We thank the reviewers 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

Anonymous Referee #1

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-RC1, 2021

This paper focuses on the injection of SO2 to the stratosphere by volcanic eruptions, and the resulting variability of the stratospheric aerosol layer. It presents a new volcanic SO2 emission database, derived from a collection of satellite instruments, covering the period 1990-2019. It also presents results from a chemistry climate model which uses the updated injection database, and compares the results of the model to various satellite data sets, focusing on the multi-wavelength aerosol optical depth and instantaneous radiative forcing produced by the aerosols.

The construction of such detailed SO2 injection estimates covering the 1990-2019 period is an impressive accomplishment. It is also to my knowledge quite novel, as I believe it is the first attempt to produce SO2 injection values from sulfate aerosol extinction measurements. Unfortunately, the description of the methods used to produce these estimates is lacking. Furthermore, assumptions and choices made in the methods are not given justification. More detailed comments are included in "Major Comments" below.

Chemistry climate model simulations using the new SO2 injection data set are performed and some results shown. Good agreement with observations is achieved, but there is insufficient analysis to provide any improved understanding of the physical or chemical processes that control stratospheric aerosol evolution.

Major comments:

The description of how SO2 amounts were calculated lacks sufficient detail. I am not aware of any other study that has estimated SO2 injection amounts based on aerosol extinction measurements. This is thus a novel technique, but the method used is not described beyond a few statements along the lines of "The SO2 mixing ratio perturbation is derived from the extinction perturbation observed in a 10-day period beginning about a week after the eruption by dividing by air density, multiplying by a constant and subtracting a typical background." This explains extremely little: what constant is used, and why? How is the typical background determined? How well can the volume of the aerosol cloud be estimated a week after eruption from the satellite measurements? SAGE in particular has a very sparse sampling density, how does this impact the estimates? Can the method be validated? It would seem that the method could be applied to SAGE and OSIRIS during periods of overlap with MIPAS and the values from the new method compared to the "direct" MIPAS measurements. This would help increase confidence in the method, and provide some idea of the uncertainties in the estimates.

-> Response: We added a detailed explanation with case studies in Appendix B:

"The eruption of Reventador in the tropics in November 2002 has shown to be an ideal case where simultaneous observations of all satellite sensors were available so that the direct SO_2 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 if its observations at 530 and 1025 nm are interpolated to 750 nm. Here we first use the ratio between model calculated sulfate volume mixing ratio and its share on extinction in low latitudes of the lower stratosphere which is typically $1.2\times10^{12}\times$air density (in molecules/cm^3). This works for medium size eruptions and data available over about four weeks following the eruptions, and if no other events occur less than about four weeks before which is the case for the Reventador eruption. 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 is relatively close in time, the factor has to be <1 to remove the influence of the previous event. For Reventador the factor is 1 (for OSIRIS 0.8 is slightly better). From all instruments the derived injected SO_2 mass is very close to 77 kt as shown in Table 2. The spatial patterns are similar, except when the zonal wind causes a shift in longitude due to the time lag from conversion of SO_2 to aerosol, see Figure B1. In the case of SAGE, the alternate method of Grainger et al. (1995) involving aerosol surface area density (SAD) and aerosol volume density is more suitable to remove cloud perturbation. It is assumed that sulfate mixing ratios correspond to the SO_2 injected. Some uncertainty remains from removing the background which we have done by subtracting a fraction of the derived SO_2 at the longitude where it has a minimum, i.e. the longitude where the effect of the volcano is smallest for all altitudes. Integrated injected SO_2 masses for all examples are provided in Table B1.
For the eruption of Merapi in November 2010 the satellite instruments do not agree. From OSIRIS about 70% more injected SO_2 is derived than from MIPAS, i.e. 170 kt instead of 97 kt used in the transient simulation (see Table 2 and differences in Figure 10). GOMOS has too sparse data here to obtain a proper integral directly but patterns are similar (Figure B2). If other information is available, the gaps can be filled with likely values in the region where the plume was seen, a method which had to be applied also to some events seen by OSIRIS in 2018 and 2019 for which the data were sparse.
For high latitude eruptions the longer conversion time of SO_2 to sulfate compared to the tropics has to be considered which, together with aerosol removal processes, lead to a weaker extinction signal. To account for this a correction factor of about two in the conversion formula for OSIRIS for example for Sarychev in June 2009 leads to values consistent to the ones derived by MIPAS (Figure B3). For the low latitude eruption of Mando Hararo in the same entry of Table 2 (separated at 24° N for the integration) the factor 1 is still appropriate."

I highly recommend that the emission database be provided as an electronic supplement (e.g., csv or xls), to allow it to be readily used by other researchers.
-> Response: The input data files and model output of EMAC used here are stored at DKRZ, Hamburg, the volcanic inventory and the output for radiative forcing also at WDCC https://doi.org/10.26050/WDCC/SSIRC_3

The table, as text, presently takes up almost 8 pages of the manuscript: it would be more efficient to visualize the data somehow and include the values as supplemental information. Also, I strongly suggest that the format of the table be modified so that each individual eruption be listed per row, even if there are multiple eruptions on a given date. This will greatly improve the ease in which the data can be read within a computer program and thus used in other studies.
-> Response: SO_2 mixing ratios from the volcanic emission inventory are shown in Fig. 6. 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 as formatted ascii at this link: https://doi.org/10.26050/WDCC/SSIRC_3

The model results show good agreement with observations, but it's impossible to know whether the improved agreement (compared to prior works from the same group) is a result of the updated SO2 injection data, or to model improvements or changes in model resolution. Given the theme of the ACP journal, the reader expects that this work should improve our understanding of the chemical and/or physical processes that control stratospheric aerosol evolution, but it remains unclear if there is any improvement in understanding being extracted from the study. Nor is there any real motivation or objectives stated in the introduction for the model simulations.
-> Response: The improvements are mostly due to use of more satellite data for volcanic SO2 using a novel method. This is mentioned now at several places including the abstract. Model improvements include consideration of aerosol effects in the photolysis rate calculation (minor effects here) and compared to Bingen et al. (2017) a finer horizontal resolution (see section 3).

Specific comments:

L11: “Reproduce” is too strong
-> Corrected: are consistent with

L12: Here it is said that “slight deviations … were found only for the large volcanic eruption of Pinatubo in 1991”, but later in the document deviations in other time periods, e.g., 2010 are discussed, so this is inconsistent.
-> Removed: ~~only~~

L19: precise language is needed here, is this the peak radiative forcing produced by a typical “small” eruption, or the time average forcing from these eruptions? And what is a small or medium eruption? Also, it’s not clear how this number is estimated, a value of 0.10 W/m^2 is not mentioned in the results or conclusions, and if it comes from Fig 11, how is the effect of small eruptions separated from that of “background” sulfur (e.g., DMS, OCS) transported into the stratosphere via atmospheric circulation?
-> Response: This number can be taken from Fig. 11 for volcanically quiescent periods or periods between medium size eruptions. Background is about 0.04 W/m^2 (not shown explicitly, taken from a simulation with much less volcanoes). This includes organics and dust.

L22-24: references needed for these statements.
-> Cited: Kloss et al., 2019 and Vernier et al., 2011

L25: I believe Bruehl et al., 2015 were not making the actual measurements of the size distribution of stratospheric aerosols. Better reference needed.
-> Cited: Wilson et al., 2008

L31: part of the aims stated here is apparently related to the interaction of aerosols with ozone, but this is not shown in the manuscript.
-> Response: text modified but this is not quantified in the text except for a number related to forcing.

L34ff: Reference(s) needed.
-> Cited: Vernier et al., 2011

L37: I am skeptical of a 3-year upper limit on the impact from volcanic eruptions: if ocean temperatures are a part of “climate”, then there is good evidence that volcanic impacts on climate can last much longer than 3 years (e.g., McGregor et al., 2015). Obviously the period of impact depends on many factors, but we should be careful to not overly simplify statements which might be misleading to some readers.
-> Corrected: " These changes influence in turn the radiative forcing at tropopause altitudes (or at the top of the atmosphere) for several years after the eruptions (Timmreck et al. 2012) and can even have a more prolonged impact on the global climate (McGregor et al. 2015)."

L44: Reference(s)?
-> Added citation: Solomon et al. 2011

L65: Some information should be given on how the SO2 column data was used, especially in regards to how a stratospheric component was estimated from the full column.
-> Response: In the case of data gaps for the four main satellite instruments used, the data from additional satellites e.g. TOMS, OMI or OMPS are applied to double check, if available. Especially in 2018 and 2019 the OSIRIS data are so sparse that constraints from instruments like OMPS or analogues events of previous years have to be superimposed for some eruptions.

L130: The gaps in spatial coverage of the OSIRIS data at 17 km extend significantly beyond the polar night: they seem to extend even in best cases to 20-30deg. Some rephrasing needed.
-> Response: OSIRIS provides a surface coverage from 82° S--82° N, except in polar winter when there is no sunlight and except in the Southern Hemisphere winter when tangent point is not illuminated by the sun

L136: It's not apparent how the sensitivity to clouds can be seen in Figure 4.
-> Response: At altitudes near and below the tropopause, the OSIRIS measurements are sensitive to clouds that may be interpreted as elevated aerosols. This is likely contributing to larger background extinction values measured below approximately 17 km in the tropics, as can be seen in Figure 4.

L140: How is the correction factor determined? This sounds suspiciously like numbers have been chosen only to produce best agreement.
-> 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 is relatively close in time, the factor has to be <1 to remove the influence of the previous event."

L157: The study of Grainger et al. (1995) does not seem to provide a relationship between SAD and SO2 mixing ratio. More explanation needed.
-> Response: This is skipped here. Now we write in Section 4: "For SAGE II in most cases the SO2 mixing ratio is derived using the parameterisation of Grainger et al. (1995) which converts SAD to volume density as a first step. We use the pressure and temperature provided to convert from mass density to a volume mixing ratio, assuming that observed sulfate is produced from injected SO2 some weeks ago. With this method it is easier to correct for cloud contamination than by using the extinction directly as above for the other instruments. Case studies for three events, comparing SO2 results from the different satellites and the different conversion methods are presented in Appendix B".

L190: It is not clear how differences in the "vertical transport of tracers, like dust and water vapor or ozone" between model resolutions has any importance to the present study.
-> Response: This is part of the general setting of the model simulations and has been moved to the appendix.

L216: What parameters?
-> Response: We removed "parameters" by: "aerosol optical properties like wavelength- dependent particle extinction cross section, single scattering albedo, and asymmetry parameter for each aerosol mode from AEROPT (Dietmüller et al. 2016) "

L218ff: The double radiation call most likely calculates the "instantaneous radiative forcing". It is important to be clear about this and consistent with the terminology.
-> Added: "instantaneous" radiative forcing

L219: There is a double radiation call, but how exactly is the radiative forcing calculated?
-> Correction: "instantaneous" radiative forcing, "..by taking the difference of the net total fluxes at 100 hPa or TOA."

L220: Not understanding this, are you diagnosing the impact of volcanic aerosol on upper stratospheric UV absorption? Nothing like this is shown in the results.
-> Response: The description of the RAD\FUBRAD sub-submodel is part of the general setting of the model simulations and has been moved to the appendix.

L241: What is the justification for the lower limits to the vertical integration given? You use 12 km as the lower limit in high latitudes, but the climatological tropopause height in high latitudes is 9-10 km. Conversely, you use 14 km in low latitudes, but the tropopause there is around 17 km. A thorough explanation for these counterintuitive thresholds will need to be given.
-> Response: 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 (e.g. effects of frontal systems). 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 in our extraction scheme we exclude cloud contaminated regions (see section 4 and 2).

L251: An “integration time” has not been introduced, it is not clear what this means in terms of the method.
-> Response: The temporal resolution of the satellite data is 5 days (for MIPAS, GOMOS and OSIRIS). The integration time is the case dependent time period used for the single eruptions. “The integration time (i.e the used time period)”

Table 2: There are a number of cases where the number of values do not match between the different columns in a particular row, e.g., 11 Feb 1990, 19 Aug 1992, 18 Sep 1996. Expanding the table so each eruption is listed in a single row would help this issue, as well as improve the machine readability of the table more generally. There is also a case (14 Jan 2002) where values are listed within brackets, and I did not find an explanation for what this means.
-> Response: These issues have been corrected, the table is available now in a Fortran formatted form and it has to be consistent to the 3D netcdf SO2 pertubation files provided in https://doi.org/10.26050/WDCC/SSIRC_3, which often contains multiple events. A single event list with more than 500 entries would not be acceptable for the ACP readership.

Table 2: The methods used produce an estimate of about 17 Tg for Pinatubo, which is in line with direct measurements of SO2 (e.g., Guo et al., 2004), but in contrast to recent model studies which suggest the effective injection for Pinatubo was much less (e.g., Mills et al., 2016; Dhomse et al., 2014). Some discussion of this issue would fit well into the paper.
-> Added in sec. 6.2:
“In this study about 17 Tg SO2 are injected for the Pinatubo eruption (Guo et al. 2004). Model comparisons by Timmreck et al. (2018) show that the span of used injections varies between 10 Tg SO2 (e. g. Dhomse et al. (2014); 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 the injected sulfur mass.
On the other hand, filling the gaps in the SAGE data just by horizontal linear interpolation increases the peak AOD by about a factor of 2, which is close to the GloSSAC compilation. In Figure 10 the AVHRR (Advanced Very High Resolution Radiometer) by Long and Stowe (1994) at 630 nm are included, which are close to our simulations (if converted to 750 nm their AOD would be slightly less or to 550 nm slightly larger).
When comparing the EMAC simulations (red line in Figure 9) with the simulation of Schmidt et al. (2018, fig.1) (black line in Figure 9, lower panel) it can be recognized that a smaller value for the peak of the Pinatubo eruption occurs, but here it needs to be considered that Schmidt et al. (2018) are using monthly global-means. This has the consequence that the signal of single eruptions is blurred and smaller sized eruptions cannot be easily identified.”

L269: Mixing ratios appear quite variable, what is meant here by “typical”?
-> Response: altitude range added in text: (“25 to 29 km”), this refers also MIPAS observations in that altitude (Höpfner et al. 2013).

L271: What upper limit is referred to here?
-> Added in text: “consistent to the Smithsonian reports, SAGE and TOMS”.

L281: References should be included to support this statement on the transport of aerosols from Nabro.
-> Done: Clarisse et al. 2014

L290: “The comparison of the simulated and observed SO2 values” is really hard to do since Figures 1 and 6 use different units and color schemes. It would be helpful to extract the MIPAS years from the simulations and show them with the same units and color scheme in comparison to the observations.
-> Response: figure 1 replotted using the same unit and color scheme as in figure 6.

L293: Is the statement on SO2 lifetimes made here a result of this study, or are the lifetimes equivalent to those given by Hoepfner et al. (2015)? If the result is the same as Hoepfner et al., (2015), that should be explicitly stated. If estimated lifetime are different from Hopefner et al. (2015), how and why?
-> Response: "Generally, the conversion of SO2 to sulfate aerosol particles depends on several factors, such as the altitude, latitude, or season of the eruption and takes according to Höpfner et al. (2015) about 13, 23

and 32 days in 10–14, 14–18 and 18–22 km altitude, respectively, in midlatitudes. 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)."

L300: This sentence seems to say that stratospheric aerosol optical properties were calculated using a range of different aerosol types (sulfate, dust etc.). Is this correct, or is the sentence just misleading?
-> Response: Yes, this is correct. See Dietmüller et al. (2016):
"Aerosol species explicitly considered are water soluble inorganic ions (WASO), black carbon (BC), organic carbon (OC), sea salt (SS), mineral dust (DU), and aerosol water (H2O). The refractive indices for those aerosol species are extracted from various data sources (most of the data are compiled in the HITRAN2004 database) and include wavelength dependencies. … The refractive indices for each aerosol mode required as input for the lookup tables are calculated assuming an internal mixture of the aerosol components for the hydrophilic modes. A mean refractive index is calculated for each mode wavelength combination by averaging the refractive indices of the individual components weighted with their volume contributions. The corresponding Mie size parameters are derived from the median radii of the log-normally distributed modes and the respective wavelengths. The wavelength- dependent particle extinction cross section, single scattering albedo, and asymmetry parameter for each mode are then obtained from the lookup table for the appropriate modal width (σ)."

L330: The OSIRIS data is converted from 750 nm to 550 nm, which is fine, but this contradicts the statement just a couple sentences earlier that "Unlike most other studies, the stratospheric AOD is compared at the original wavelengths derived from different optical channels of the satellite instrument measurements."
-> Response: The model calculates the AOD and other optical properties directly, derived from the original wavelengths of the satellite data. For additional comparisons only, the satellite data of OSIRIS were converted to 550 nm by the cited authors.

L333: The statement that "differences after the large Pinatubo eruption in 1991 between the model simulations and the SAGE II observations are related to the "saturation" effects of the satellite instrument" seems much too confidently worded. It seems quite possible that "saturation" effects explain some of the difference, but how certain can you be sure that it is the only, or even the primary reason? In the tropics, the simulated AOD appears to be ~3 times larger than the SAGE II measurements—is it likely that the SAGE II measurement is so strong an underestimate of the true total AOD?
-> Response: See AVHRR data now included. Filling the SAGE data gaps in the lower stratosphere just by linear interpolation in the first year after the eruption increases AOD by about a factor of 2.

Fig. 11: The ERBE measurements are not described at all in the text. Are they anomalies? What is the global coverage of the measurements? Likewise, the data from Solomon is only mentioned in passing in the text, and a little more detail should be included on how those radiative forcing estimates were calculated.
-> Response: The ERBE data shown in Fig. 11 are estimated averages for solar forcing at the top of the atmosphere, with 72-day means in the near-global (60◦ S–60◦ N) data set. Details are in Wong et al., (2006) and Toohey et al., (2011). (in caption of Fig.11)

L352: "The new model simulations with the additional volcanic eruptions (red line) are closer to the calculated estimates from satellite extinction measurements of SAGE, GOMOS and CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) by Solomon et al. (2011) (green crosses) than in previous studies (e.g., Brühl et al. (2015))." This statement, a concrete conclusion of the study, is impossible for the reader to verify without accessing the prior study, finding the relevant figure, and trying to visually compare the two. This is asking too much of the reader. Please include the result of Bruehl et al. (2015) directly on Fig 11 here so we can directly assess the validity of this statement.
-> Response: For direct comparison, the results of Brühl et al. (2015, 2018) (pink lines) are now included in Fig. 9-11.

L361: Are the results of Minnis et al. (1993) equivalent to the ERBE data shown in Fig 11? Please clarify.
-> Response: Minnis et al. (1993) derived their results from the ERBE data.

L362: clarify that the *simulated* AOD drops too quickly compared to the observations.
-> Corrected: “the simulated instantaneous global negative radiative forcing drops again too quickly for the EMAC simulations (red line) after the Pinatubo eruption compared to the observations.”

L374: “2019” is not an eruption.
-> Corrected: including the Raikoke eruption in 2019

L375ff: This paragraph is quoting results from other papers, not showing work from this study. If these statements are important, they should be moved out of the Results section or linked directly with results of the study.
-> Corrected: moved to introduction

L385: The fact that this study uses a higher resolution model than previous studies should have been mentioned earlier, in the model description and/or introduction.
-> Added in Sec. 3.1:
“For these model simulations, a higher horizontal resolution T63 (1.87°×1.87°), instead of T42 (2.81°×2.81°) in Bingen et al. (2017), was chosen.”

L386: This appears to be a result of the study by Bruehl et al. (2018), which would be important in describing the experiment earlier in the manuscript but not here in the conclusions.
-> Response: Skipped here (see above).

L388: The SAGE II and OSIRIS extinction measurements are not really “newly available”, some version of this data has been available for many years. The estimation of SO2 from these data sets is quite new—it’s what this paper is presenting!
-> Response: The resolution of the updated versions is improved, so OSIRIS data has allowed comparisons up to be extended to 2019; additional SAGE II data was also used extend the comparison back to 1990 (together with the Smithsonian database).

L402ff: This conclusion is not supported by the results: there is no quantification of the impact the increased number of eruptions included in the database has on the radiative forcing, or its level of agreement with observations.
-> Removed.

L408ff: This is an interesting conclusion, but it is not supported by the results. There is no demonstration that including the injections below the tropical tropopause improves the agreement. Even a comparison with prior studies will not prove necessarily support the statement since those prior studies used a different resolution model.
-> Response: Remarks on that are added at several places in other sections.

L418: This is not a new result, as it has been shown by prior studies.
-> Response: Paragraph skipped. Instead we included now in the paragraph beginning with:
“Our volcanic sulfur emission inventory….” “The inclusion of plenty of small size eruptions reaching the UTLS has the consequence that stratospheric aerosol optical depth and radiative forcing does not decrease to almost zero between medium size eruptions in agreement with observations, in contrast to a lot of other studies.”

L422: The impact of volcanic aerosol on tropical upwelling is not diagnosed in this study. Prior studies have explored this, but statements like this can not be included in the conclusions of this work if there are no new results shown to support it and build upon prior work.
-> Moved to introduction.

L437ff: This paragraph talks about meteoritic dust, which was not investigated in the study. Perhaps simply adding a sentence or two on the agreement between the model and observed aerosol extinction in the upper stratosphere to motivate the discussion of meteoritic dust would help the reader follow the logic here.

-> Removed: Remark added to sec.6.1: “Above about 24 km altitude, EMAC underestimates the observations because in the model meteoric dust particles were not considered.”

448: Confirming the findings of the IPCC report is, firstly, incorrectly phrased, since the IPCC report only summarizes and reports findings gathered from the published literature. It would be more important to compare the results here with the primary sources, including studies that have been published since the IPCC AR5 (e.g., Schmidt et al., 2018). Second, confirming some general results from prior studies does not make an overwhelming case for publication. What does this study add to the understanding of volcanic radiative forcing that wasn’t known before?
-> Response: Paragraph shortened, restricted on new results.

L450: Radiative forcing is stated to be that at the surface here, where Fig 11 is said to be RF at the tropopause. Also the numbers quoted here don’t seem to agree with Fig 11. It would be best to only refer to calculations for which the results are shown in the paper.
-> Response: Typo corrected, numbers refer to difference to quiescent periods which is now mentioned.

Editorial comments:

Line 9: Volcanic SO2 is not “pollution” in the usual sense of the word, suggest it be cut here.
-> Corrected.

L49: “Distribution”?
-> Corrected.

L53: “constitute a source of background…”
->Corrected.

L55: Awkwardly phrased: the processes aren’t structured, the paper is, and not strictly according to processes.
-> Corrected: "This paper is structured as follows:"

L80: I’ve never seen pptv written with v as a subscript, is this a new standard?
-> Corrected.

L111: confusingly phrased.
-> Corrected concerning cited Figures. We have changed the sentence in:
“Figures 2 and 3 show the aerosol extinction from the GOMOS instrument at wavelengths of 550 nm (Figure 2) and 750 nm (Figure 3), respectively. In both cases, a gridded aerosol extinction dataset is used (CCI-GOMOS dataset in version 3.00, see Bingen et al., 2017).”

Added References:

Dhomse, S. S., Emmerson, K. M., Mann, G. W., Bellouin, N., Carslaw, K. S., Chipperfield, M. P., Hommel, R., Abraham, N. L., Telford, P., Braesicke, P., Dalvi, M., Johnson, C. E., O’Connor, F., Morgenstern, O., Pyle, J. A., Deshler, T., Zawodny, J. M. and Thomason, L. W.: Aerosol microphysics simulations of the Mt.~Pinatubo eruption with the UM-UKCA composition-climate model, Atmos. Chem. Phys., 14(20), 11221–11246, doi:10.5194/acp-14-11221-2014, 2014.

Guo, S., Bluth, G. J. S., Rose, W. I., Watson, I. M. and Prata, A. J.: Re-evaluation of SO 2 release of the 15 June 1991 Pinatubo eruption using ultraviolet and infrared satellite sensors, Geochemistry Geophys. Geosystems, 5(4), Q04001, doi:10.1029/2003GC000654, 2004.

McGregor, H. V., Evans, M. N., Goosse, H., Leduc, G., Martrat, B., Addison, J. A., Mortyn, P. G., Oppo, D. W., Seidenkrantz, M.-S., Sicre, M.-A., Phipps, S. J., Selvaraj, K., Thirumalai, K., Filipsson, H. L. and Ersek, V.: Robust

global ocean cooling trend for the preindustrial Common Era, Nat. Geosci., 8(9), 671–677, doi:10.1038/ngeo2510, 2015.

Mills, M. J., Schmidt, A., Easter, R., Solomon, S., Kinnison, D. E., Ghan, S. J., Neely, R. R., Marsh, D. R., Conley, A., Bardeen, C. G. and Gettelman, A.: Global volcanic aerosol properties derived from emissions, 1990-2014, using CESM1(WACCM), J. Geophys. Res. Atmos., 121(5), 2332–2348, doi:10.1002/2015JD024290, 2016.

Schmidt, A., Mills, M. J., Ghan, S., Gregory, J. M., Allan, R. P., Andrews, T., Bardeen, C. G., Conley, A., Forster, P. M., Gettelman, A., Portmann, R. W., Solomon, S. and Toon, O. B.: Volcanic Radiative Forcing From 1979 to 2015, J. Geophys. Res. Atmos., 123(22), 12,491-12,508, doi:10.1029/2018JD028776, 2018.

(more in the acpd-preprint)

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.

Comment on acp-2021-654

Anonymous Referee #3

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-RC3, 2021

Review of “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.

Using various (occultation and limb based) satellite instruments, with vertical SO2profiles from different satellite instruments and chemistry climate simulations, this study characterizes the influence of stratospheric volcanic aerosols for the period between 1990 and 2019. The results show that small but relatively frequently eruptions contribute to the stratospheric aerosol layer and could cause a global radiative forcing in the order of−0.1 Wm−2 at the tropopause. In specific, the objective of this study was to generate a detailed volcanic sulfur emission inventory, to improve the EMAC model simulations of the global stratospheric aerosol and sulfate burden, and to compute the volcano-induced radiative forcing through validation with satellite data.

Honestly, the paper keeps me a bit loss, as I am not sure if it is a more scientifically or more technically oriented paper. The scientific objective is not clear to me in particular the added value to the recent literature. I am wondering if the paper would not better fit in Earth System Science data (ESSD, https://www.earth-system-science-data.net/) or in Geoscientific Model Development (GMD, https://www.geoscientific-model-development.net/). The topic of the paper is in general very suitable for ACP but the paper needs major substantial revisions before publishing in ACP, see my major comments below.

Major comments:

The introduction needs a complete rewriting, less text book more scientific background with respect to the questions to be addressed. The paper is a successor of Brühl et al. (2015; 2018) and Bingen et al. (2017) but I miss a clear separation and explanation about the added values of this paper compared to its predecessors. The better horizontal resolution has already been discussed in Brühl et al. (2018), so the new aspect, as far as I understood it, is the increased amount of volcanic eruptions and the extend time period by using new satellite data.
-> Response: Abstract and introduction are expended to address this. The resolution is now only mentioned in the model description section

I completely miss references to recent literature in the introduction with respect to radiative forcing estimates of recent eruptions. There are several publications e.g. Andersson et al., (2015); Friberg et al., (2018), Schmidt et al., (2018); Kloss et al;(2021) just to name a few which have addressed the radiative forcing of small to moderate volcanic eruptions in the recent years. These papers have to be cited and differences/added values to their work have to be addressed in the introduction.
-> Response: Most included in introduction now, see replies to other referees and “Friberg et al. (2018) included the whole time series of CALIOP data from 2006 to 2015 and derived stratospheric AOD using reanalysis data for the tropopause, but mentions only medium size eruptions explicitly. Radiative forcing is estimated there from multiplying AOD with -25, an approach which is valid only for purely scattering aerosol”. Kloss et al. (2021) cited in sections 6.2 (“Our northern hemisphere results for AOD of about 0.025 for Raikoke (550nm) agree within uncertainties with Kloss et al. (2021) who use different satellites and different modelling approaches”) and 6.3 (“The value for Raikoke/Ulawun is within the range discussed in Kloss et al. (2021)”) now. This paper was not available when we wrote the first version, thank you.

The discussion needs also to be rewritten. As mentioned above the lack of references of recent literature is astonishing. The results of the study need to be discussed in the context of recent literature, e.g. what do we learn from this paper, what we didn’t know before from previous studies.

-> Response: Done, see comments to other referees.

I am also wondering about the importance of the small eruptions for the global radiative forcing. It would be interesting if you neglect all small eruptions below a certain threshold values e.g. 10 kT SO2, how this would really change the global radiative forcing. What is range of uncertainties, the range of interannual variability in background periods? Estimates about the uncertainty range are completely missing in the paper.
-> Response: See comparison with results of Schmidt et al. (2018) in Fig 11. They neglect also larger eruptions than your threshold, however. See also remark to referee 1 on background.

Last but not least, differences between the model simulations and satellite measurement need not to be the only cause of missing SO2 sources. There could be several other reasons for the discrepancies (transport, microphysics), neither model simulations nor satellite measurements are perfect. This has to be discussed here as well.
-> Response: Mentioned at several places.

Specific comments

Abstract, line 17: "significantly" is a big word. I did not find any significance tests in the paper.
-> Removed: "significantly"

Page 3, which SSTs do you use? I suppose you run only one ensemble members did you check for the influence of internal variability at least in short sensitivity studies?
-> Response: The CCM is nudged to ERAI which includes SSTs (see section 3).

Description of the EMAC module could be reduced, to only the parts which are really relevant for the paper,. e.g. the calculation of the radiative forcing. This part could be more elaborated. More detailed model descriptions can be put in the appendix.
-> Response: Parts of the model description are moved to the appendix:
"As EMAC is a very complex chemistry climate model it contains many submodels and functions which are essential for running the simulations but are not directly related to the sulfur cycle, these are mentioned in Appendix A. In this section we focus on the sulfur cycle."

Page 12, lines 245-247 It would be nice to see a comparison with Carn et al (2017) and other recent emission data
-> Response: You compare then apples and oranges, a hint is given in Appendix C.

Table 2: It would be nice to see (e.g. with different color) which entries are new or changed with respect to the previous data set.
-> Response: Eruptions from in Bingen et al. (2017) and Brühl et al. (2018) are marked in italics in the table: Based on a previous study from Brühl et al. (2018) with scaling factors for T63 and already published in an earlier version in Bingen et al. (2017) *(in italics).*

Will the data set be published?
-> Response: Yes, the data set is published on WDCC: https://doi.org/10.26050/WDCC/SSIRC_3.

Page 21, line 279 "strong" I wouldn't call Kasatochi or Raikoke a strong eruption
-> Corrected: "medium strong"

Figures 9, 10, 11: A comparison with Brühl et al. (2015) for the Pinatubo period and with Brühl et al (2018) for 2002 to 2012 would be nice, to better asses the improvements of this study. Also a validation with GloSSAC (Thomason et al., 2018; Kovilakam et al., 2020) would more than beneficial.
-> Response: Pink lines for comparison with Brühl et al. 2015 are added in fig. 9-11. GloSSAC gives no additional information here since it is derived from data shown in the figure but we can include it in the upper panel

since it covers a longer time period than the blue line. Nevertheless, we include a black line in Fig 9 since it is interesting for Pinatubo and the period 2012 to 2018 (based on V2 of Kovilakam).

Section 6.3: Any reason why you look at the tropopause? What is the uncertainty range in your forcing estimates?
-> Response: This is because of the comparison with Solomon et al 2011. For Pinatubo it differs not much from the value at TOA.

Figure 11 I recommend a comparison with Schmidt et al (2018) here
-> Response: 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 black line.

Page 409, 410: "This was demonstrated to be essential for correctly assessing the extinction coefficient in volcanically quiescent periods." By whom? Maybe I have overseen it but I didn't find it in the paper.
-> Reformulated: "This was demonstrated to be important for correctly modelling the AOD in volcanically quiescent periods". Convection was mentioned in earlier sections.

Page 445, Which studies?
-> Check given reference and Brühl et al. (2018).

Added References:

Andersson, S. M., Martinsson, B. G., Vernier, J. P., Friberg, J., Brenninkmeijer, C. A. M., Hermann, M., Van Velthoven, P. F. J., and Zahn, A.: Significant radiative impact of volcanic aerosol in the lowermost stratosphere, Nat. Commun., 6, 1–8,https://doi.org/10.1038/ncomms8692, 2015.

Bingen, C., Robert, C. E., Stebel, K., Brühl, C., Schallock, J., Vanhellemont, F., Mateshvili, N., Höpfner, M., Trickl, T., Barnes, J. E., Jumelet, J., Vernier, J.-P., Popp, T., de Leeuw,G., and Pinnock, S.: Stratospheric aerosol data records for the climate change initiative: Development, validation and application to chemistry-climate modelling, Remote Sensingof Environment, 203, 296–321,https://doi.org/10.1016/j.rse.2017.06.002, 2017.

Brühl, C., Lelieveld, J., Tost, H., Höpfner, M., and Glatthor, N.: Stratospheric sulphur and its implications for radiative forcing simulated by the chemistry climate model EMAC, J. Geophys. Res.-Atmos. 120, 2103–2118, https://doi.org/10.1002/2014JD022430, 2015.

Brühl, C., Schallock, J., Klingmüller, K., Robert, C., Bingen, C., Clarisse, L., Heckel, A., and North, P.: Stratospheric aerosol radiative forcing simulated by the chemistry climate model EMAC using aerosol CCI satellite data, Atmospheric Chemistry and Physics, 18,1–15, https://doi.org/https://doi.org/10.5194/acp-18-12845-2018, 2018.

Friberg, J., Martinsson, B. G., Andersson, S. M., and Sandvik, O. S.: Volcanic impact on the climate – the stratospheric aerosol load in the period 2006–2015, Atmos. Chem. Phys., 18, 11149–11169, https://doi.org/10.5194/acp-18-11149-2018, 2018

Kloss, C., Berthet, G., Sellitto, P., Ploeger, F., Taha, G., Tidiga, M., Eremenko, M., Bossolasco, A., Jégou, F., Renard, J.-B., and Legras, B.: Stratospheric aerosol layer perturbation caused by the 2019 Raikoke and Ulawun eruptions and their radiative forcing, Atmos. Chem. Phys., 21, 535–560, https://doi.org/10.5194/acp-21-535-2021, 2021.

Kovilakam, M., Thomason, L. W., Ernest, N., Rieger, L., Bourassa, A., and Millán, L.: The Global Space-based Stratospheric Aerosol Climatology (version 2.0): 1979–2018, Earth Syst. Sci. Data, 12, 2607–2634, https://doi.org/10.5194/essd-12-2607-2020, 2020.

Schmidt, A., Mills, M. J., Ghan, S., Gregory, J. M., Allan, R. P., Andrews, T., Bardeen, C.G., Conley, A., Forster, P. M., Gettelman, A., Portmann, R. W., Solomon, S., and Toon, O.B.: Volcanic Radiative Forcing From 1979 to 2015, J. Geophys. Res.-Atmos.,123,12491–12508, https://doi.org/10.1029/2018jd028776, 2018.

Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., and Peter, T.: A global space-based stratospheric aerosol climatology: 1979–2016, Earth Syst. Sci. Data, 10, 469–492,https://doi.org/10.5194/essd-10-469-2018, 2018.