We thank Anonymous Referee #2 for the helpful suggestions and comments. Our point by point answers to the comments are presented below. Referee comments are in bold and our replies in body text.

Referee #2 comments:

The main conclusion of the paper is that the impacts of a volcanic eruption would be "significantly different depending on whether the eruption occurs during SRM deployment or into a clean background stratosphere" (pg 21857, lines 17-19). In fact, the results could be used to argue the opposite.

In Fig 2, the peak sulfate burden reached after a volcanic eruption during continuous SRM ("SRM Cont") is identical to the sum of the individual results of SRM alone and an eruption alone. (There are some differences in burden between months 10 and 20, but these seem small compared to the peak burden, and the significance of this difference is impossible to judge from single ensemble members). Similarly, global mean temperature and precipitation responses shown in Fig 4 are consistent (within error bars) for the "Volc" and "SRM Cont" experiments, which implies that the background state has little-to-no impact on the climate response to a volcanic eruption

The discrepancy between the author's conclusions and those above originates from the definition of what exactly "during SRM deployment" means. The paper puts most of its focus on simulations of a scenario where SRM is employed before the eruption, but suspended upon the occurrence of the eruption. (The authors argue that this the most likely scenario, but one could counterargue that the politics of SRM deployment would be so complicated that it is near impossible to state with any confidence what might be most likely.) Since the difference between the results of the "Volc" and "SRM Cont" simulations are small (as argued above), the conclusions described in the abstract and conclusion section are not due to the existence of SRM before the eruption, but to the sudden suspension of SRM at the time of the eruption. This (as the authors acknowledge within the manuscript) is a trivial result, and very specific to the scenario constructed. Unfortunately, this is not well communicated in the text of the abstract nor the conclusions, and I think many readers could understand that the study finds a significant decrease in the volcanic response under continuous SRM. A readers confusion might be justified, since a comparison of the simulations of volcanic eruptions under background and continuous SRM is the more natural experiment, where only one parameter is being changed between the simulations. The authors' choice to fo- cus primarily on the SRM scenario with a sudden stop therefore increases the chance for misinterpretation and seems to oversell the role of aerosol size in the radiative and climate impacts of volcanic eruptions, at least for the scenarios explored here.

While we would still argue that suspending SRM following a massive eruption is the most likely scenario (both because of safety and financial considerations), we admit that it is impossible to say for certain whether SRM would be suspended completely, immediately after the eruption or later (if at all). Here we have studied two boundary cases, one where injections are suspended completely immediately after the eruption and another where injections are continued despite the eruption. A "real-life" scenario could be either one of these, or a combination of these two (i.e. SRM is suspended at some point after the eruption or the amount of injected sulfur is decreased). Therefore, the reviewer‘s suggestions on reformulating parts of the abstract and conclusions have been taken into account.

The abstract now reads:

" -- *According to our simulations the radiative impacts of the eruption and SRM are not additive and the radiative and climate impacts of the eruption depend strongly on whether SRM is continued or* *suspended after the eruption. In the former case, the peak burden of the additional stratospheric sulfate as well as changes in global mean precipitation are fairly similar regardless of whether the eruption takes place in a SRM or non-SRM world. However, the maximum increase in the global forcing is approximately 30% lower compared to a case when the eruption occurs in an unperturbed atmosphere. In addition, the recovery of the stratospheric sulfur burden and forcing is significantly faster in the concurrent case because the eruption during the SRM leads to a smaller number of larger sulfate particles compared to the eruption in a non-SRM world. On the other hand, if SRM is suspended immediately after the eruption, the peak increase in global forcing is about 40% lower compared to a corresponding eruption into a clean background atmosphere. In addition, the recovery of the stratospheric sulfur burden and forcing is significantly faster in the concurrent case. In this simulation, a volcanic eruption leads to only about 1/3 of the peak global ensemble-mean cooling compared to an eruption under unperturbed atmospheric conditions. Furthermore, the global cooling signal is seen only for the 12 months after the eruption in the former scenario compared to over 40 months in the latter. In terms of global precipitation rate, we obtain a 36% smaller decrease in the first year after the eruption and again a clearly faster recovery in the concurrent eruption and SRM scenario, which is suspended after the eruption. We also found that an explosive eruption could lead to significantly different regional climate responses depending on whether it takes place during geoengineering or into an unperturbed background atmosphere. Our results imply that observations from previous large eruptions, such as Mt Pinatubo in 1991, are not directly applicable when estimating the potential consequences of a volcanic eruption during stratospheric geoengineering.“*

The relevant parts of the results now read:

“Figure 4a also shows that on average a volcanic eruption during continued SRM *(simulation SRM Cont, red line) leads to on average 33% smaller cooling for next three years after the eruption than under unperturbed atmospheric conditions. If SRM is suspended (SRM Volc), the maximum value of the global cooling is only about 1/3 (i.e. less than 0.14 K at maximum for the ensemble mean) compared to an eruption to the non-geoengineered background stratosphere (simulation Volc). This is consistent with the clearly smaller radiative forcings.”*

The relevant parts of the conclusions now read:

"According to our simulations, the impacts of a volcanic eruption during SRM depend strongly on whether SRM is continued or halted after the eruption. In the former case, the peak additional forcing is about 30% lower and the global cooling 33% smaller than compared to an eruption taking place in non-SRM world. However, the peak additional burden and changes in global mean precipitation are fairly similar regardless of whether the eruption takes place in a SRM or non-SRM world. On the other hand, if SRM is stopped immediately after the eruption, the peak burden is 24% and forcing 40% lower and reached earlier compared to the case with unperturbed atmosphere. Furthermore, the forcing from the eruption declines significantly faster, implying that if SRM was stopped after the eruption, it would need to be restarted relatively soon (in our scenario within 10 months) after the eruption to maintain the pre-eruption forcing level."

The paper describes differences in regional precipitation patterns between the different scenarios. The significance of the precipitation anomalies in the three volcanic scenarios is weak outside of the tropics, due likely to the high natural variability and the relatively short runs and few ensemble members. Therefore, the conclusion that precipitation anomalies after an eruption during SRM would be different than during background conditions is not well supported by the results, since it is not shown whether the differences between the scenarios is significant.

We agree and have already explicitly stated the lack of statistical significance both in the text and in the figures (the hatching indicates regions with statistically significant changes). However, we decided to show also the regional precipitation changes, since they are a much discussed aspect of geoengineering. The uncertainty in regional precipitation is not caused by our specific simulation set-ups, but related to natural variability in precipitation and generally models' capability to simulate changes in precipitation. However, the text has now been modified to recognize these uncertainties even more explicitly.

Furthermore, a major caveat is that if SRM is ever employed, it will be under elevated CO2 concentrations, which are not accounted for here. The combination of SRM and CO2 forcing would lead to different total temperature and precipitation anomalies than shown in panels (a) of Figs 5 and 6, therefore the impact of volcanic forcing (e.g., on the meridional temperature gradient) would likely be different in the real world case (with SRM and CO2 forcing) than in these simulations with only SRM.

We agree that SRM without warming from CO2 is an unrealistic scenario. Here we want to study and compare impacts from SRM and/or volcanic eruption in a simplified set-up. However, we have now highlighted the point raised by the reviewer in the text:

"However, it should be noted that here we have studied an unrealistic scenario where SRM is implemented without global warming. If warming from increased greenhouse gases had been included in the scenarios, the temperature gradient could be very different in simulation SRM which could lead to different precipitation patterns. There is also a large natural variability in the precipitation rates and as the precipitation changes after the eruption are

relatively small, our results are statistically significant only in a relatively small area (hatching in Fig. 6). "

Appendix A, which validates the model results compared to observations of the Pinatubo eruption, serves a valuable purpose in regards to the paper and is justifiably included as an appendix to sharpen the focus of the main text. If anything, Appendix A would benefit from more material, a comparison of simulated and observed AOD seems to be missing. On the other hand, Appendix B contains results which seem to be outside the scope of the rest of the paper, and only briefly described and explained with little more than speculative reasoning. Either these results should be incorporated into the main results of the paper and more thoroughly explained, or removed from the manuscript.

While a more detailed comparison of the model to observations in Appendix A could be interesting, it is out of the scope of this study. For ECHAM-HAM a more detailed comparison has already made (Niemeier et al 2009, Toohey et al 2011). However this would require changes in mode widths which is used to represent size distribution. After these changes, modelling tropospheric aerosols is not feasible anymore. Here we have used different microphysical aerosol model (M7 -> SALSA) which allows us simulate both stratospheric and tropospheric emissions with basic model setup. Compared to the previous comparisons, using SALSA instead of M7 affect mainly to particle size which is evaluated by comparing effective radius and global sulfate burden (which basically shows lifetime of particle which is related to the particle size). We think that these would be relevant scenarios to study here but the results are not that surprising or significant that those should be included in the main results.

Regarding Appendix B, we feel that it is important to remind the reader that our conclusions are to some extent dependent on the chosen scenario, and that for example different eruption location or season could lead to different results. In the first drafts of this manuscript the appendix material was incorporated within the main text, but several of the co-authors found the main storyline easier to follow once it was moved to an appendix. We have therefore left Appendix B as it was.

All minor comments have been fixed if not otherwise said.

P21838 L10: "decay" might be a better word choice than "recovery"

P21839 L23: "very likely" seems a strong statement given uncertainties in eruption return rates and the term of (hypothetical) SRM. For instance, between Katmai and Pinatubo passed almost 80 years. So it seems one could easily do SRM for 30-50 years and not have a major eruption.

"very likely" changed to "possible"

P21840 L8,11: no 's on Max Planck Institute

P21840 L10: The Niemeier and Stier references don't pertain to HAM-SALSA specifically.

P21841 L10: There is a stratospheric version of HAM. It might be explained why you have not used this version.

We are using sectional microphysical aerosol model SALSA instead of modal model M7. Using a stratospheric model of M7 basically mean using narrower mode width for coarse particles. In SALSA basic configuration is suitable for simulate stratospheric aerosols.

Sec 2.2: It is not clear in the experimental description (esp. for the SALSA simulations) how many ensemble members have been run, and whether the results shown are ensemble means or single ensemble members.

In this section we just want to describe scenarios and these should be included in previous section.

We added to section 2.1.1: "*Only one MAECHAM5-HAM-SALSA simulation has been performed for each of the studied scenarios to obtain the aerosol optical fields for the ESM simulations. Only for Volc we have carried out a five member ensemble to address potential forcing uncertainties (Appendix A).*"

P21845 L14: How well constrained is the 8.5 Tg S for Pinatubo?

This value is based on TOMS/TOVS and MLS satellite observations (Guo et al.,2004b; Read et al., 1993), and is the same as was used in previous studies with ECHAM-HAM (Toohey et al 2011 and Niemeier et al 2009).

P21846 L24: "yr-1" should be removed?**

P21847 L14: The SW forcing seems to oscillate a fair degree, so it might be wrong to call it "steady state"at least, you might clarify that the -6.22 Wm-2 is an average value.**

"Steady state forcing" is changed to *"average global mean forcing"*

P21848 L14: Here we learn that Fig 2 shows a single ensemble member. Why not take the ensemble mean (and show the variability of the ensemble with error bars)?

Since running ECHAM-HAM-SALSA (i.e. explicit aerosol microphysics model within a full atmospheric model) is computationally heavy and since there was not that large a variation in global mean values between single ensemble members, we decided not to run multimember ensembles for all the scenarios.

P21849 L1-5,21-30: The significance of the difference between the "SRM Cont" and the sum of the "Volc" and "SRM" experiments is not convincing, with only single ensemble members shown and no error bars, see major comments.

We now mention in the abstract and conclusions that the differences between these two cases are mostly small.

P21851 L24: Another example of a statement which could be easily misinterpreted, since the SRM scenario used here includes a suspension at the time of the eruption, which is not readily clear from the immediate context.

This is now changed: *"Figure 4a also shows that on average a volcanic eruption during continued SRM (simulation SRM Cont, red line) leads to on average 33% smaller cooling for next three years after the eruption than under unperturbed atmospheric conditions. If SRM is suspended (SRM Volc), the maximum value of the global cooling is only about 1/3 (i.e. less than 0.14 K at maximum for the ensemble mean) compared to an eruption to the non-geoengineered background stratosphere (simulation Volc). This is consistent with the clearly smaller radiative forcings."*

P21853 L13: Driscoll et al., 2012 shows that models don't produce the winter warming

pattern.

This is correct, CMIP5 models have problems to reproduce the winter warming. However some of the CMIP5 models produce a warming in winter, but much weaker than observed. This sentence is now rewritten to:

"Winter warming after a volcanic eruption has been seen also in observations (e.g. Robock and Mao 1992, Fischer et al., 2007)), though the current generation of CMIP5 models has problems to reproduce thee NH postvolcanic winter warming pattern (Driscoll et al. 2012) "

P21855 L16: This idea has been around for much longer than since 2013, see e.g., Bala, G., P. B. Duffy, and K. E. Taylor (2008), Impact of geoengineering schemes on the global hydrological cycle., Proc. Natl. Acad. Sci. U. S. A., 105(22), 7664–9, doi:10.1073/pnas.0711648105.

Added

P21856 L21: In the SE Pacific yes, but it's not clear if this opposite response holds for "most of the tropics".

This has been specified to concern Pacific and Atlantic.

P21858 L28: This explanation for the widening of the ITCZ in the simulations seems inconsistent with the understanding of the observed widening of the tropical belt being related to a decrease in the meridional temperature gradient. See Adam, O., T. Schneider, and N. Harnik (2014), Role of Changes in Mean Temperatures versus Temperature Gradients in the Recent Widening of the Hadley Circulation, J. Clim., 27(19), 7450–7461, doi:10.1175/JCLI-D-14-00140.1

Our results are not contradictory to Adam et al. Here the temperature gradient increases in SRM which would lead to the narrowing of the Hadley cell. However making conclusions about ITCZ would require a larger ensemble for further detailed investigations. Based on this, we have now removed the text concerning ITCZ and added discussion about different temperature pattern within the Tropics, which would explain the different precipitation patterns between the scenarios.

Text is now rewritten as follows:

““Similar to the temperature change, our simulations indicate that a tropical volcanic eruption impacts precipitation patterns differently in unperturbed and SRM conditions. In fact, a volcanic eruption during geoengineering (SRM Volc and SRM Cont) leads to an opposite precipitation change pattern than an eruption to the unperturbed atmosphere (Volc) over the tropic area in Pacific and Atlantic (Fig. 6c and 6d). In these areas, a volcanic eruption during SRM leads to the increase in the evaporation flux at the surface during the first year after the eruption, whereas the evaporation flux decreases if the eruption takes place in unperturbed conditions. This is caused by different tropical temperature responses between the simulations (Fig. 5). Compared to the pre-eruption values, in simulations SRM and Volc, equatorial SST anomalies (latitudes 0 N - 10 N) are relatively colder than the SST anomalies over latitudes 10 N - 20 N. In simulation SRM, the difference in SST anomaly between these areas is -0.02 K and in simulation Volc, it is -0.05 K. On the other hand, in simulations SRM Volc and SRM Conc, equatorial SST anomalies are relatively warmer than those over latitudes 10 N - 20 N. In SRM Volc, the difference in temperature anomaly is 0.13 K and in SRM Cont it is 0.05 K. However, these changes in precipitation are not significant and a larger ensemble would be necessary for further detailed investigations.“