more quiescent period?
-> Response: The VEI index was developed for the volcanic explosivity, but is not a direct indicator for the climate relevance of volcanic eruptions, e.g. the eruption of Eyjafjallajökull in 2010 VEI4 had almost no influence on the stratosphere.

Indeed, there was a relative volcanically quiescent period from 1999–2002 (Schmidt et al. 2018), there are also less entries in the Smithsonian database, but this does not explain the low number of detected volcanic eruptions in the years directly after Pinatubo (1992-1998).

Lines 293-294: a brief comparison with observations in Carn et al. 2016 would be welcome here (I think they suggest even lower UTLS e-folding time). Also you say yourself here that the conversion time is about 2 weeks, which seems to strongly undermine your chosen 8-17 day time window to constrain SO2 emission from satellites?
-> Inserted:
" The conversion of SO2 to sulfate aerosol particles depends on several factors, such as the altitude, latitude, or season of the eruption and takes about 13, 23 and 32 days in 10–14, 14–18 and 18–22 km altitude (Höpfner et al 2015, midlatitudes), while Carn et al. (2016) report an e-folding time varying between 2-40 days. The range agrees with our simulations (and assumptions in section 4). "

Line 303-304: please clarify what you mean by "feedback to atmospheric dynamics" and cite appropriate references
-> Response: This means that radiative heating implies an enhancement of upward motion (or cooling a descent).

Line 309-310: the reader has to look at three different figures and compare them to verify this statement. It would me much better if you could present equivalent observations and model plots on the same figure and different panels. This would greatly facilitate modelobservation comparisons.
-> Response: In figure 8 we added an additional panel with observations from SAGE II and OSIRIS for the direct comparisons with the model simulations for aerosol extinction at 750 nm wavelength:
"Figure 8. Comparison of aerosol extinction for 750 nm wavelength at 17 km altitude between the model simulations (lower panels) and SAGE II and OSIRIS satellite data (upper panel). EMAC simulation of the stratospheric aerosol extinction are given on a logarithmic scale log(1/km) for 750 nm wavelength from January 1991–August 2019 based on the volcanic sulfur emission inventory (Table 2), in horizontal T63 resolution of zonal mean at 17 km altitude (middle) and in vertical distribution for tropical regions20° S–20° N (bottom). Maximum and minimum values appear above (dark red) and below (violet) the color keys, respectively."

Line 326: the vast majority of studies use SAOD at 550nm like you (e.g. Schmidt et al. 2018), and also 1020nm (e.g. Aubry et al. 2021) which is another standard one for some instruments? So this statement seem really not justified and should be removed or modulated.
-> Modified.

Line 331: clarify that the AOD of 0.4 is in the tropics and isn't a global mean value -> done: "with a stratospheric AOD of 0.4 in the tropics"
-> Corrected: "with a stratospheric AOD of 0.4 in the tropics"

Line 334: There could be other factors explaining model-observation differences in the post-Pinatubo period including flaws in the model (as evident from the different decay timescales) and uncertainty in the SO2 mass, or at least the "climatically relevant" portion of it (you use 17Tg, other studies use as little as 10 which should be briefly discussed; see Zhu et al. 2021, Mills et al. 2016, Schmidt et al. 2018).
-> Response: In this study about 17 Tg SO2 are injected for the Pinatubo eruption. Model comparisons by Timmreck et al. (2018) show that the range of used injections varies between 10 Tg SO2 (e. g. Mills et al. 2016, Schmidt et al. 2018) and 20 Tg SO2 (e.g. English et al. 2013). Thus, this study is in the middle range of injected sulfur mass.

Line 337: unless major eruptions are missing, is it really likely that imperfections in your inventory explain the large SAOD differences over 1993-1996?
-> Response: Between 1993 and 1996 the reduction of the stratospheric AOD in the model simulations is faster than indicated by the satellite observations and in Schmidt et al. (2018). This indicates that the removal of stratospheric aerosol is still too rapid from applying the modal model. Schmidt et al. (2018) show a slower decrease in AOD after the Pinatubo eruption. This could indicate that EMAC still needs better fine-tuning of the size distribution modes, or adding modes in the aerosol submodel to improve the aerosol removal in the stratosphere. Here the sectional aerosol model used by Schmidt et al. (2018) might have an advantage. Additionally, smaller volcanic eruptions might be missing, in view of the low number of identified events in the years after the Pinatubo eruption.

Figure 11: it may be better to show horizontal bars (with a length of 1 year) instead of green crosses as these are time-average measurement and it would facilitate comparison with your high-resolution output?
-> Done: Green crosses are replaced by bars with a length of 1 year.

Figure 11: Here and on Figure 9 and 10, could you not show for comparison the simulations from at least Bruhl et al. (2015) and maybe Schmidt et al. 2018 assuming their data are available with the paper? Discussing the differences would really improve the discussion.
-> Response: Pink lines with the results from Brühl et al. (2015/2018) are added to Fig. 9+11. Schmidt et al. (2018) is available for global AOD at 550 nm and added as black line in Fig. 9, AOD at 750 nm (Fig. 10) is not shown by Schmidt et al. (2018). In Fig. 11 the data for volcanic effective radiative forcing from Schmidt et al. (2018) is added as a black line.

Legend of Figure 11: specify the time resolution of the ERBE data. Is there no other observational estimate of radiative forcing to complement observations shown? E.g. CERES data?
-> Response: 72-day means are used in the near-global data set of ERBE (Toohey et al. 2011 fig. 2). In the AOD-figure we included AVHRR.
-> Question: CERES? Can you provide a reference?

Line 354: "previous studies"-> show their data and discuss comparison? On that note making sure that your key outputs (SAOD/radiative forcing time series) are easily available is important and I don't think it's the case yet? Key outputs should not be made "available upon request" but should ideally be provided as SI or in a data repository.
-> Response: The output for radiative forcing is now available at WDCC: (https://doi.org/10.26050/WDCC/SSIRC_3).

Line 359: For reference, can you indicate the SO2 mass for Merapi used in your and other (e.g. Carn et al. 2016) inventories? Overall, it would be really useful to have a comparison of your inventories with other standard ones, in particular those used in ISA-MIP (Timmreck et al. 2018). Another potentially useful reference, showing how different inventories affect the SAOD prediction by a simple model, is Aubry et al. (2020) (see Figure 8 there).

-> Response: See also Höpfner et al. (2015), Carn does not provide the stratospheric fraction. A range from our data sources is in Appendix B.

Figure 12: could you discuss how these results compare with recent studies, e.g. Rieger et al. (2020) or Stocker et al. (2019)
-> Response: Comparing our results with Rieger et al. (2020) shows that our results:
"...corresponds quite well with the results of Rieger et al. (2020) showing a maximum of instantaneous solar heating rate of 0.5 K/day in the tropics near 24 km plus thermal heating rates of about 0.2–0.3 K/day. "

Section 7: Overall I find that some of the most natural lines of discussion (and accompanying analyses) are completely missing including:

i) comparison of your new inventory with other ones, including Carn et al.;
-> Response: A table for comparison of annual global volcanic SO2 emissions between this study and Carn et al. (2016) is added in Appendix C.

ii) comparison of your new simulations with other equivalent ones, including Schmidt et al (2018) and Brühl et al (2015);
-> Response: A comparison with Schmidt et al. (2018) and Brühl et al. (2015) is added in Fig. 9 +11 and discussed in the text.

iii) discussion of how your 3D-plume injection strategy compares to a point injection.
-> Response: With our model this cannot be done since a proper module for point sources is not available yet. Possible problems of the point source approach are addressed in the introduction now.

Line 385: provide numbers (e.g. latitude resolution at equator) that make it easier for the reader to understand the difference between these resolutions.
-> Added: T42L90 (2.81°×2.81°) to T63L90 (1.87°×1.87°). This you find now only in section 3.

Line 410: Missing reference?
-> Response: This is novel and expanded. Sentence corrected.

Line 429-430: Zhu et al. (2020) should be cited here
-> Cited: Zhu et al. (2020)

Line 429-432: On model difference/setup and how it may affect simulated aerosol properties, Clyne et al. (2021) is an important difference and should be discussed here and elsewhere.
-> Response: Clyne et al. (2021) is cited now in section 6.2 in connection with the sectional model of Schmidt since they show similar differences between modal and sectional models concerning the behavior after a major eruption (Fig .9). The major eruptions are not the main scope of our paper.

Lines 448-455: This whole paragraph doesn't acknowledge the contribution of previous studies when most of the statements made are not really new. First maybe you should refer to the AR6 report now that it is out instead of the AR5? Second, for radiative forcing estimate, the contribution of Schmidt et al. (2018) should be acknowledged and you should compare in details your forcing estimates to theirs. Third, for temperature effects, you should cite the papers by Santer and co-authors (2014, 2015) and Schmidt et al. (2018). I personally think that the novel aspects of your paper would be highlighted better if you ended it on key points related to the new inventory and the 3D plume injection method.
-> Response: Parts added in abstract, introduction and conclusions. Surface temperature is not the focus of this study since we nudge the troposphere and prescribe a time dependent SST. Comparison with Schmidt is done in Section 6.2 and 6.3.

Added References:

Aubry, Thomas J., et al. "Climate change modulates the stratospheric volcanic sulfate aerosol lifecycle and radiative forcing from tropical eruptions." Nature Communications 12.1 (2021): 1-16.

Brühl, C., et al. "Stratospheric sulfur and its implications for radiative forcing simulated by the chemistry climate model EMAC." Journal of Geophysical Research: Atmospheres 120.5 (2015): 2103-2118.

Carn, S. A., Lieven Clarisse, and Alfred J. Prata. "Multi-decadal satellite measurements of global volcanic degassing." Journal of Volcanology and Geothermal Research 311 (2016): 99-134.

Clyne, Margot, et al. "Model physics and chemistry causing intermodel disagreement within the VolMIP-Tambora Interactive Stratospheric Aerosol ensemble." Atmospheric Chemistry and Physics 21.5 (2021): 3317-3343.

Fasullo, J. T., et al. "The amplifying influence of increased ocean stratification on a future year without a summer." Nature communications 8.1 (2017): 1-10.

Kovilakam, Mahesh, et al. "The Global Space-based Stratospheric Aerosol Climatology (version 2.0): 1979–2018." Earth System Science Data 12.4 (2020): 2607-2634.

Rieger, Landon A., et al. "Quantifying CanESM5 and EAMv1 sensitivities to Mt. Pinatubo volcanic forcing for the CMIP6 historical experiment." Geoscientific Model Development 13.10 (2020): 4831-4843.

Schmidt, Anja, et al. "Volcanic radiative forcing from 1979 to 2015." Journal of Geophysical Research: Atmospheres 123.22 (2018): 12491-12508.

Swindles, Graeme T., et al. "Climatic control on Icelandic volcanic activity during the mid-Holocene." Geology 46.1 (2018): 47-50.

Timmreck, Claudia, et al. "The interactive stratospheric aerosol model intercomparison project (ISA-MIP): Motivation and experimental design." Geoscientific Model Development 11.7 (2018): 2581-2608.

Zhu, Yunqian, et al. "Persisting volcanic ash particles impact stratospheric SO 2 lifetime and aerosol optical properties." Nature communications 11.1 (2020): 1-11.