Authors' reply to reviewers' comments on 'Multi-model comparison of the volcanic sulfate deposition from the 1815 eruption of Mt. Tambora' by Marshall et al.

We thank the Reviewers for their detailed and constructive assessment of our manuscript. Their comments and suggestions have greatly helped to improve our paper. Our reply begins with some general changes we have made to the paper, followed by our point-by-point replies. Reviewer comments are in italics and coloured dark grey.

General changes

In addition to the reviewer comments, upon revision, the authors have also made the following corrections (see also tracked changed document for minor corrections):

A spin-up issue was discovered in the SOCOL-AER runs and these have since been repeated. In the new runs the carbonyl sulphide (OCS) level was also set to 337 pptv as opposed to ~500 pptv (Table 1). As a consequence, all figures in the manuscript have been updated. Overall the main results in comparison to the other models have not changed, but sulfate deposition over the ice sheets has increased for both background and Tambora deposition. The text has been updated accordingly.

In the revised manuscript we have used the full model names instead of abbreviations to avoid confusion with the use of sub-models as names.

We have rephrased the text under Table 2 regarding the injection details for each model. Despite a uniform injection in CESM1(WACCM) and UM-UKCA, both these injections were divided amongst several grid boxes and each model grid did this differently. As such we thought it was misleading to state that for CESM1(WACCM) the overlap of model levels resulted in the emission fluxes peaking in the centre of the plume and to say nothing of UM-UKCA, where similar effects may be occurring. Instead we have added the sentence: "as the models are not on regular grids and their vertical resolutions differ, the distribution of the emission over the model grid boxes cannot be exactly the same. As a result, the injection profiles differed slightly between the models"

We have also added an additional author, James Pope, who was invaluable in setting up the UM-UKCA runs. We apologise for the oversight in not including him.

Point-by-point reply

Reviewer 1: C. Gao

This work is an important contribution to intermodel comparison and evaluation of volcanic simulation. The model derived relationship between volcanic sulfur injection and ice core volcanic aerosol deposition, if the results converge, can provide critical information to verify or improve the ice-core-based reconstruction of past volcanism. The paper is well written; the results are clearly presented and discussed. I would like to recommend the paper to be published in this journal after addressing the following points:

We thank Chaochao Gao for the helpful comments and have addressed these below.

1. In section 2 "Model set-up and ice core data". Are the four models the only aerosol models available for Tambora simulation? If that's the case, please state; if not, please explain if there is any criteria taken to choose the models.

There were no selection criteria as the four models are the only aerosol models available that were able to simulate deposition. A fifth model (CAMB-UPMC-M2D) was included in the initial Tambora experiment as part of Zanchettin et al. (2016), but the model did not simulate deposition and was therefore excluded from the analysis in our paper. We have added this statement to Sect. 2 and have also updated the abstract and introduction to specify that five models simulated the eruption as part of the VolMIP pre-study.

Also, please briefly describe model performance in previous studies.

We have added the following paragraph to describe model performance when simulating the eruption of Mt. Pinatubo in 1991:

"All four models simulate the 1991 eruption of Mt. Pinatubo in reasonable agreement with observations of the sulfate burden, aerosol optical depth and stratospheric heating (Niemeier et

al., 2009; Toohey et al., 2011; Dhomse et al., 2014; Sheng et al., 2015b; Mills et al., 2016), giving confidence in the models' overall abilities to accurately simulate the atmospheric and climatic effects of a large-magnitude eruption. However, the models vary in the details regarding the model-observation comparisons. For example, MAECHAM5-HAM (Niemeier et al., 2009) and SOCOL-AER (Sheng et al., 2015b) simulated a too rapid aerosol decay and UM-UKCA (Dhomse et al., 2014) had a low bias in the model-simulated aerosol effective radius compared to observations. Possible reasons for these differences include omitted or under-represented influences from meteoric particles, too large sedimentation and cross-tropopause transport and too fast transport from tropics to high latitudes. Conversely, the models differ in the amount of emitted SO_2 required to achieve good comparisons to observations with the mass of SO_2 emitted by the four models ranging from 10 Tg for UM-UKCA (Dhomse et al., 2014) and CESM1(WACCM) (Mills et al., 2016; 2017) to 12-14 Tg for SOCOL-AER (Sheng et al., 2015b) to 17 Tg for MAECHAM5-HAM (Niemeier et al., 2009; Toohey et al., 2011). For this reason, the use of a common protocol in this study (Sect. 2.2) enables us to better attribute potential differences in the results to model processes rather than to the eruption source parameters."

2. In section 3 "Results"

2.1 The argument of "models have similar background sulfate deposition patterns" sounds weak to me, please provide more quantitative evidence/analysis to support the argument. The color scale in Figure 1 and few other figures seems misleading, please consider using a monotone color scale.

We included the sentence as a statement with the aim of highlighting what can be seen in Figure 1; that the patterns of deposition are similar across the models as described in the opening sentences, despite differences in the magnitude. To make things clearer we have referred to Figure 1. To show quantitatively that the background deposition is similar at the poles we have also moved Figure S4, which shows the ice core fluxes versus the model-simulated values, from the supplementary material to the main text. This is now Figure 3 (please also see reply to point 2.3).

We have considered alternative colour scales but find that the use of this diverging scale highlights the regional variations in deposition more clearly. We therefore retain the original colour scale.

2.2 A short description of model configuration, especially those closely related to transport dynamic would be helpful to understand the difference in the results. A short summary of the model performance and its implication at the end of this section would be nice.

We have extended section 2 to include a more detailed paragraph at the beginning to describe the models, with additional section sub-headings. The models all include parameterizations of key aerosol processes such as nucleation, condensation and coagulation and simulate the transport of aerosol through sedimentation and large-scale circulation by the Brewer-Dobson circulation. There are differences in these parameterizations across the models. However, the authors feel that it is beyond the scope of this paper to address these in detail, given that it would be impossible with these simulations alone to determine the relative importance of each process, in addition to differences arising from model resolution and the individual deposition schemes. The inter-model differences found in this study are motivation for more dedicated multi-model comparison projects, which explicitly look at differences in model processes, such as the Interactive Stratospheric Aerosol Modelling Intercomparison Project (ISA-MIP). We have introduced this project in the conclusion.

We have added a brief description of the processes that are included in each model as well as highlighting in this section that the representation of the QBO differs in each model, which will impact the initial transport of sulfate aerosol. We have also added a brief description of the deposition schemes used in the models and further references for each model, which the reader can explore. We have also moved the details from Table 1 in the supplementary to Table 1 in the main paper.

The deposition schemes in SOCOL-AER are relatively simple and are not related to the precipitation. We have therefore removed references to precipitation-deposition correlations for SOCOL-AER throughout the paper.

At the end of section 3.1 we have extended the last paragraph to explain each model's polar deposition and have added the following text to summarize the model performance and the implications:

"Overall, the magnitude of the deposited sulfate in CESM1(WACCM) and MAECHAM5-HAM, where deposition to the ice sheets is dominated by wet deposition, is expected to be

driven by the snow accumulation rates across the ice sheets, which are well represented by all models (Fig. S3). In UM-UKCA, although the polar deposition is correlated with the polar precipitation, the ice sheet sulfate deposition mostly occurs by dry deposition. This is because this model de-activates nucleation scavenging if more than a threshold fraction of the cloud water is present as ice, greatly reducing the aerosol scavenging in polar regions. In SOCOL-AER, fewer regional details are captured since the deposition scheme is simpler and is not connected to precipitation, and therefore the deposition mostly reflects the tropospheric distribution of sulfate.

In summary the models simulate similar overall patterns of background sulfate deposition fluxes, although there are differences in the regional details and magnitude. The similarities and realistic deposition patterns amongst the models suggests that the background sulfate emissions, transport and deposition processes are reasonably parameterized. Although SOCOL-AER is less able to simulate regional details, its simplified deposition scheme is still sufficient for the analysis of inter-hemispheric differences and the temporal evolution of deposition.

2.3 P7L3-4, "the four models simulate similar background sulfate deposition patterns and magnitudes and compare well to pre-industrial ice core sulfate fluxes", please provide a table lists all pairs of the model-ice core values to support this statement.

We agree that a table of the comparisons would be useful, and indeed this information was provided in supplementary Figure 4, which showed scatter plots of the ice core fluxes versus those simulated in each model. We have therefore moved this figure to the main text and added a 1:1 line so that the reader can more easily see how each model-simulated value compares to each ice core for both Antarctica and Greenland, and also to the other models. As such, all of the comparisons can be seen in this one figure and we do not think it is necessary to have a table as well. This figure can now be directly compared to the same figure as for the Tambora sulfate ice core deposition fluxes (Figure 7) so the reader can more easily compare how the models perform under the background and perturbed conditions.

The new SOCOL-AER runs show higher deposition to the ice sheets and therefore we have rephrased this sentence to not include SOCOL-AER, where the comparisons are worse than for the other three models. We have also added a sentence to reference the new figure (now Figure 3).

2.4 The differences between wet and dry deposition across models have been discussed at various part of the paper, please explain in more depth what are the implication of these differences.

The differences between wet and dry deposition across the models are due to the individual deposition scheme parameterizations. If dry deposition is more spatially uniform across the ice sheets than wet deposition, then the spatial pattern of sulfate deposition flux is a function of the wet deposition, but the total magnitude, dependent on the proportion of wet vs dry deposition. In comparison to differences between sedimentation, aerosol transport and stratosphere-troposphere exchange between the models, we do not think that differences in the wet and dry partitioning is important in explaining these results.

To address this uncertainty, however, we have added the following text to the discussion (Sect. 4.1.3):

"The differences between wet and dry deposition simulated across the models are due to the individual deposition scheme parameterizations. The implication of these differences in dictating the resulting total sulfate deposition remains uncertain. However, since inter-model differences in volcanic sulfate deposition patterns appear unrelated to differences between climatological wet and dry deposition patterns, the proportion of wet vs dry deposition is likely of secondary importance compared to differences between the models in aerosol transport processes including sedimentation and stratosphere-troposphere exchange."

2.5 Some of the discussions are overlapping or repeating, for example, the temporal evolution of different models in the last paragraph of section 3.2.1 and section 3.2.2.
Please consider combine them to shorten the discussion.

We thank the reviewer for highlighting this and have rearranged section 3.2 and section 3.3 to avoid any major repetitions. We have first focussed on the global sulfate deposition (both spatial and temporal) (section 3.2.1) and then the ice sheet deposition (both spatial and temporal) (3.2.2). We have moved the description of the temporal evolution of sulfate burdens to section 3.3, which is now a dedicated section for all results relating to sulfate burdens and the relationship between burden and deposition. We have renamed the sections accordingly.

We have also moved the paragraphs on winds to the discussion (section 4.1.2).

3. In section 4 "Discussion"

3.1 Again, some of the discussions repeat the results in section 3. Please focus more on discussing the implication of the results, for example, the causes of the difference in model simulated Tambora deposition.

We have re-written the discussion section to focus more on explaining the differences in model simulated deposition for Tambora and have reduced the emphasis on precipitation (much of this text has been removed). We have split the discussion into the following sections to address each reason (sulfate formation and transport, stratospheric winds and polar vortices, and the deposition schemes) in turn:

4.1 Differences in deposited sulfate
4.1.1 Volcanic sulfate formation and transport
4.1.2 Dynamical effects
4.1.3 Deposition schemes
4.2 Implications for model differences in simulated deposition

We have also removed the sentence referring to aerosol size in different modes (page 13, lines 10-13) since this may be misleading without a comprehensive investigation of aerosol particle size.

3.2 Please discuss why models cannot give a converged simulation of Tambora deposition, while they were able to simulate the preindustrial background sulfate deposition well.

Because the models are able to simulate the background deposition reasonably well, the inter-model differences in volcanic sulfate deposition are most likely due to differences in the volcanic aerosol formation and aerosol size due to differences in aerosol microphysics, stratospheric aerosol transport and stratosphere-troposphere exchange. In the background most of the deposited sulfate is of tropospheric origin. Differences in deposition may also become more pronounced in the perturbed case than in the background due to the higher sulfate aerosol burden. Scavenging and deposition parameterizations are highly uncertain, and the chance that

such parameterizations become unrealistic under the large sulfate aerosol loadings associated with a Tambora eruption cannot be discounted, and should be explored in future work.

We hope that the re-written discussion emphasizes these reasons. We have also re-written the conclusions to highlight these reasons.

3.3 P14L28-P15L2, the discussion in these lines seems unnecessary to me since the focus of this study is on model intercomparison.

Although we agree that this discussion is surplus to the inter-model focus, we argue that this paragraph is still useful to further put the results in context especially if a reader were to use the ice core comparison to infer model skill. We therefore feel it is necessary to report some of the issues with ice core derived sulfate fluxes.

4. In section 5 "Conclusions"
4.1 P16 L10-11, "Our derived BTD factors highlight uncertainties ..." Not necessary the actual uncertainties between the atmospheric burden and ice sheet deposition, but the uncertainties in the model's ability to derive the relationship. Therefore, I would recommend the authors to rephrase the sentence to make this distinguish.

The authors agree this was misleading and have rephrased the sentence as follows:
"Our range in derived BTD factors highlights uncertainties in the relationship between atmospheric sulfate burden and ice sheet deposited sulfate as simulated by models"

4.2 P16L19-21, "Using..... will provide the opportunity to better understand model diversity and to advance our understanding of the climate response to large volcanic eruptions". It is true that using the same prescribed forcings could help us to better understand model diversity, but only true to advance the understanding of the volcanic climate response if the prescribed forcings are assumed to be correct. And if that is the case, what about the goal of this VolMIP study to improve the ice-core-based reconstruction?

Agreed. We have removed the sentence about advancing our understanding.

Review 2: Anonymous

This study does a careful comparison of the sulfate aerosol deposition from the Mt. Tambora eruption to test the models ability to simulate deposition observed in the ice core record as well as the assumption made to back out SO2 injections from ice core sulfate signals. This work is done using a variety of models that include microphysical aerosol modules and highlights some of the successes as well as the continued work that needs to be done. I do find that it is a clearly written paper with results that would be of interest to the ACP community and would recommend publication with only a few mostly minor comments for authors to address.

We thank Reviewer 2 for their helpful comments and have addressed these below.

The main comment I have would be related to needing some additional discussion of hydroxyl radical (OH). I would like the authors to include more details on OH in the main paper and possibly consider a figure in the main text or supplement showing something like profiles of its tropical concentrations in the various models (both background and perturbed if applicable). The table S1 would be best to include in the actual paper rather then the supplement. 3 Models have interactive OH and 1 prescribed. Given OH's critical role in the conversion of SO2 into sulfate aerosol and given the differences in sulfate aerosol evolution in the different models it would be really helpful to look at whether any differences in the amount of OH or its distribution can explain the differences in sulfate conversion noted on page 8 lines 17-20.

The authors acknowledge the importance of OH in dictating the initial aerosol formation but the effects of OH are second order to differences in aerosol sedimentation and large-scale transport when explaining the inter-model deposition. However, we agree that the paper will be enhanced by adding these details.

We have combined the details of Table S1 with Table 1 and have added an additional paragraph in section 2 to describe the models, which also highlights the OH details in each model. Photolysis rates are not impacted by the sulfate aerosol in any of the models and we have added this statement. Columns labelled "injection height" and "location of injection" are not included in the new Table 1. This information is instead specified under Table 2.

We find that the models have similar background concentrations and distributions of stratospheric OH and have added this statement to the discussion (section 4.1.1) and included this figure in the supplementary (Figure S5). We find that in the models with interactive OH the tropical OH is depleted in the first ~2 months after the eruption. We have also added a supplementary figure of the percentage change in tropical OH concentrations in the first 8 months after the eruption (Figure S6). This figure illustrates the average tropical depletion but depletion rates in the actual aerosol plume would be much higher.

We have introduced the depletion of SO_2 in section 3.3 and have added the following text to section 4.1.1:

"MAECHAM5-HAM is the only model that has prescribed OH (Table 1). OH may become depleted in dense volcanic clouds by reaction with SO_2, affecting the rate of sulfate aerosol formation (Bekki, 1995). The background stratospheric OH concentrations are similar between the models (Fig. S5) but in SOCOL-AER, UM-UKCA and CESM1(WACCM), in the first 2 months after the eruption, stratospheric tropical OH becomes depleted, with ensemble mean peak reductions of between 15-33% (Fig. S6). This reduces the rate of sulfate aerosol formation compared to MAECHAM5-HAM where the SO_2 will be more rapidly oxidised, and explains the later peaks in sulfate burdens in these models."

For the models with interactive OH does the sulfate aerosol impact photolysis rates, which would decrease OH formation and slow conversion. Does stratospheric water vapor increase in these runs, increasing OH production? Do any of the models deplete OH when reacting with SO2? If so it would be important to note in the text, if not mention as a source of uncertainty in the sulfate conversion.

Photolysis rates are not impacted by the sulfate aerosol in any of the models and we have added this statement to section 2. Because studies have shown that this effect is not as important as reductions in OH due to depletion by SO_2 (e.g. Mills et al., 2017, JGR), we do not think it is necessary to add any more discussion here. The models do deplete the OH when reacting with SO_2 (see answer above).

Not all the models outputted the stratospheric water vapour, so we have been unable to investigate this in detail. We find that in all models the OH concentration increases after the

initial depletion (due to oxidation of SO_2); in SOCOL-AER, where the stratospheric water vapour was available, we find that the stratospheric water vapour does increase synchronously with the OH increase. These findings are shown in the figures below:

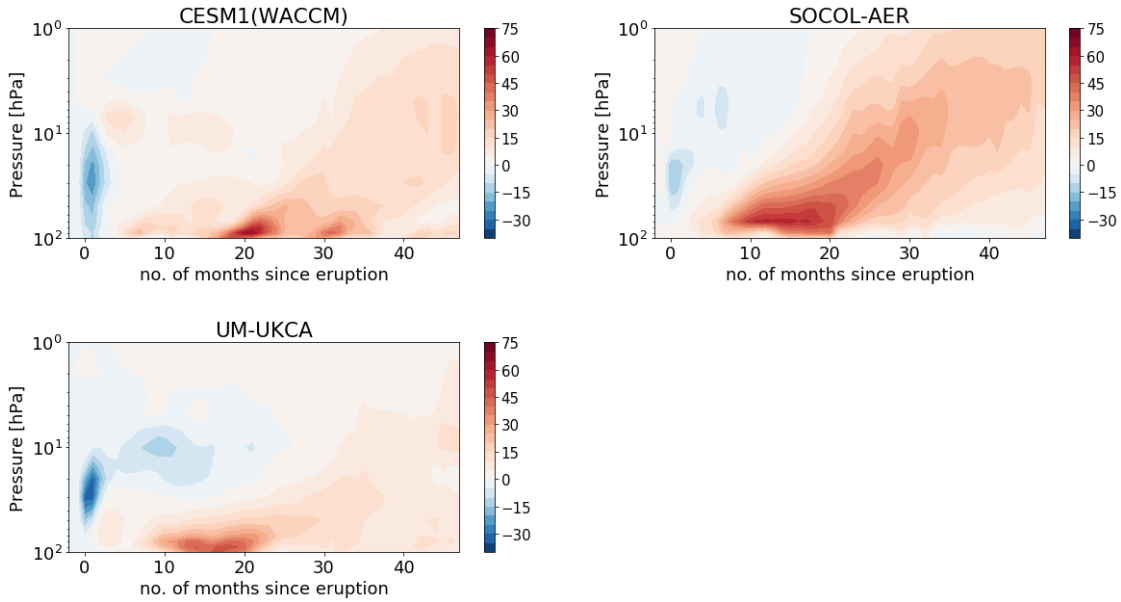

Figure 1: Percentage change in tropical OH for each model (ensemble mean)

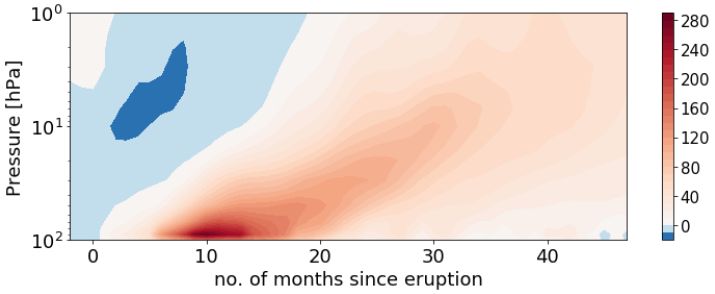

Figure 2: Percentage change in stratospheric water vapour in SOCOL-AER (ensemble mean). Stratospheric water vapour does increase similar to the increase in OH (see Fig. 1).

Given that the later increase in OH does not affect the sulfate aerosol burden (since it is now decaying) and the focus of the paper on deposition, we argue that additional details on stratospheric water vapour are beyond the scope of the paper.

page 5 line 2 you should add "as emitting" after simulated or something similar.

We have rephrased the sentence as follows: "The eruption was simulated by emitting the SO_2 over 24 hours on 1 April"

page 6 line 20-22 Is there a notable difference between UKCA and other models in this regard that would be worth discussing it seems like a potentially important point concerning the focus of this paper w.r.t. deposition schemes.

We have found that UM-UKCA de-activates nucleation scavenging in clouds with a lot of ice, which accounts for why deposition is predominantly dry over the ice sheets. We have mentioned this in the text when explaining the deposition differences, i.e., in section 3.1. Please also see the replies to reviewer 1 point 2.2.

In general, given the focus of this paper a brief mention of the deposition scheme used in each model and reference would be very helpful.

We agree that these details would be very helpful so have added a description of these schemes in section 2 along with additional references.

page 13 lines 13-14 More discussion about OH here and earlier would be helpful

Please see replies to earlier comments and tracked changes in the revised manuscript.

page 13 lines 15-17 I don't think it is beyond the scope of this paper to show or discuss the OH since it is critical in the formation of the sulfate aerosols and could help address difference in the the sulfate burdens

Please see replies to earlier comments. We have included an additional figure (Figure S6) to show the changes in OH after the eruption.

page 14 lines 19-20 sentence starting with "Even if the models were perfect" I would recommend removing this sentence, it is not necessary and confusing

We agree and have removed the sentence.

Table S1 SOCOL is listed as 8S location of injection the rest are equator is the a typo or real difference in injection latitude.

This is real. We argue that since it is close enough to the equator it can still be defined as an equatorial eruption. We do not think there are significant implications of this that warrant further discussion in the paper. Since combining the information from Table S1 to Table 1, this information has been added as a statement under Table 2.

Multi-model comparison of the volcanic sulfate deposition from the 1815 eruption of Mt. Tambora

Lauren Marshall[1], Anja Schmidt[1,*], Matthew Toohey[2,3], Ken S. Carslaw[1], Graham W. Mann[1,4], Michael Sigl[5], Myriam Khodri[6], Claudia Timmreck[3], Davide Zanchettin[7], William Ball[8,9], Slimane Bekki[10], James S.A. Brooke[11], Sandip Dhomse[1], Colin Johnson[12], Jean-Francois Lamarque[13], Allegra LeGrande[14], Michael J. Mills[13], Ulrike Niemeier[3], James O. Pope[15], Virginie Poulain[6], Alan Robock[15,6], Eugene Rozanov[8,9], Andrea Stenke[8], Timofei Sukhodolov[9], Simone Tilmes[13], Kostas Tsigaridis[14], Fiona Tummon[8]

[1]Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, UK
[2] GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany
[3] Max Planck Institute for Meteorology, Hamburg, Germany
[4]National Centre for Atmospheric Science, University of Leeds, UK
[5]Laboratory of Environmental Chemistry, Paul Scherrer Institut, 5232 Villigen, Switzerland
[6]IRD/IPSL/Laboratoire d'Océanographie et du Climat, Paris, France
[7]Department of Environmental Sciences, Informatics and Statistics, University Ca' Foscari of Venice, Mestre, Italy
[8]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[9]PMOD/WRC, Davos, Switzerland
[10]LATMOS-IPSL, Université UPMC/Paris-Sorbonne, Université UVSQ/Paris Saclay, CNRS/INSU, Paris, France
[11]School of Chemistry, University of Leeds, UK
[12]Met Office Hadley Centre, Exeter, UK
[13]Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA
[14]NASA Goddard Institute for Space Studies and Center for Climate Systems Research, Columbia University, New York, NY, USA
[15]British Antarctic Survey, Cambridge, UK
[15,6]Department of Environmental Sciences, Rutgers University, New Brunswick, NJ, USA
[*]now at Department of Chemistry, University of Cambridge, UK and Department of Geography, University of Cambridge, UK

Correspondence to: Lauren Marshall (eelrm@leeds.ac.uk)

Abstract. The eruption of Mt. Tambora in 1815 was the largest volcanic eruption of the past 500 years. The eruption had significant climatic impacts, leading to the 1816 'Year Without a Summer' and remains a valuable event from which to understand the climatic effects of large stratospheric volcanic sulfur dioxide injections. The eruption also resulted in one of the strongest and most easily identifiable volcanic sulfate signals in polar ice cores, which are widely used to reconstruct the timing and atmospheric sulfate loading of past eruptions. As part of the Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP), ~~four~~five state-of-the-art global aerosol models simulated this eruption. We analyse both simulated background (no Tambora) and volcanic (with Tambora) sulfate deposition to polar regions and compare to ice core records. The models simulate overall similar patterns of background sulfate deposition, although there are differences in regional details and magnitude. ~~Background sulfate deposition is of similar magnitude across all models and compares well to ice core records.~~ However, the volcanic sulfate deposition varies considerably between the models with differences in timing, spatial pattern and magnitude~~, between the models.~~ Mean simulated deposited sulfate on Antarctica ranges from 19 to 264 kg km^{-2}, and on Greenland from 31 to 194 kg km^{-2}, as compared to the mean ice core-derived estimates of roughly ~~40~~ 50 kg km^{-2}, for both Greenland and Antarctica. The ratio of the hemispheric atmospheric sulfate aerosol burden after the eruption to the average ice sheet deposited sulfate varies between models by up to a factor of 15. Sources of this inter-model variability include differences in both the formation and the transport of sulfate aerosol. Our results suggest that deriving relationships between sulfate deposited on ice sheets and atmospheric sulfate burdens from model simulations may be associated with greater uncertainties than previously thought.~~highlight the uncertainties and difficulties in deriving historic volcanic aerosol radiative forcing of climate, based on measured volcanic sulfate in polar ice cores.~~

1 Introduction

Mt. Tambora in Indonesia (8.2°S, 118.0°E) erupted in April 1815 (e.g. Oppenheimer, 2003) and had a considerable impact on climate, leading to widespread tropical and Northern Hemisphere mean cooling of ~1°C and a 'Year Without a Summer' in 1816 (e.g. Raible et al., 2016). Volcanic sulfate aerosol, produced from the oxidation of sulfur dioxide (SO_2) emitted into the atmosphere by volcanoes, is transported throughout the atmosphere and deposited to the surface by both wet and dry processes, some of which is eventually incorporated into polar ice (e.g. Robock, 2000). Bipolar volcanic sulfate deposition signals are presumed to result from tropical eruptions, whereby sulfur entering the tropical stratosphere is converted to sulfate aerosol, which is transported globally by the Brewer-Dobson circulation (e.g. Trepte et al., 1993; Langway et al., 1995; Robock, 2000; Gao et al., 2007). Polar ice core deposition signals typically start around 0.5-1 year after a large tropical eruption and remain elevated for approximately 2 - 3 years (Robock and Free, 1995; Sigl et al., 2015). Throughout the last 2500 years, polar ice core records show over 200 sulfate spikes, which have been used to estimate the timing, evolution and magnitude of radiative forcing of climate caused by volcanic eruptions during this period (Sigl et al., 2015). The 1815 eruption of Mt. Tambora produced the 6th largest bipolar sulfate signal in the last 2500 years (Sigl et al., 2015).

Determining the stratospheric aerosol properties of the 1815 Mt. Tambora eruption such as spatial extent of the sulfate aerosol cloud, aerosol optical depth and aerosol size distribution bears substantial uncertainties, which ultimately affects the quantification of its climatic impacts using climate models. As part of the Model Intercomparison Project on the climatic response to volcanic forcing (VolMIP) (Zanchettin et al., 2016), which is a Coupled Model Intercomparison Project Phase 6 (CMIP6) endorsed activity (Eyring et al., 2016), coordinated simulations of the 1815 eruption of Mt. Tambora were performed with five~~four~~ state-of-the-art global aerosol models. Our study, motivated by the uncertainty that remains in the climatic forcing from this eruption, investigates the sources of uncertainty in the sulfate deposition to polar regions in these simulations, and discusses implications for reconstructions of historic volcanic forcing.

Previous reconstructions of volcanic sulfate aerosol properties used to force climate models scaled the average sulfate deposited on Antarctica and Greenland to the hemispheric atmospheric sulfate aerosol burden (e.g. Gao et al., 2007; Crowley & Unterman~~n~~, 2013; Sigl et al., 2015). Scaling factors (ratios of the hemispheric sulfate aerosol burden to the sulfate deposited at the poles) were based on the ratio of these two quantities as observed after the eruption of Mt. Pinatubo in 1991 and from the estimated atmospheric burden and measured deposited radioactive material after nuclear bomb tests. Previous climate model simulations of the ratio between atmospheric sulfate burden and polar deposited sulfate and were also used to derive the scaling factors (Gao et al., 2007; 2008). These scaling factors may not hold for larger eruptions where volcanic sulfate aerosol particles can grow larger, increasing their sedimentation rate (e.g. Pinto et al., 1989; Timmreck et al., 2009). Toohey et al. (2013) also found that differences in the dynamical response to large-magnitude eruptions changed the spatial distribution of the deposited sulfate. Furthermore, available ice core measurements are not evenly distributed over both ice caps, and large spatial variations in the sulfate deposition fluxes can exist between individual ice cores due to differences in local accumulation rates and sulfate redistribution by snow drift (Clausen and Hammer, 1988; Zielinski et al., 1997; Cole-Dai et al., 1997, 2000; Wolff et al., 2005; Gao et al., 2006, 2007). It is therefore important that a range of ice core records from different geographical regions is used to estimate the average volcanic sulfate deposited on each ice cap. Previous studies using only a few ice cores to reconstruct volcanic forcing histories may be biased (e.g. Zielinski, 1995; Zielinski et al., 1996; Crowley, 2000), although it has been demonstrated that deposition fluxes derived from single ice cores at high accumulation sites are representative of total ice sheet deposition (Toohey and Sigl, 2017). Gao et al. (2007), who analysed 44 ice cores to investigate the spatial distribution of volcanic sulfate deposition during the last millennium, found larger average deposited sulfate on Greenland (mean deposition of 59 kg km^{-2}, using 22 ice cores) than on Antarctica (mean deposition of 51 kg km^{-2}, 17 ice cores) for the eruption of Mt. Tambora. However, Sigl et al. (2015) found, using additional high temporal-resolution ice core records in Antarctica (Sigl et al., 2014), average Antarctic deposited sulfate of 46 kg km^{-2}, and a smaller average deposited sulfate on Greenland of 40 kg km^{-2}, with both averages smaller than the averages provided by Gao et al. (2007). Although in Sigl et al. (2015) the Antarctic average was derived with 17 ice core records**,** the Greenland average was calculated from only two ice cores (NEEM and NGRIP) compared to the 22 cores used for Greenland in Gao et al. (2007).

Previous modelling studies that have investigated the sulfate deposition from the 1815 eruption of Mt. Tambora have failed to reproduce the magnitude of the measured deposited sulfate on both ice caps compared to ice core records, although

the models were able to capture the spatial pattern (Gao et al., 2007; Toohey et al., 2013). Gao et al. (2007) found the model-simulated mean deposited sulfate to be a factor of 2~~two~~ greater than the ice core-derived estimate, with average Antarctic deposited sulfate of 113 kg km^{-2} and smaller Greenland deposited sulfate of 78 kg km^{-2}. Toohey et al. (2013), on the other hand, found higher deposition to Greenland, and although matching the spatial pattern of deposited sulfate on Antarctica remarkably well, found model-simulated mean deposited sulfate to be ~4.7 times greater than inferred from ice cores. Differences between simulated and measured deposited sulfate could be caused by inaccuracies in the model representation of several physical processes such as the formation and transport of sulfate aerosol, sedimentation, cross-tropopause transport and deposition processes (e.g. Hamill et al., 1997; SPARC, 2006). Neither of the models used by Toohey et al. (2013) and Gao et al. (2007) included a representation of the Quasi-Biennial Oscillation (QBO), which may significantly impact the initial aerosol dispersion (e.g. Trepte et al., 1993). Furthermore, uncertainties exist in the source parameters used for simulating the eruption in models such as the SO_2 emission magnitude and emission height.

In general, sulfate deposited on the polar ice caps is only a small fraction of the sulfate deposited globally (e.g. Toohey et al., 2013) and there remains uncertainty surrounding the partitioning of the 1815 Mt. Tambora volcanic sulfate aerosol between both hemispheres. Model results can aid in the interpretation of the ice core estimates, by allowing us to assess the relationship between the simulated atmospheric sulfate aerosol burdens and the simulated deposited sulfate.

In this paper we focus on the model-simulated sulfate deposition and the implications for reconstructions of historic volcanic forcing by analysing the deposited sulfate simulated by four global aerosol models and comparing to ice core records. Sect.~~ion~~ 2 describes the model simulations and ice core records. In S~~s~~ect.~~ion~~ 3 we assess the sulfate deposition simulated under both background (no Tambora) (Sect. 3.1) and volcanically perturbed (with Tambora) conditions (Sect. 3.2) and compare the simulated deposited sulfate to ice core measurements. We investigate the relationship between hemispheric atmospheric sulfate burdens and mean ice sheet deposited sulfate in S~~s~~ect.~~ion~~ 3.3 and explore reasons for model differences in S~~s~~ect.~~ion~~ 4. Conclusions are presented in S~~s~~ect.~~ion~~ 5.

2 Model~~s set up~~ and ice core data

2.1 Model descriptions

Of the five models that took part in the coordinated simulations of the 1815 eruption of Mt. Tambora (Zanchettin et al. 2016) only four simulated the sulfate deposition and are therefore included in our study. Model details are listed in Table 1. ~~The four global aerosol models included in this study are listed in Table 1.~~ In each model aerosol formation and growth is simulated through parameterizations of nucleation, condensation and coagulation. Three of the four models have modal aerosol schemes in that the aerosol particle size distribution is represented by several log-normal modes. SOCOL-AER has~~In~~ a sectional scheme where the aerosol particle size distribution is represented by 40 discrete size bins ~~(e.g. SOCOL)~~. The models simulate the transport of stratospheric aerosol through sedimentation and large-scale circulation by the Brewer-Dobson circulation. The QBO is simulated by all models except for MAECHAM5-HAM and is either internally generated (UM-UKCA) or nudged

(CESM1(WACCM) and SOCOL-AER). In CESM1(WACCM), MAECHAM5-HAM and UM-UKCA dry deposition schemes are resistance-based and wet deposition is parameterized based on model precipitation and convective processes, with aerosol removal calculated via first-order loss processes representing in-cloud and below-cloud scavenging (Stier et al., 2005; Mann et al., 2010; Lamarque et al., 2012; Liu et al., 2012; Bellouin et al., 2013; Kipling et al., 2013). In SOCOL-AER dry deposition is calculated by multiplying concentrations in the lowest model level by fixed values depending on surface cover. Wet deposition in SOCOL-AER is not related to the precipitation in the model and tropospheric wet removal rates are 5 days mean lifetime for H_2SO_4 (Sheng et al., 2015a). Apart from MAECHAM5-HAM, the models include interactive hydroxyl radical (OH) chemistry, allowing OH concentrations to evolve throughout the simulations (Sect. 4.1.1). Photolysis rates are not impacted by sulfate aerosol in any of the models. ~~A more comprehensive list of model specifications is provided in the supplementary material (Table S1).~~

All four models simulate the 1991 eruption of Mt. Pinatubo in reasonable agreement with observations of the sulfate burden, aerosol optical depth and stratospheric heating (Niemeier et al., 2009; Toohey et al., 2011; Dhomse et al., 2014; Sheng et al., 2015b; Mills et al., 2016), giving confidence in the models' overall abilities to accurately simulate the atmospheric and climatic effects of a large-magnitude eruption. However, the models vary in the details regarding the model-observation comparisons. For example, MAECHAM5-HAM (Niemeier et al., 2009) and SOCOL-AER (Sheng et al., 2015b) simulated a too rapid aerosol decay and UM-UKCA (Dhomse et al., 2014) had a low bias in the model-simulated aerosol effective radius compared to observations. Possible reasons for these differences include omitted or under-represented influences from meteoric particles, too large sedimentation and cross-tropopause transport and too fast transport from tropics to high latitudes. Conversely, the models differ in the amount of emitted SO_2 required to achieve good comparisons to observations with the mass of SO_2 emitted by the four models ranging from 10 Tg for UM-UKCA (Dhomse et al., 2014) and CESM1(WACCM) (Mills et al., 2016; 2017) to 12-14 Tg for SOCOL-AER (Sheng et al., 2015b) to 17 Tg for MAECHAM5-HAM (Niemeier et al., 2009; Toohey et al., 2011). For this reason, the use of a common protocol in this study (Sect. 2.2) enables us to better attribute potential differences in the results to model processes rather than to the eruption source parameters.

2.2 Experiment setup

The parameters used for the Mt. Tambora simulations are listed in Table 2. Each model simulated the eruption by emitting 60 Tg of SO_2 at the approximate location of Mt. Tambora ~~(Table S1),~~ between approximately 22-26 km (see details in Table 2 regarding the injection details for each model) and during the easterly QBO phase. This SO_2 emission estimate is based on both petrological and ice core estimates (Self et al., 2004; Gao et al., 2008), however there remains uncertainty regarding the amount of SO_2 emitted, which could range between ~30 – 80 Tg SO_2 (e.g. Stoffel et al., 2015). Nevertheless, 60 Tg SO_2 remains our best estimate. There is also uncertainty in the altitude of the emission and QBO phase due to the lack of observations. Therefore, the injection altitude and QBO phase were chosen to match those of the 1991 Mt. Pinatubo eruption based on satellite and lidar observations (McCormick and Veiga, 1992; Read et al., 1993; Herzog and Graf, 2010). The eruption was simulated by emitting the SO_2 over ~~for~~ 24 hours on 1 April.

In MAECHAM5-HAM, SOCOL-AER and UM-UKCA simulations were atmosphere-only with prescribed pre-industrial ~~S~~sea surface temperatures. In CESM1(WACCM) the simulations were run in a pre-industrial coupled atmosphere-ocean mode.~~,~~ ~~greenhouse gas concentrations, tropospheric aerosols and ozone were set to C~~climatological pre-industrial settings were used for greenhouse gas concentrations, tropospheric aerosols and ozone as defined by each modelling group. The simulations were run for five years and included five ensemble members, except for CESM1(WACCM), which had three members only. The models include additional species and processes compared to earlier modelling studies of Mt. Tambora (e.g. Gao et al., 2007; Toohey et al., 2013). UM-UKCA for example includes meteoric smoke particles (Brooke et al., 2017) and an internally generated QBO. Model output is in the form of monthly means.

2.3 Ice core data

The ice cores used in this analysis are provided in Tables S~~2~~1-S2~~3~~. The Antarctic ice cores are the most extensive array of annually resolved cores that have been used to reconstruct historic volcanic forcing (Sigl et al., 2014; 2015). Greenland ice core records have been compiled from several studies (Table S1~~2~~). Further ice core measurements of the natural background sulfate deposition fluxes were taken from Lamarque et al. (2013).

Sulfate deposition fluxes are derived from ice cores by multiplying measured sulfate concentrations by the annual ice accumulation rate. To derive the volcanic sulfate deposition flux contribution the natural sulfate background level (e.g. due to marine biogenic sulfur emissions) is calculated in each ice core (the non-volcanic contribution) and a threshold flux value is chosen, above which sulfate is assumed to be of volcanic origin. The ice core-derived volcanic sulfate deposition flux is then calculated as the difference between a year with the volcanic contribution and the mean of the non-volcanic years, and the resulting reported volcanic sulfate deposition flux is the sum of the fluxes in these perturbed years (Ferris et al., 2011; Cole-Dai et al., 2013; Sigl et al., 2013). Our comparable model-simulated volcanic deposition flux is calculated as the sum of the sulfate deposition anomaly (perturbed run minus control run) over the duration of the deposition signal (~2-4 years). For SOCOL, which has a sectional aerosol scheme, diagnostics are available for the wet and dry components of the sulfate deposition. For modal models, each of these components is further split into the contribution from each aerosol size mode simulated in the models (nucleation, Aitken, accumulation, coarse). The sulfate deposition flux is calculated (comparable to the ice core derived values) as the sum of all of these wet and dry components with a composition of SO_4^{2-} only. In the following sections, we define 'total deposition' as referring to the sum of wet and dry deposition fluxes. We define 'volcanic sulfate deposition' to specify the sulfate deposition flux anomaly due to the eruption of Mt. Tambora, and use 'cumulative deposited sulfate' to specify the time-integrated volcanic sulfate deposition fluxes.

To compare the model-simulated results with ice core values, we calculate two statistical metrics; the Normalized Mean Bias (NMB) and correlation coefficient (r). NMB is defined by:

$$NMB = \frac{\sum_{i=1}^{N}(M_i - O_i)}{\sum_{i=1}^{N}(O_i)}, \quad (1)$$

where O_i is the ice core derived sulfate deposition and M_i is the simulated sulfate deposition in the model grid box containing the ice core. N is the number of ice core records. For both NMB and r, each ice core is given equal weighting. We define a high correlation as $r > 0.7$ and low correlation as $r < 0.3$.

3 Results

3.1 Pre-industrial background sulfate deposition

Figure 1 shows the average annual ~~mean~~ sulfate deposition fluxes in the pre-industrial control simulations (no Tambora) for each model. Areas of high background sulfate deposition fluxes are found in close proximity to sulfur emission sources such as continuously degassing volcanoes (e.g. in South America and Indonesia) and along and near mid-latitude storm tracks (30°–60°) where aerosol is removed effectively by precipitation (except SOCOL-AER where the deposition is not affected by precipitation). Continuously degassing volcanic emissions are not included in MAECHAM5-HAM. Sulfate deposition fluxes are higher over the oceans than over the land, mainly due to the emission of marine dimethyl sulfide (DMS). In general, Fig. 1 shows that the models have similar background sulfate deposition patterns, with the global mean total (wet + dry) sulfate deposition flux ranging from 78 kg SO_4 km^{-2} yr^{-1} (CESM1(WACCM)) to 173 kg SO_4 km^{-2} yr^{-1} (UM-UKCA).

We find that the pre-industrial background ~~pre-industrial~~ global mean atmospheric sulfate burdens are similar between CESM1(WACCM), MAECHAM5-HAM and SOCOL-AER, but ~2-3 times larger in UM-UKCA (Fig. S1). Sulfur source species included in each model are listed in Table ~~S~~1. Although the models have similar background sulfate deposition patterns, the partitioning of wet and dry deposition fluxes differs markedly between the models (Fig. 1, Table 3). MAECHAM5-HAM deposits very little sulfate by dry processes compared to the other models with annual global -total dry deposited sulfate a factor of 40 less than the global -total wet deposited sulfate. In SOCOL-AER, dry deposited sulfate is approximately half the magnitude of wet deposited sulfate.

The sulfate deposited on Antarctica and Greenland is a very small fraction (less than 1%) of the sulfate deposited globally. In UM-UKCA the sulfate deposited on the polar ice sheets is dominated by dry deposition, which is supported by observations (Legrand and Mayewski, 1997), especially in the Antarctic interior (Wolff, 2012). In contrast, in MAECHAM5-HAM, SOCOL-AER and CESM1(WACCM) the sulfate deposited on the polar ice sheets is dominated by wet deposition ~~(i.e. through precipitation)~~, suggesting an issue with the deposition or precipitation representation. However, we find that the simulated total precipitation compares well between models both globally and over the poles (Fig. S2-S3) indicating the differences in wet and dry deposition partitioning are due to each model's deposition schemes.

The annual global -total deposition for both SO_2 and SO_4 is listed in Table 3 for each model. Included for reference is the equivalent preindustrial SO_X (SO_2 + SO_4) deposition from the multi-model mean of the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) (Lamarque et al., 2013, their Table S4a). The ACCMIP simulations were set up as time-slice experiments and the multi-model mean listed is an average of 6 models. UM-UKCA compares well to the ACCMIP multi-model mean for dry SO_X, but the wet SO_X is 7 Tg S yr^{-1} higher and the SO_4 deposition (29 Tg S yr^{-1}) is also

much higher when compared to the other models (13-19 Tg S yr^{-1}). MAECHAM5-HAM has a similar total for wet SO_X compared to the ACCMIP multi-model mean, but dry deposition is a factor of 4 lower. CESM1(WACCM) has a similar total for wet SO_X deposition compared to the ACCMIP multi-model mean but total SO_X is 5 Tg S yr^{-1} lower. SOCOL-AER simulates the highest dry SO_X (18 Tg S yr^{-1}) and total SO_X (44 Tg S yr^{-1}) with total SO_X 10 Tg S yr^{-1} greater than the ACCMIP multi-model mean.~~matches the ACCMIP multi-model mean for total SO_X~~.

Following the analysis of Lamarque et al. (2013) we have taken the average sulfate deposition fluxes from 1850 to 1860 (a non-volcanic period) in several ice cores from Antarctica and Greenland and compared the ice core fluxes to the modelled polar sulfate deposition fluxes in the control simulations (Fig. 2). Ice core meta-data are included in the supplementary information (Table S2).

Overall, CESM1(WACCM), MAECHAM5-HAM and UM-UKCA~~the four models~~ simulate similar background polar sulfate deposition patterns and magnitudes and compare well to pre-industrial ice core sulfate fluxes. Scatter plots of the ice core fluxes versus those simulated by each model are shown in Fig. 3. SOCOL-AER simulates slightly higher deposition with reduced regional variability compared to the other models (Fig. 2). However, ~~C~~compared to the ice cores, all models capture the lower sulfate deposition in the interior of Antarctica and higher sulfate deposition toward the coast. The models overestimate Antarctic deposition, particularly ~~However, model simulated sulfate deposition~~ in West Antarctica ~~is higher;~~ Antarctic NMB (Sec. 2.3, Eq. (1)) range from 1.3 (UM-UKCA) to 3.9~~2.9~~ (SOCOL-AER) but~~,~~ ~~w~~We find that the model-simulated Antarctic sulfate deposition and Antarctic ice core values are highly correlated for all models with *r* above 0.9 ~~for all models~~ (Fig. 3~~S4~~). Deposition over the Arctic is also well captured, with MAECHAM5-HAM and CESM1(WACCM) slightly underestimating the sulfate deposition fluxes, both with NMB of -0.1. UM-UKCA has a very small positive NMB of 0.01~~09~~ but SOCOL-AER has the highest Arctic deposition with a NMB of 1.7~~1~~. None of the models capture the low flux recorded in Alaska as also found by Lamarque et al. (2013).

The background polar sulfate deposition flux is highly correlated with the simulated mean polar precipitation for CESM1(WACCM), MAECHAM5-HAM and UM-UKCA~~ all models, which in general can explain the simulated deposition patterns~~. Correlation coefficients in the Arctic (60° to 90°) are between 0.8 (~~SOCOL,~~ MAECHAM5-HAM and UM-UKCA) and 0.9 (CESM1(WACCM)). The correlation coefficients are slightly higher in the Antarctic (-60° to -90°) with $r = 0.9$ for all models. In SOCOL-AER, the higher NMB between simulated polar sulfate deposition fluxes and ice core values is due to the more simplified deposition scheme in this model, which is not connected to the model's simulated precipitation. We find that the Antarctic precipitation in each model matches measured accumulation rates in ice cores (Fig. S3), and with a high correlation with *r* values of between 0.7 (SOCOL-AER, included for reference) to 0.9 (CESM1(WACCM), UM-UKCA). UM-UKCA and CESM1(WACCM) have very small NMB of ~0.1. MAECHAM5-HAM has a slightly higher NMB of 0.6 and SOCOL-AER a NMB of 0.~~7~~8. In the Arctic, the models also capture the precipitation reasonably well compared to the accumulation in the ice cores, with NMB of between 0.1 (UM-UKCA) and 0.5 (CESM1(WACCM)) but low correlation coefficients (*r* lies between 0.1-0.2 for all models). Thus, compared to the ice cores the models capture the magnitude and spatial pattern of the background polar precipitation. Overall, the magnitude of the deposited sulfate in CESM1(WACCM)

and, MAECHAM5-HAM and SOCOL, where deposition to the ice sheets deposition is dominated by wet deposition, is expected to be likely driven by the snow accumulation rates across the ice sheets, which are is well represented by all models (Fig. S3). In UM-UKCA, although the polar deposition is correlated with the polar precipitation, the ice sheet sulfate deposition mostly occurs by dry deposition. This is because this model de-activates nucleation scavenging if more than a threshold fraction of the cloud water is present as ice, greatly reducing the aerosol scavenging in polar regions. In SOCOL-AER, fewer regional details are captured since the deposition scheme is simpler and is not connected to precipitation, and therefore the deposition mostly reflects the tropospheric distribution of sulfate.

In summary the models simulate similar overall patterns of background sulfate deposition fluxes, although there are differences in the regional details and magnitude. The similarities and realistic deposition patterns amongst the models suggests that the background sulfate emissions, transport and deposition processes are reasonably parameterized. Although SOCOL-AER is less able to simulate regional details, its simplified deposition scheme is still sufficient for the analysis of inter-hemispheric differences and the temporal evolution of deposition.

3.2 1815 Mt. Tambora eruption sulfate deposition

3.2.1 Inter-model comparison of Global sulfate deposition

Figure 34 shows the zonal ensemble mean monthly volcanic sulfate deposition (left) and cumulative deposited sulfate (right) simulated by each model and highlights inter-model differences in the timing and spatial distribution of the deposited sulfate. Deposition occurs rapidly in MAECHAM5-HAM with 35% of the global -total deposition occurring in 1815 and the majority (60%) occurring in 1816. SOCOL-AER simulates the sulfate deposition starting slightly later than in MAECHAM5-HAM, with the majority of the deposition (75%) occurring in 1816. In contrast, only 9% of deposition in UM-UKCA occurs in 1815, with 55% in 1816 and 29% in 1817. In CESM1(WACCM) the deposition occurs even later, with no deposition occurring in 1815. Instead, 32% is deposited in 1816, 46% in 1817 and 17% in 1818. Deposition is longest in duration in CESM1(WACCM) and global -total sulfate deposition remains elevated at the end of the simulation (Fig. 54). In MAECHAM5-HAM deposition returns to near background levels by ~30 months after the eruption and ~40 months for UM-UKCA and SOCOL-AER. We find individual ensemble members are similar for each model and the ensemble spread in the global -total volcanic sulfate deposition over time is small, as shown in Fig. 54.

In UM-UKCA and CESM1(WACCM) most of the volcanic sulfate is deposited at mid-latitudes (30-60º). This contrasts with MAECHAM, where the deposition is globally more uniform, with greater deposition in the polar regions, and high deposited sulfate exceeding 360 kg SO_4 km^{-2} over West Antarctica, which is completely absent in the other models. In SOCOL-AER, deposition is greatest in the Southern Hemisphere mid-latitudes.

The models vary in the simulated relative contribution of wet deposition of sulfate and dry deposition of sulfate to the global -total cumulative deposited sulfate (Table 4), although the global -total is always dominated by wet deposition, as was also the case with the background sulfate deposition (Fig. 1, Table 3). Dry deposited sulfate in MAECHAM5-HAM is a

factor 15 lower than the dry deposited sulfate simulated by UM-UKCA and CESM1(WACCM). SOCOL-AER also simulates fairly low dry deposited sulfate (1.0~~0.8~~ Tg S).

~~The temporal and spatial evolution of the volcanic sulfate deposition ultimately reflects the evolution of the atmospheric volcanic sulfate burdens and the precipitation in each model. Figure 5 shows the zonal mean monthly mean and global total monthly mean atmospheric volcanic sulfate burdens for each model. MAECHAM has the fastest conversion of SO_2 to sulfate aerosol, with the global peak sulfate burden occurring only 4 months after the eruption (Fig. 5b). UKCA and SOCOL are next with the peak global sulfate burden occurring 6-7 months after the eruption, but the global burden in SOCOL decays a lot quicker than in UKCA. The global burden in WACCM peaks 12 months after the eruption and remains elevated until the end of the simulation and hence deposition in WACCM is longer lived. In all models there is stronger transport of the sulfate aerosol to the Southern Hemisphere compared to the Northern Hemisphere (Fig. 5a) likely due to the Brewer Dobson Circulation, which is stronger in the winter hemisphere.~~

3.2.2 ~~Multi-model comparison to ice core records~~Ice sheet sulfate deposition

Although the ~~all four~~ models simulated similar pre-industrial background ~~pre-industrial~~ (no Tambora) polar sulfate deposition (with the exception of SOCOL-AER) (Fig. 2), the simulated polar volcanic sulfate deposition varies in time, ~~both~~ spatially and in magnitude between the models. Figure 6 shows the simulated cumulative deposited sulfate for each model compared to the cumulative deposited sulfate measured in ice cores from Greenland and Antarctica for the 1815 Mt. Tambora eruption.

In general, the ice cores from Antarctica show lower volcanic sulfate deposition in East Antarctica and higher deposition over the Antarctic Peninsula, with deposited sulfate ranging from 13~~2.7~~ kg SO_4 km^{-2} (East Antarctica, core NUS07-7) to 133 kg SO_4 km^{-2} (Antarctic Peninsula, core Siple). In Greenland the ice core estimates range from 25 kg SO_4 km^{-2} (B20) to 85~~.4~~ kg SO_4 km^{-2} (D3).

We find MAECHAM5-HAM and SOCOL-AER simulate too much deposited sulfate on Antarctica and Greenland compared to the ice cores records (also seen in Toohey et al., 2013), whereas UM-UKCA and CESM1(WACCM) simulate deposited sulfate much closer to the ice core values (Fig. 6). For Antarctica the NMB are 3.9 for MAECHAM5-HAM, 2.0~~1~~ for SOCOL-AER, -0.5 for CESM1(WACCM) and -0.7 for UM-UKCA. For Greenland the biases are slightly lower: 2.6 for MAECHAM5-HAM, 1.8~~3~~ for SOCOL-AER, 0.1 for CESM1(WACCM) and -0.5 for UM-UKCA. However, although MAECHAM5-HAM is the model with the highest bias between the simulated cumulative deposited sulfate and ice core values, we find that the simulated Antarctic cumulative deposited sulfate in MAECHAM5-HAM is highly spatially correlated with the ice core values (r = 0.8) and Greenland deposition is moderately correlated (r = 0.6). Hence MAECHAM5-HAM captures the spatial pattern of the deposited sulfate, especially in Antarctica, with greater deposition on the Antarctic Peninsula and lower deposition in East Antarctica, but the magnitude of the deposition is ~~a factor~~ ~3.7 times too large. Figure 7 shows the ice core values versus the model-simulated cumulative deposited sulfate. Correlation coefficients are less than ~0.5 for all models except MAECHAM5-HAM, although these models have lower mean biases. A figure where the simulated deposition

in MAECHAM5-HAM has been reduced by a factor of 3 to illustrate the well-captured spatial pattern of deposition is included in the supplementary information, Fig. S~~5~~4 (SOCOL-AER is also included in this figure). Both UM-UKCA and CESM1(WACCM), which are the higher resolution models, simulate a strong gradient in deposition between the low deposition over land and high deposition over sea and although they match the magnitude of the cumulative deposited sulfate more closely on the ice sheets than SOCOL-AER and MAECHAM5-HAM, they fail to produce the high values of cumulative deposited sulfate on the Antarctic Peninsula. ~~Although cumulative deposited sulfate in SOCOL is not highly correlated with the ice core values across the whole of the Antarctic ice sheet, this model does simulate higher cumulative deposited sulfate over the Antarctic Peninsula.~~

The polar deposition in UM-UKCA and CESM1(WACCM) more closely follows the models' precipitation field, with correlation coefficients between the polar (60º-90º) precipitation (averaged over the four years after the eruption) and polar cumulative deposited sulfate (in the four years after the eruption) of 0.7 and 0.9 respectively. Polar correlation coefficients for ~~SOCOL are very low at 0.2 in the Arctic and 0.1 in the Antarctic and for~~ MAECHAM5-HAM are 0.6 in the Arctic and 0.4 in the Antarctic. Figure 8 shows the zonal mean precipitation and zonal mean cumulative deposited sulfate in each model. The precipitation in the models is very similar, suggesting that the differences in model-simulated volcanic sulfate deposition arise from differences in the transport of the sulfate aerosol to the polar regions and/or the deposition schemes themselves. The ice sheet sulfate deposition in UM-UKCA remains dominated by dry deposition.

Figure 9 shows for each model the simulated area-mean volcanic sulfate deposition to the Antarctic and Greenland ice sheets over time, compared to two of the highest resolved and most precisely dated ice cores (~~for each model. Included for comparison are two of the highest resolved and most precisely dated ice cores (~~D4: McConnell et al., 2007; DIV: Sigl et al., 2014). We find that deposition to both ice sheets peaks first in MAECHAM5-HAM, followed by SOCOL-AER, then UM-UKCA and CESM1(WACCM). The main phase of deposition recorded in the two ice cores falls in time between that simulated by MAECHAM5-HAM and the other models. Compared to DIV and D4, the deposition to the ice sheets in MAECHAM5-HAM is too quick, but too slow in CESM1(WACCM) and UM-UKCA, although the timing is still relatively well captured for all models. The onset and duration of deposition to the ice sheets simulated by SOCOL-AER is most comparable to the two ice cores, suggesting a good representation of the volcanic aerosol evolution, but simulated deposition is too large (see Fig. 6) ~~and does not exhibit the variability in time apparent in the two ice core deposition time series. WACCM simulates the greatest variability in ice sheet deposition over time, with several peaks and troughs in the deposition time series. The ice sheet deposition simulated in WACCM is longer in duration than measured in the two ice cores (see also Fig. 10 for further time series of the model simulated ice sheet deposition).~~ The timing of the ice sheet deposition is further explored in Sect. 3.3.

3.3 ~~Relationship between hemispheric atmospheric sulfate burdens and sulfate deposited on ice sheets~~Ice sheet sulfate deposition and relationship to sulfate burdens

The temporal and spatial evolution of the volcanic sulfate deposition ultimately reflects the evolution of the atmospheric volcanic sulfate burdens. Figure 10 shows the zonal mean monthly-mean and global -total monthly-mean

atmospheric volcanic sulfate burdens for each model. MAECHAM5-HAM has the fastest conversion of SO_2 to sulfate aerosol, with the global peak sulfate burden occurring only 4 months after the eruption (Fig. 10b). This fast conversion is likely due to the lack of interactive OH in the model (Table 1), since OH does not become depleted by reaction with SO_2. In UM-UKCA and SOCOL-AER the peak global sulfate burden occurs 6-7 months after the eruption, but the global burden in SOCOL-AER decays more rapidly than in UM-UKCA. The global burden in CESM1(WACCM) peaks 12 months after the eruption and remains elevated for another 3.5 years (until the end of the simulation) and hence deposition in CESM1(WACCM) is longer-lived. The delay in full conversion of SO_2 to sulfate aerosol in these models is due to initial depletion of OH, which we explore further in Sect. 4.1.1. In all models there is stronger transport of the sulfate aerosol to the Southern Hemisphere (SH) compared to the Northern Hemisphere (NH) (Fig. 10a) due to the Brewer-Dobson Circulation, which is stronger in the winter hemisphere.

Here we consider the relationships between the NH sulfate burden vs. the SH sulfate burden, the cumulative sulfate deposited on Antarctica vs. Greenland, and most importantly, the relationship between the hemispheric sulfate burdens and the sulfate deposited on each ice sheet.

~~In order to derive the relationships between the hemispheric atmospheric sulfate burdens and the sulfate deposited on the ice sheets, we calculate the cumulative mean sulfate deposited on Antarctica and Greenland, calculated as the area-weighted mean deposited sulfate on each ice sheet once a land-sea mask has been applied. We calculate the ratio between Antarctica and Greenland deposited sulfate and the ratio between the Southern Hemisphere (SH) peak atmospheric sulfate burden and Northern Hemisphere (NH) peak atmospheric sulfate burden. Next we calculate the ratio between the hemispheric peak atmospheric sulfate burdens [Tg SO_4] (representing the total amount of sulfate aerosol that has formed) and the average amount of sulfate deposited on each ice sheet [kg SO_4 km^{-2}] for each of the models. We refer to this ratio as the Burden-To-Deposition factor (BTD), which is equivalent to the scaling factors derived by Gao et al. (2007).~~

In all models the SH peak atmospheric sulfate burden is greater than the NH peak atmospheric sulfate burden (Table 5) ~~due to seasonal preferential transport of the sulfate aerosol to the SH in all simulations (Fig. 5a)~~. Ratios between the SH and NH peak burdens are between 1.4 and 1.9. However, despite the larger SH burden, only MAECHAM5-HAM and SOCOL-AER simulate greater Antarctica mean deposited sulfate than in Greenland. CESM1(WACCM) has the smallest deposition ratio (0.3) with mean Greenland deposited sulfate of 109 kg SO_4 km^{-2} compared to 36 kg SO_4 km^{-2} in Antarctica. MAECHAM5-HAM and SOCOL-AER have the closest deposition ratio to that derived by Sigl et al. (2015), but with mean deposited sulfate ~4~~3~~-6 times larger than the Sigl et al. (2015) estimates~~ice core estimates~~. Conversely, and as simulated in UM-UKCA and CESM1(WACCM), the mean deposited sulfate deduced by Gao et al. (2007) for the eruption of Mt. Tambora showed slightly more mean deposited sulfate on Greenland relative to Antarctica, with a ratio of 0.9, although this ratio is still much larger than in UM-UKCA and CESM1(WACCM). In contrast to MAECHAM5-HAM and SOCOL-AER, where the deposition ratio mirrors the hemispheric split of the sulfate aerosol, deposition ratios for both UM-UKCA and CESM1(WACCM) are dissimilar to the ratio of the hemispheric peak burdens.

Figure 11 shows the simulated deposition to each ice sheet over time as in Fig. 9, except we compare to the hemispheric sulfate burdens.~~To explore possible mechanisms for the differences between the hemispheric atmospheric sulfate~~

burden ratios and deposition ratios, we have examined the temporal evolution of the ice sheet deposition compared with the sulfate burdens (Fig. 10). In MAECHAM5-HAM the NH sulfate burden peaks only 2 months after the eruption and the SH burden peaks 4 months after the eruption. The ice sheet deposition follows suit with the majority of deposition to Greenland occurring 8 months after the eruption and peak deposition to Antarctica occurring 14 months after the eruption. However, in the other models the SH burden peaks before the NH peak burden. The SH burden is greatest between 5-7 months after the eruption in these models and the NH burden peaks between 10-12 months after the eruption. In contrast to MAECHAM5-HAM, there are no clear separate peaks between the deposition to each ice sheet. In SOCOL-AER, both the majority of Greenland and Antarctic deposition occurs between 10-20 months after the eruption, which was found to compare well to the timing recorded in two ice cores (Fig. 9). In UM-UKCA and CESM1(WACCM) the main phase of deposition is longer lived and occurs between 10-30 months after the eruption. and there is more variability in the deposition timeseries compared to MAECHAM and SOCOL. Overall, decay of the atmospheric sulfate burden and deposition to the ice sheets in MAECHAM5-HAM is rapid, occurring within the first 20 months after the eruption, suggesting a fast transport of sulfate aerosol to the poles. We find that in the first ~8 months after the eruption the sulfate burden in UM-UKCA and CESM1(WACCM) is restricted between ~60 °S and ~40 °N (Fig. 105a), with strong gradients in sulfate burden across the SH polar vortex and NH subtropical edge, whereas more sulfate is transported to the poles in MAECHAM5-HAM and SOCOL-AER. Reasons for this are explored in Sect. 4.

The strength of the background climatological winds differs slightly in each model, with the strongest polar jets simulated in WACCM, and weakest in MAECHAM (maximum zonal mean climatological zonal winds are 52 m s^{-1} in WACCM and 32 m s^{-1} in MAECHAM (Fig. S6)). All models simulate a strengthening in the NH and SH polar zonal winds in the first year after the eruption; WACCM simulates the largest zonal mean anomalies and MAECHAM the weakest. All models simulate a clear strengthening of the southern polar vortex ~3-8 months after the eruption (defined here as the zonal mean zonal wind at the grid boxes closest to 60°S and 10 hPa), with a peak anomaly of 30 m s^{-1} simulated in WACCM (Fig. S7). Figure 11 shows the zonal mean zonal wind averaged over the first year after the eruption in each model. Peak zonal mean SH polar zonal wind is 58 m s^{-1} in WACCM, 45 m s^{-1} in UKCA, 44 m s^{-1} in SOCOL and 38 m s^{-1} in MAECHAM.

The polar vortex inhibits the transport of stratospheric aerosol to the poles (e.g. Schoeberl & Hartmann, 1991), with the stronger polar vortex in WACCM likely restricting some of the sulfate deposition in Antarctica compared to MAECHAM, which simulates the weakest winds. Differences in the polar vortex strength in each model may therefore contribute to the differences in simulated polar sulfate deposition but more work is needed to assess the dynamical response to the eruption of Tambora in these models. This remains beyond the scope of the current paper.

Next, we calculate the ratio between the hemispheric peak atmospheric sulfate burdens [Tg SO_4] (representing the total amount of sulfate aerosol that has formed) and the average amount of sulfate deposited on each ice sheet [kg SO_4 km^{-2}] for each of the models. We refer to this ratio as the Burden-To-Deposition factor (BTD), which is equivalent to the scaling factors derived by Gao et al. (2007) calculated from the observed relationship between the atmospheric burden and deposition of radioactive material after nuclear bomb tests. examine the range in BTD factors across the models. BTD factors are

important for estimating the hemispheric atmospheric sulfate burden and subsequently estimating the forcing of historical volcanic eruptions based on ice core sulfate deposition records (Sect. 1). We calculate the BTD factors for both NH (NH_BTD) and SH (SH_BTD), ~~, defined as the ratio between the hemispheric peak atmospheric sulfate burden [Tg SO_4] and the mean ice sheet deposited sulfate [kg SO_4 km^{-2}] ~~(Table 6). BTD factors for MAECHAM5-HAM are the same for both the NH and SH, as in Gao et al. (2007), but a factor 5 lower than Gao et al. (2007). CESM1(WACCM), SOCOL-AER and UM-UKCA simulate smaller NH_BTD than SH_BTD, but these factors are different in each model, with the NH_ BTD ranging from 0.22~~7~~ *10^9 km^{-2} (SOCOL-AER) to 0.97 *10^9 km^{-2} (UM-UKCA) and the SH_BTD from 0.3~~3~~4 *10^9 km^{-2} (SOCOL-AER) to 2.91 *10^9 km^{-2} (UM-UKCA). All models simulate a NH_BTD less than 1 *10^9 km^{-2}, but SH_BTD are less than 1 *10^9 km^{-2} for only MAECHAM5-HAM and SOCOL-AER due to the much larger Antarctic deposition in these models compared to UM-UKCA and CESM1(WACCM). The multi-model mean NH_BTD factor is 0.42~~4~~ *10^9 km^{-2} (~60% smaller than in Gao et al. (2007)) and multi-model mean SH_BTD factor is 1.27~~6~~ *10^9 km^{-2} (~30% greater than in Gao et al. (2007)). ~~These factors are ~40% different to the mean factors derived by Gao et al. (2007), which were calculated from the observed relationship between the atmospheric burden and deposition of radioactive material after nuclear bomb tests~~. We also find variability in the BTD factors across the individual ensemble members for each model arising due to internal variability, but ensemble spread is smaller than the inter-model spread.

We also test the sensitivity of the derived model BTD factors in Table 6 if we take polar deposition (60°-90°N/S) as opposed to ice sheet deposition, given that both UM-UKCA and CESM1(WACCM) simulate strong gradients in cumulative deposited sulfate across the land-sea boundary (Fig. 6). We find that the BTD factors remain similar for SOCOL-AER and MAECHAM5-HAM, but are reduced by up to a factor of 3 in UM-UKCA ~~and WACCM~~ due to the mean polar cumulative deposited sulfate being greater than the mean ice sheet cumulative deposited sulfate (Table S3~~4~~). In CESM1(WACCM) the SH_BTD is also reduced by a factor of 3, but the NH_BTD remains similar. Overall, the ~~This reduces the~~ spread in the BTD factors between the models decreases and results in a reduction of the multi-model mean NH_BTD factor from 0.42~~4~~ *10^9 km^{-2} to 0.28~~7~~ *10^9 km^{-2} and the SH_BTD from 1.27~~6~~ *10^9 km^{-2} to 0.54~~3~~ *10^9 km^{-2}.

4 Discussion

4.1 Differences in deposited sulfate

The spatial pattern and magnitude of deposited sulfate depends on the sources of atmospheric SO_2, the transport and mixing of the sulfate aerosol formed throughout the stratosphere and across the tropopause, and wet and dry deposition processes ~~as well as its deposition by either precipitation or gravitational sedimentation~~ (e.g. Hamill et al., 1997; Kremser et al., 2016). In the pre-industrial background state (no Tambora) (Fig. 1), all four models examined simulate similar patterns of sulfate deposition, with more sulfate deposited at the mid-latitudes and in oceans and near SO_2 sources such as continuously degassing volcanoes. In the polar regions, the models also simulate~~d~~ similar sulfate deposition (with the exception of SOCOL-AER) ~~and matched the magnitude and spatial pattern recorded~~ with reasonable comparison to~~in~~ ice core records (Fig. 2). This

indicates that the models ~~may be~~ are realistically~~accurately~~ simulating aspects of the formation and transport of background sulfate aerosol and subsequent deposition processes.~~, albeit with differences in the ratio of wet and dry deposition. UKCA was the only model to simulate a greater proportion of background deposition to the ice sheets occurring by dry processes. We find that all models simulate similar precipitation, which compares well to the annual snow accumulation recorded in ice cores and that the polar sulfate deposition is highly correlated with the polar precipitation.~~

However, under the volcanically-perturbed conditions (with Tambora), the simulated volcanic sulfate deposition differs between all models, with differences in timing, spatial pattern and magnitude. Compared to ice core records of cumulative deposited sulfate for 1815 Mt. Tambora, MAECHAM5-HAM and SOCOL-AER simulate much ~~too~~ higher deposition to polar ice sheets~~regions~~, which is ~3-5 times greater than the mean ice core-derived estimates by Gao et al. (2007) and Sigl et al. (2015). UM-UKCA and CESM1(WACCM) simulate deposition closer in magnitude to the ice core records although in UM-UKCA the sulfate deposited on both ice sheets is ~2 times too small compared to the mean ice core-derived estimates.~~,~~ In CESM1(WACCM) the sulfate deposited on Antarctica is slightly too small but ~2 times greater in Greenland compared to the mean ice core-derived estimates.~~too little deposition to polar regions. In UKCA and WACCM, the polar cumulative deposited sulfate more closely follows the spatial pattern of precipitation, although on the actual ice sheets, deposition in UKCA is dominated by dry deposition. In MAECHAM and SOCOL the polar cumulative deposited sulfate is more enhanced compared to the precipitation (Fig. 8). However, MAECHAM simulates the best spatial correlations with the ice cores, further indicating that precipitation is not the sole driver of the spatial variability of deposited volcanic sulfate.~~ Considering the models are more comparable in the background state it is likely that the inter-model differences in volcanic deposition are due to differences in the formation of the volcanic aerosol, the stratospheric transport of volcanic aerosol and stratosphere-troposphere exchange, since in the background state, most of the deposited sulfate is of tropospheric origin. These processes are discussed in the following sections.

4.1.1 Volcanic sulfate formation and transport

The timing and duration of sulfate deposition mirrors that of the atmospheric sulfate burdens. In MAECHAM5-HAM the atmospheric sulfate burden peaks sooner and decays more quickly than in the other models, and ice sheet deposition occurs more rapidly (within the first 2 years after the eruption). The atmospheric sulfate burden in CESM1(WACCM) is still elevated 4 years after the eruption, and hence the deposition signal is also longer-lived (Fig. 5). MAECHAM5-HAM is the only model that has prescribed OH (Table 1). OH may become depleted in dense volcanic clouds by reaction with SO_2, affecting the rate of sulfate aerosol formation (Bekki, 1995). The background stratospheric OH concentrations are similar between the models (Fig. S5) but in SOCOL-AER, UM-UKCA and CESM1(WACCM), in the first 2 months after the eruption, stratospheric tropical OH becomes depleted, with ensemble mean peak reductions of between 15-33% (Fig. S6). This reduces the rate of sulfate aerosol formation compared to MAECHAM5-HAM where the SO_2 will be more rapidly oxidised, and explains the later peaks in sulfate burdens in these models.

The rapid decay of the sulfate burden in MAECHAM5-HAM also indicates that this model could have faster accumulation of particles and stronger sedimentation compared to the other models. Although beyond the scope of this paper a more detailed examination of the aerosol microphysical processes and the size of the aerosol particles, on which sedimentation is dependent, will facilitate a greater understanding of some of the model differences identified here.

The high biases in cumulative deposited sulfate in MAECHAM5-HAM and SOCOL-AER compared to ice cores may be caused by a high bias in poleward aerosol transport (e.g. Stenke et al., 2013; Toohey et al., 2013). ~~As these models were able to match the ice-core fluxes in the background state, it stands that in the Tambora case, the poleward transport of aerosol in UKCA and WACCM is too weak or mid-latitude deposition is too strong. In MAECHAM and SOCOL the polar volcanic deposition is too strong likely because of a high bias in poleward aerosol transport (Toohey et al., 2013).~~ MAECHAM5-HAM and SOCOL-AER also have the lowest resolution of the four models (Table 1), which may contribute to the high deposition bias since stratospheric circulation and cross-tropopause transport is better represented in higher resolution models (e.g. Toohey et al., 2013). Gao et al. (2007), using the GISS ModelE, found that simulated deposited sulfate over the poles after the eruption of Mt. Tambora was a factor of 2~~two~~ too large, but that the spatial pattern of deposition recorded in ice cores was well captured. GISS ModelE had a much lower resolution than the models used here (4° x 5°) and a simplified scheme for stratospheric aerosol microphysics. SOCOL-AER and MAECHAM5-HAM have the same dynamical cores and therefore we expect transport to be similar in these models. Hence, the differences in simulated volcanic deposition between SOCOL-AER and MAECHAM5-HAM are likely due to aerosol growth and sedimentation, and the deposition schemes. In UM-UKCA and CESM1(WACCM) the poleward transport of volcanic aerosol may be too weak or mid-latitude deposition too strong.~~It is also likely that smoother topography in these lower resolution models will influence the spatial pattern of deposition.~~

~~The timing and duration of sulfate deposition mirrors that of the atmospheric sulfate burdens. In MAECHAM the atmospheric sulfate burden peaks sooner and decays more quickly than in the other models, and the ice sheet deposition occurs more rapidly (within the first 2 years after the eruption). The atmospheric sulfate burden in WACCM is still elevated 5 years after the eruption, and hence the deposition signal is also longer lived; deposition still occurs in 1819, four years after the eruption (Fig. 4).~~

~~The size of the volcanic sulfate aerosol particles will impact the simulated sulfate deposition by affecting the particle sedimentation rates. Most of the volcanic sulfate aerosol resides in the accumulation mode for UKCA and MAECHAM, but in the larger coarse mode for WACCM, and WACCM has the strongest mid-latitude volcanic deposition where cross-tropopause transport occurs (Fig. 3). Furthermore, MAECHAM is the only model that has prescribed OH as opposed to interactive OH. OH may become depleted in dense volcanic clouds by reaction with SO_2, affecting the rate of sulfate aerosol formation (Bekki, 1995). The lack of interactive OH in MAECHAM should result in more rapid oxidation of the volcanic SO_2 into sulfate aerosol compared to the other models. Although it is beyond the scope of this paper to address all of the model differences in sulfur chemistry and the formation of the volcanic aerosol, the chemistry and aerosol microphysics are ultimately important factors affecting deposition patterns. Further work on the aerosol size and growth processes will allow a more detailed understanding of some of the model differences identified here.~~

4.1.2 Dynamical effects

The direction and strength of the stratospheric winds impacts the transport of sulfate aerosol and hence where it is deposited. UM-UKCA, CESM1(WACCM) and SOCOL-AER have similarly defined QBOs with downward propagating easterly and westerly winds, with the eruption simulated in the easterly phase. MAECHAM5-HAM does not include a QBO and although stratospheric winds were easterly in the MAECHAM5-HAM simulations, we find that these winds are ~20 m s^{-1} weaker than the easterly phase winds in the other models (Fig. S7~~8~~). This may contribute~~explain~~ to the quicker transport and subsequent deposition to the poles in the MAECHAM5-HAM simulations~~, as strong easterly winds restrict the meridional movement of aerosol (Trepte and Hitchman, 1992).~~

In addition to mid-latitude tropopause folds, a further location of cross-tropopause transport of sulfate aerosol is the polar winter vortex (e.g. SPARC, 2006; Kremser et al., 2016). The polar vortex inhibits poleward transport (e.g. Shoeberl & Hartmann, 1991), and it has been suggested that variations in the strength of the polar vortex may modulate volcanic aerosol transport and deposition to polar ice sheets (Toohey et al, 2013). We find that the strength of the background climatological winds differs slightly across the models, with the strongest polar jets simulated in CESM1(WACCM), and the weakest in MAECHAM5-HAM (maximum zonal mean ~~climatological~~ zonal winds are 52 m s^{-1} in CESM1(WACCM) and 32 m s^{-1} in MAECHAM5-HAM (Fig. S8)). All models simulate a strengthening in the NH and SH polar zonal winds in the first year after the eruption; CESM1(WACCM) simulates the largest zonal mean anomalies and MAECHAM5-HAM the weakest. Figure 12 shows the zonal mean zonal wind averaged over the first year after the eruption in each model. Peak zonal mean SH polar zonal wind is 58 m s^{-1} in CESM1(WACCM), 46 m s^{-1} in SOCOL-AER, 45 m s^{-1} in UM-UKCA and 38 m s^{-1} in MAECHAM5-HAM. Inter-model differences in polar vortex strength may therefore contribute to differences in polar sulfate deposition. Following this hypothesis, the strong SH polar vortex simulated by CESM1(WACCM) may contribute to the lower deposited sulfate on Antarctica ~~large SH_BTD values for~~in this model, and likewise, the relatively weaker polar vortex in MAECHAM5-HAM may contribute to the greater deposited sulfate on Antarctica~~small SH_BTD values~~. On the other hand, UM-UKCA simulates average polar vortex winds, but the smallest deposited sulfate on Antarctica~~largest SH_BTD values~~. Therefore, it appears to be a combination of factors that drive the inter-model differences in simulated polar volcanic sulfate deposition.

4.1.3 Deposition schemes

Differences in the deposition schemes contribute to the inter-model differences. The simplified scheme in SOCOL-AER results in deposition following more closely the atmospheric distribution of sulfate. The differences between wet and dry deposition simulated across the models are due to the individual deposition scheme parameterizations. The implication of these differences in dictating the resulting total sulfate deposition remains uncertain. However, since inter-model differences in volcanic sulfate deposition patterns appear unrelated to differences between climatological wet and dry deposition patterns, the proportion of wet vs dry deposition is likely of secondary importance compared to differences between the models in

aerosol transport processes including sedimentation and stratosphere-troposphere exchange. Smoother topography in the lower resolution models will also influence the spatial pattern of deposition.

The realistic deposition of background sulfate suggests that the scavenging and deposition processes in the models are reasonably parameterized, and thus that inter-model differences in the Tambora case are more likely due to differences in stratospheric transport and stratosphere-troposphere exchange as described above. However, due to the higher sulfate burdens in the perturbed case, differences in deposition due to the schemes may become more pronounced. Scavenging and deposition parameterizations are highly uncertain, and the chance that such parameterizations become unrealistic under the large sulfate aerosol loadings associated with a Tambora eruption cannot be discounted, and should be explored in future work.

4.2 Implications for ~~the relationship between hemispheric atmospheric sulfate burdens and sulfate deposited on ice sheets~~model differences in simulated sulfate deposition

Using just four global aerosol models, we find large differences in the mean deposited sulfate on the Antarctic~~a~~ and Greenland ice sheets. The multi-model mean BTD factors, which relate~~s~~ the atmospheric sulfate burdens to the deposition at the ice sheets, differ from the estimates by Gao et al. (2007) by ~60% in the NH and ~30% in the SH~~ed on average by ~40% from the estimates by Gao et al. (2007)~~, although the Gao et al. (2007) estimates are within or close to the multi-model spread. We find that the multi-model spread in BTD factors is reduced when we take a polar cap average of deposition as opposed to the average ice sheet deposition because simulated deposition is more similar amongst the models when a greater area average is considered. Due to the large gradient between land and sea deposition simulated in CESM1(WACCM) and UM-UKCA, mean polar deposition does not represent the mean ice sheet deposition and BTD factors are therefore sensitive to the areas chosen to represent the polar/ice sheet deposition. This makes it difficult to estimate accurately the relationship between ice sheet deposition and sulfate aerosol loading in the models. We highlight this to emphasize caution when determining BTD factors in future modelling studies. Furthermore, although these simulations aimed to follow a common protocol, the injection setup did differ between models due to differences in the ways modelling groups interpret and simulate a volcanic injection (Table 2). This is a common problem in multi-model comparisons and makes it more difficult to isolate and attribute model differences.

We did not expect the models to be able to simulate the exact deposition at each ice core location given the large natural variability in local weather and snow patterns, which will be different in the models, uncertainties in estimating ice core volcanic sulfate deposition and in model inputs (e.g. magnitude and altitude of the volcanic sulfur emission), and fundamentally, that in the real world there was only one realization of weather. ~~Even if the models were perfect, they could not predict the actual deposition pattern as they could not simulate the actual weather that occurred~~. Regarding the model inputs, there are no direct~~good~~ observations of the injection altitude of SO_2 from the 1815 Mt. Tambora eruption, and often simulations are initiated with SO_2 injected at heights lower than the estimated injection altitude, to account for self-lofting as the aerosol forms. Inter-model uncertainty is also initiated as soon as models convert the same input emission to their grids. Here, simulations followed a~~the~~ common protocol, but it may be that to better simulate the eruption of Mt. Tambora, sensitivity

to injection height should be explored. Models may also contain inaccuracies due to uncertain physical representations and coarse resolution, and several ice core locations will be represented by the same model grid box. The differing resolutions between the models also means that the number of grid boxes and area that defines each ice sheet differs slightly between the models. Sulfate deposition fluxes have a large spatial variability due to differences in precipitation, the local synoptic conditions at the time of deposition, and post-deposition movement through wind (Fisher et al., 1985; Robock and Free, 1995; Wolff et al., 2005). Deposition fluxes can vary by orders of magnitude, even between ice cores that are located close to each other. For example, in a very low accumulation site in Antarctica (Dome C), Gautier et al. (2016) found that in five cores drilled 1 m apart, two cores missed the Tambora sulfate flux signal completely, which they attributed to snow drift and surface roughness. They reported that the mean flux between these five cores is uncertain by ~30%, highlighting the uncertainties in sulfate fluxes reported from single cores at such a low-accumulation site. This appears to be an extreme case, however, and the 1815 Mt. Tambora signal is clearly identifiable in all other Antarctic ice cores (Sigl et al., 2014).

Furthermore, the phase of the QBO at the time of the eruption is unknown, and here we have only used simulations where the SO_2 is emitted during the easterly phase. Toohey et al. (2013) also found that deposition to the poles varied as a function of SO_2 injection magnitude and season, and their simulations of a Tambora-like eruption showed greater deposition to Greenland than Antarctica. However, the volcanic eruptions in Toohey et al. (2013) were simulated in January and July, located at 15°N, which may explain the bias towards Greenland deposition. Further work is required on the influence of the QBO phase and injection height on the Antarctic and Greenland deposition efficiency.

Our multi-model mean NH_BTD factor is ~60% lower than previously derived (Gao et al., 2007), which if used to estimate the NH atmospheric sulfate burden of other historic tropical eruptions from their mean Greenland deposited sulfate, would result in lower burdens, and likely less volcanic cooling. Model-simulated NH cooling following large-magnitude volcanic eruptions has been overestimated in the past (e.g. Stoffel et al., 2015; Zanchettin et al., 2016). However, our multi-model mean SH_BTD factor is ~230% greater than Gao et al. (2007), which would result in a larger SH sulfate burden estimate. Applying our BTD factors (NH: 0.424 and SH: 1.276) to the mean Greenland and Antarctic deposited sulfate from the 1257 Samalas eruption (90 kg km^{-2} and 73 kg km^{-2}, respectively (Sigl et al., 2015)), results in a considerable hemispheric asymmetry in the estimated sulfate burdens. We calculate a SH burden ~2.5 times the NH burden, despite the eruption occurring in the tropics. This could result in further differences in aerosol optical depth and volcanic aerosol radiative forcing, and hemispheric asymmetry in atmospheric sulfate burdens has been shown to shift the inter-tropical convergence zone leading to precipitation anomalies (e.g. Haywood et al., 2013). However, this asymmetry seems unlikely, given that cooling in the SH after large tropical eruptions appears limited in proxy records (Neukom et al., 2014).

5 Conclusions

We have analysed the volcanic sulfate deposition in model simulations of the 1815 eruption of Mt. Tambora using four state-of-the-art global aerosol models (CESM1(WACCM), MAECHAM5-HAM, SOCOL-AER and UM-UKCA) and

compared the simulated deposited sulfate to a comprehensive array of ice core records. We have also investigated the simulated sulfate deposition ~~to the poles~~ under ~~both~~ pre-industrial background conditions (without the eruption of Mt. ~~(no~~ Tambora)~~. and volcanically perturbed (with Tambora) conditions~~. Although the~~all~~ models simulate~~d~~ relatively similar background sulfate deposition fluxes, ~~which compare well to polar ice core records,~~ the models differ substantially in their simulation of the Mt. Tambora volcanic sulfate deposition, with differences in the timing, spatial pattern and magnitude. CESM1(WACCM) and UM-UKCA simulate similar~~comparable~~ deposition patterns, with the majority of sulfate deposited at the mid-latitude storm belts. On the ice sheets, UM-UKCA simulates too little deposited sulfate compared to mean ice core-derived estimates (~2 times too small). In CESM1(WACCM) deposited sulfate on Antarctica is also slightly too small, but deposited sulfate on Greenland is ~2 times too large compared to the mean ice core-derived estimate.~~ but underestimate the volcanic sulfate deposited on the ice sheets, with strong gradients in the deposited sulfate between land and sea.~~ In MAECHAM5-HAM and SOCOL-AER, the sulfate is deposited further across the globe and these models simulate ~3-5 times too much deposited sulfate on the ice sheets compared to mean ice core-derived estimates. However, MAECHAM5-HAM is the only model to capture the ~~simulates a~~ spatial pattern of deposited sulfate compared~~similar~~ to ice cores, especially in Antarctica.~~ but the magnitude is ~3 times too large.~~. Ice sheet deposited sulfate in SOCOL is also too large compared to ice cores.~~

Because the background deposition is more comparable between the models than in the perturbed case, differences in the volcanic sulfate deposition are likely due to differences in the formation of the volcanic aerosol, the stratospheric transport of volcanic aerosol and stratosphere-troposphere exchange. In addition, differences in deposition due the deposition schemes may become more pronounced under the higher sulfate loading. ~~The models simulate similar precipitation rates and patterns, which compare well to the annual snow accumulation recorded in ice cores. Hence, differences in the model-simulated volcanic sulfate deposition likely arise due to differences in the simulated atmospheric sulfate aerosol burdens and aerosol transport. In WACCM and UKCA the volcanic sulfate deposition is more related to the precipitation fields, with more sulfate deposited at the mid-latitude storm belts, than in MAECHAM and SOCOL. In MAECHAM and SOCOL, enhanced transport of sulfate aerosol to the poles dictates the resulting high ice sheet sulfate deposition.~~ We suggest that ~~a combination of~~ differences in model resolution, modelled stratospheric winds, aerosol microphysics and sedimentation and deposition schemes ~~in each model~~ have all contributed to the range in model-simulated~~resulting differences in~~ volcanic sulfate deposition.

We have calculated Burden-To-Deposition (BTD) factors between the mean deposited sulfate on each ice sheet and the corresponding hemispheric peak atmospheric sulfate burden for the Mt. Tambora simulations ~~(Table 6)~~. The BTD factors differ by up to a factor of 15 between the models. The multi-model mean BTD factors also differ ~~on average by ~40%~~ to BTD factors currently used to deduce historical volcanic forcing (e.g. Gao et al., 2007; 2008; Sigl et al., 2015). Our range in derived BTD factors highlights uncertainties in the relationship between atmospheric sulfate burden and ice sheet deposited sulfate as simulated by models.~~for this eruption.~~

~~We find very large differences in model-simulated volcanic sulfate deposition from the 1815 eruption of Mt. Tambora and in the relationship between atmospheric sulfate burden and ice sheet deposited sulfate~~. Given that GISS ModelE (Gao et al., 2007) did as good a job at simulating the deposited sulfate from this eruption as these newer, higher-resolution models,

which also have more sophisticated treatments of gas-to-aerosol conversion, and the fact that the four models used here simulate very different sulfate deposition, it remains an open research question as to the optimal model configuration for this problem. A detailed analysis of the differences in sulfur chemistry and the aerosol formation and transport in each model will further aid in the interpretation of these results. Dedicated multi-model comparison projects with process-oriented comparisons, such as the Interactive Stratospheric Aerosol Modelling Intercomparison Project (ISA-MIP) (Timmreck et al., 2016), will be imperative to disentangling the reasons for model differences. Using idealized prescribed aerosol forcings such as Easy Volcanic Aerosol (Toohey et al., 2016) in future VolMIP experiments, will also provide the opportunity to better understand model diversity. ~~and to advance our understanding of the climate response to large volcanic eruptions.~~ Simulations of other large-magnitude volcanic eruptions will also enable the calculation of additional multi-model BTD factors, which will aid in the calculation of historic volcanic forcing.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the multi-journal special issue "The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP)". It does not belong to a conference.

Acknowledgments. Lauren Marshall is supported by the Natural Environment Research Council (NERC), UK through the Leeds-York NERC Doctoral Training Partnership. Anja Schmidt was funded by an Academic Research Fellowship from the University of Leeds. Alan Robock is supported by U.S. National Science Foundation grant AGS-1430051. Claudia Timmreck received funding from the German Federal Ministry of Education and Research (BMBF), research program "MiKliP"(FKZ: 01LP1517B). Ulrike Niemeier, Claudia Timmreck and Slimane Bekki acknowledge support ~~received funding~~ from the European Union FP7 project "STRATOCLIM" (FP7-ENV.2013.6.1-2; Project 603557). MAECHAM5-HAM simulations were performed at the German climate Computer Centre (DKRZ). Fiona Tummon was funded by Swiss National Science Foundation grant 20F121_138017. This study benefited from the support of the Labex L-IPSL which is funded by the ANR (Grant #ANR-10-LABX-0018). Matthew Toohey acknowledges support by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority programme "Antarctic Research with comparative investigations in Arctic ice areas" through grant TO 967/1-1. Eugene Rozanov and Timofei Sukhodolov acknowledge support from the Swiss National Science Foundation under grant 200021_169241 (VEC). James Pope was funded by NERC grant NEK/K012150/1. The National Center for Atmospheric Research is funded by the National Science Foundation. We thank Chaochao Gao and an anonymous reviewer for their comments and insight on the manuscript, which greatly helped to improve the paper.

References

Bekki, S.: Oxidation of volcanic SO2: A sink for stratospheric OH and H2O, Geophys. Res. Lett., 22(8), 913–916, doi:10.1029/95GL00534, 1995.

Bellouin, N. , Mann, G. W., Woodhouse, M. T., Johnson, C., Carslaw, K. S., and Dalvi, M.: Impact of the modal aerosol scheme GLOMAP-mode on aerosol forcing in the Hadley Centre Global Environmental Model, Atmos. Chem. Phys., 13, 3027–3044, 2013.

Brooke, J. S. A, Feng, W., Carrillo-Sánchez, J. D. , Mann, G. W., James, A. D., Bardeen, C. G., and Plane, J. M. C.: Meteoric Smoke Deposition in the Polar Regions: A Comparison of Measurements With Global Atmospheric Models, J. Geophys. Res.: Atmospheres, 122, 11,112–11,130, doi:10.1002/2017JD027143, 2017.

Clausen, H. B., and Hammer, C. U.: The laki and tambora eruptions as revealed in greenland ice cores from 11 locations, Ann. Glaciol., 10, 16-22, 1988.

Cole-Dai, J., Mosley-Thompson, E., and Thompson, L. G.: Annually resolved southern hemisphere volcanic history from two Antarctic ice cores, J. Geophys. Res., 102(D14), 16761–16771, doi:10.1029/97JD01394, 1997.

Cole-Dai, J., Mosley-Thompson, E., Wight, S. P., and Thompson, L. G.: A 4100-year record of explosive volcanism from an East Antarctica ice core, J. Geophys. Res., 105(D19), 24431–24441, doi:10.1029/2000JD900254, 2000.

Cole-Dai, J., Ferris, D. G., Lanciki, A. L., Savarino, J., Thiemens, M. H., and McConnell, J. R.: Two likely stratospheric volcanic eruptions in the 1450s C.E. found in a bipolar, subannually dated 800 year ice core record, J. Geophys. Res., 118(14), 7459–7466, doi:10.1002/jgrd.50587, 2013.

Crowley, T. J.: Causes of Climate Change Over the Past 1000 Years, Science, 289(5477), 270-277, 2000.

Crowley, T. J. and Unterman, M. B.: Technical details concerning development of a 1200 yr proxy index for global volcanism, Earth Syst. Sci. Data, 5, 187–197, doi:10.5194/essd-5-187-2013, 2013.

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, 11221–11246, doi:10.5194/acp-14-11221-2014, 2014.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016.

Ferris, D. G., Cole-Dai, J., Reyes, A. R., and Budner, D. M.: South Pole ice core record of explosive volcanic eruptions in the first and second millennia A.D. and evidence of a large eruption in the tropics around 535 A.D., J. Geophys. Res., 116, D17308, doi:10.1029/2011JD015916, 2011.

Fisher, D. A., Reeh, N., and Clausen, H. B.: Stratigraphic noise in time series derived from ice cores, Ann. Glaciol., 7, 76–83, 1985.

Gao, C., Robock, A., Self, S., Witter, J.B., Steffenson, J.P., Clausen, H.B., Siggaard- Andersen, M.L., Johnsen, S., Mayewski, P.A. and Ammann, C.: The 1452 or 1453 A.D. Kuwae eruption signal derived from multiple ice core records: Greatest volcanic sulfate event of the past 700 years, J. Geophys. Res., 111, D12107, doi:10.1029/2005JD006710, 2006.

Gao, C., Oman, L., Robock, A., and Stenchikov, G. L.: Atmospheric volcanic loading derived from bipolar ice cores: Accounting for the spatial distribution of volcanic deposition, J. Geophys. Res, 112, D09109, doi:10.1029/2006JD007461, 2007.

Gao, C., Robock, A., and Ammann, C.: Volcanic forcing of climate over the past 1500 years: an improved ice core-based index for climate models, J. Geophys. Res., 113, D23111, doi:10.1029/2008JD010239, 2008.

Gautier, E., Savarino, J., Erbland, J., Lanciki, A., and Possenti, P.: Variability of sulfate signal in ice core records based on five replicate cores, Clim. Past, 12, 103–113, doi:10.5194/cp-12-103-2016, 2016.

Hamill, P., Jensen, E. J., Russell, P. B., and Bauman, J. J.: The Life Cycle of Stratospheric Aerosol Particles, Bulletin of the American Meteorological Society, 78(7), 1395–1410, 1997.

Haywood, J. M., Jones, A., Bellouin, N., and Stephenson, D.: Asymmetric forcing from stratospheric aerosols impacts Sahelian rainfall, Nat. Clim. Change, 3, 660–665, doi:10.1038/nclimate1857, 2013.

Herzog, M. and Graf, H.-F.: Applying the three-dimensional model ATHAM to volcanic plumes: dynamic of large co-ignimbrite eruptions and associated injection heights for volcanic gases, Geophys. Res. Lett., 37, L19807, doi:10.1029/2010GL044986, 2010.

Kipling, Z., Stier, P., Schwarz, J. P., Perring, A. E., Spackman, J. R., Mann, G. W., Johnson, C. E., and Telford, P. J.:

Constraints on aerosol processes in climate models from vertically-resolved aircraft observations of black carbon, Atmos. Chem. Phys., 13, 5969-5986, https://doi.org/10.5194/acp-13-5969-2013, 2013.

Kremser, S., Thomason, L. W., Hobe, M., Hermann, M., Deshler, T., Timmreck, C., Toohey, M., Stenke, A., Schwarz, J. P., Weigel, R., Fueglistaler, S., Prata, F. J., Vernier, J.-P., Schlager, H., Barnes, J. E., Antuña-Marrero, J.-C., Fairlie, D., Palm, M., Mahieu, E., Notholt, J., Rex, M., Bingen, C., Vanhellemont, F., Bourassa, A., Plane, J. M. C., Klocke, D., Carn, S. A., Clarisse, L., Trickl, T., Neely, R., James, A. D., Rieger, L., Wilson, J. C. and Meland, B.: Stratospheric aerosol–Observations, processes, and impact on climate, Rev. Geophys., 54(2), 278–335, doi:10.1002/2015RG000511, 2016.

Lamarque, J.-F., Emmons, L. K., Hess, P. G., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C. L., Holland, E. A., Lauritzen, P. H., Neu, J., Orlando, J. J., Rasch, P. J., and Tyndall, G. K.: CAM-chem: description and evaluation of interactive atmospheric chemistry in the Community Earth System Model, Geosci. Model Dev., 5, 369-411, https://doi.org/10.5194/gmd-5-369-2012, 2012.

Lamarque, J.-F., Dentener, F., McConnell, J., Ro, C.-U., Shaw, M., Vet, R., Bergmann, D., Cameron-Smith, P., Dalsoren, S., Doherty, R., Faluvegi, G., Ghan, S. J., Josse, B., Lee, Y. H., MacKenzie, I. A., Plummer, D., Shindell, D. T., Skeie, R. B., Stevenson, D. S., Strode, S., Zeng, G., Curran, M., Dahl-Jensen, D., Das, S., Fritzsche, D., and Nolan, M.: Multi-model mean nitrogen and sulfur deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): evaluation of historical and projected future changes, Atmos. Chem. Phys., 13, 7997-8018, doi:10.5194/acp-13-7997-2013, 2013.

Langway, C. C., Osada, K., Clausen, H. B., Hammer, C. U., and Shoji, H.: A 10-century comparison of prominent bipolar volcanic events in ice cores, J. Geophys. Res., 100(D8), 16241–16247, doi:10.1029/95JD01175, 1995.

Legrand, M., & Mayewski, P.: Glaciochemistry of polar ice cores: A review, Rev. Geophys., 35(3), 219–243, doi:10.1029/96RG03527, 1997.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J.-F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G., and Mitchell, D.: Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, Geosci. Model Dev., 5, 709-739, https://doi.org/10.5194/gmd-5-709-2012, 2012.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the

UKCA composition-climate model, Geosci. Model Dev., 3, 519-551, https://doi.org/10.5194/gmd-3-519-2010, 2010.

McCormick, M. P. and Veiga, R. E.: SAGE II measurements of early Pinatubo aerosols, Geophys. Res. Lett., 19, 155–158, doi:10.1029/91GL02790, 1992.

Mills, M. J., Schmidt, A., Easter, R., Solomon, S., Kinnison, D. E., Ghan, S. J., Neely III, 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, 2332–2348, doi:10.1002/2015JD024290, 2016.

Mills, M. J., Richter, J. H., Tilmes, S., Kravitz, B., MacMartin, D. G., Glanville, A. A., Tribbia, J. J., Lamarque, J.-F., Vitt, F., Schmidt, A., Gettelman, A., Hannay, C., Bacmeister, J. T., and Kinnison, D. E.: Radiative and chemical response to interactive stratospheric sulfate aerosols in fully coupled CESM1(WACCM). J. Geophys. Res.-Atmos., 122, https://doi.org/10.1002/2017JD027006, 2017

Niemeier, U., Timmreck, C., Graf, H.-F., Kinne, S., Rast, S., and Self, S.: Initial fate of fine ash and sulfur from large volcanic eruptions, Atmos. Chem. Phys., 9, 9043–9057, doi:10.5194/acp- 9-9043-2009, 2009.

Oppenheimer, C.: Climatic, environmental and human consequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815, Prog. Phys. Geog., 27, 230–259, 2003.

Pinto, J. P., Toon, O. B., and Turco, R. P.: Self-limiting physical and chemical effects in volcanic eruption clouds, J. Geophys. Res., 94, 11165–11174, doi:10.1029/JD094iD08p11165, 1989.

Raible, C. C., Brönnimann, S., Auchmann, R., Brohan, P., Frölicher, T. L., Graf, H.-F., Jones, P., Luterbacher, J., Muthers, S., Neukom, R., Robock, A., Self, S., Sudrajat, A., Timmreck, C., and Wegmann, M.: Tambora 1815 as a test case for high impact volcanic eruptions: Earth system effects, WIREs Clim Change, 7, 569–589, doi:10.1002/wcc.407, 2016.

Read, W. G., Froidevaux, L., and Waters, J. W.: Microwave limb sounder measurement of stratospheric SO_2 from the Mount Pinatubo volcano, Geophys. Res. Lett., 20, 1299–1302, doi:10.1029/93GL00831, 1993.

Robock, A.: Volcanic eruptions and climate, Rev. Geophys., 38(2), 191–219, 2000.

Robock, A., and Free, M. P.: Ice cores as an index of global volcanism from 1850 to the present, J. Geophys. Res., 100(D6), 11549-11567, doi:10.1029/95JD00825, 1995.

Schoeberl, M. R., and Hartmann, D. L.: The Dynamics of the Stratospheric Polar Vortex and Its Relation to Springtime Ozone Depletions, Science, 251(4989), 46–52, doi:10.1126/science.251.4989.46, 1991.

Self, S., Gertisser, R., Thordarson, T., Rampino, M. R., and Wolff, J. A.: Magma volume, volatile emissions, and stratospheric aerosols from the 1815 eruption of Tambora, Geophys. Res. Lett. 31, L20608, doi:10.1029/2004GL020925, 2004.

Sheng, J.-X., Weisenstein, D. K., Luo, B.-P., Rozanov, E., Stenke, A., Anet, J., Bingemer, H., and Peter, T.: Global atmospheric sulfur budget under volcanically quiescent conditions: aerosol– chemistry–climate model predictions and validation, J. Geophys. Res.-Atmos., 120, 256–276, doi:10.1002/2014JD021985, 2015a.

Sheng, J.-X., Weisenstein, D. K., Luo, B.-P., Rozanov, E., Arfeuille, F., and Peter, T.: A perturbed parameter model ensemble to investigate Mt. Pinatubo's 1991 initial sulfur mass emission, Atmos. Chem. Phys., 15, 11501-11512, https://doi.org/10.5194/acp-15-11501-2015, 2015b.

Sigl, M., McConnell, J. R., Layman, L., Maselli, O., McGwire, K., Pasteris, D., Dahl-Jensen, D., Steffensen, J. P., Vinther, B., Edwards, R., Mulvaney, R. and Kipfstuhl, S.: A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000 years, J. Geophys. Res.-Atmos., 118(3), 1151–1169, doi: 10.1029/2012JD018603, 2013.

Sigl, M., Mcconnell, J. R., Toohey, M., Curran, M., Das, S. B., Edwards, R., Isaksson, E., Kawamura, K., Kipfstuhl, S., Krüger, K., Layman, L., Maselli, O. J. Motizuki, Y., Motoyama, H., and Pasteris, D. R.: Insights from Antarctica on volcanic forcing during the Common Era, Nat. Clim. Change, 4, 693–697, doi:10.1038/nclimate2293, 2014.

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O. J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., Salzer, M., Schüpbach, S., Steffensen, J. P., Vinther, B. M., and Woodruff, T. E.: Timing and climate forcing of volcanic eruptions for the past 2,500 years, Nature, 523, 543–549, doi:10.1038/nature14565, 2015.

SPARC: SPARC Assessment of Stratospheric Aerosol Properties (ASAP). Thomason, L., and Peter, T., (Eds.), SPARC Report No. 4, WCRP-124, WMO/TD - No. 1295, 2006.

Stenke, A., Schraner, M., Rozanov, E., Egorova, T., Luo, B., and Peter, T.: The SOCOL version 3.0 chemistry–climate model: description, evaluation, and implications from an advanced transport algorithm, Geosci. Model Dev., 6, 1407-1427, https://doi.org/10.5194/gmd-6-1407-2013, 2013.

Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A. and Petzold, A.: The aerosol-climate model ECHAM5-HAM, Atmos. Chem. Phys., 5(4), 1125–1156, doi:10.5194/acp-5-1125-2005, 2005.

Stoffel, M., Khodri, M., Corona, C., Guillet, S., Poulain, V., Bekki, S., Guiot, J., Luckman, B. H., Oppenheimer, C., Lebas, N., Beniston, M. and Masson-Delmotte, V.: Estimates of volcanic induced cooling in the Northern Hemisphere over the past 1,500 years, Nat. Geosci., 8, 784–788, doi:10.1038/ngeo2526, 2015.

Timmreck, C., Lorenz, S. J., Crowley, T. J., Kinne, S., Raddatz, T. J., Thomas, M. A., and Jungclaus, J. H.: Limited temperature response to the very large AD 1258 volcanic eruption, Geophys. Res. Lett., 36(21), doi:10.1029/2009GL040083, 2009.

Timmreck, C., Mann, G.W., Aquila, V., Brühl, C., Chin, M., Dhomse, S.S., English, J.M., Hommel, R., Lee, L.A., Mills, M.J., Neely, R., Schmidt, A., Sheng, J.X., Toohey, M., and Weisenstein, D.: ISA-MIP: A coordinated intercomparison of Interactive Stratospheric Aerosol models, Geophys. Res. Abstr., 18, EGU2016-13766, EGU General Assembly 2016

Toohey, M., Krüger, K., and Timmreck, C.: Volcanic sulfate deposition to Greenland and Antarctica: A modeling sensitivity study, J. Geophys. Res.-Atmos., 118(10), 4788–4800, doi: 10.1002/jgrd.50428, 2013.

Toohey, M., Stevens, B., Schmidt, H., and Timmreck, C.: Easy Volcanic Aerosol (EVA v1.0): an idealized forcing generator for climate simulations, Geosci. Model Dev., 9, 4049-4070, doi:10.5194/gmd-9-4049-2016, 2016.

Toohey, M. and Sigl, M.: Volcanic stratospheric sulfur injections and aerosol optical depth from 500 BCE to 1900 CE, Earth Syst. Sci. Data, 9(2), 809–831, doi:10.5194/essd-9-809-2017, 2017. ~~Toohey, M. and Sigl, M.: Volcanic stratospheric sulphur injections and aerosol optical depth from 500 BCE to 1900 CE, Earth Syst. Sci. Data Discuss., in review, 2017~~

Trepte, C. R., and Hitchman, M. H.: Tropical stratospheric circulation deduced from satellite aerosol data, Nature, 355(6361), 626–628, 1992.

Trepte, C. R., Veiga, R. E., and McCormick, M. P.: The poleward dispersal of Mount Pinatubo volcanic aerosol, J. Geophys. Res.-Atmos., 98(D10), 18563–18573, doi:10.1029/93JD01362, 1993.

Wolff, E. W.: Chemical signals of past climate and environment from polar ice cores and firn air, Chem. Soc. Rev. Chem. Soc. Rev, 41(41), 6247–6258, doi:10.1039/c2cs35227c, 2012.

Wolff, E. W., Cook, E., Barnes, P. R. F., and Mulvaney, R.: Signal variability in replicate ice cores, J. Glaciol., 51(174),

462–468, doi:10.3189/172756505781829197, 2005.

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., Hegerl, G., Robock, A., Pausata, F. S. R., Ball, W. T., Bauer, S. E., Bekki, S., Dhomse, S. S., LeGrande, A. N., Mann, G. W., Marshall, L., Mills, M., Marchand, M., Niemeier, U., Poulain, V., Rozanov, E., Rubino, A., Stenke, A., Tsigaridis, K., and Tummon, F.: The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): experimental design and forcing input data for CMIP6, Geosci. Model Dev., 9, 2701-2719, doi:10.5194/gmd-9-2701-2016, 2016.

Zielinski, G. A.: Stratospheric loading and optical depth estimates of explosive volcanism over the last 2100 years derived from the Greenland Ice Sheet Project 2 ice core, J. Geophys. Res., 100(D10), 20937, doi:10.1029/95JD01751, 1995.

Zielinski, G. A., Mayewski, P. A., Meeker, L. D., Whitlow, S., and Twickler, M. S.: A 110,000-Yr Record of Explosive Volcanism from the GISP2 (Greenland) Ice Core, Quaternary Res., 45(2), 109–118, doi:10.1006/qres.1996.0013, 1996.

Zielinski, G.A., Mayewski, P.A., Meeker, L.D., Grönvold, K., Germani, M.S., Whitlow, S., Twickler, M.S. and Taylor, K.: Volcanic aerosol records and tephrochronology of the Summit, Greenland, ice cores, J. Geophys. Res.-Oceans, 102(C12), 26625-26640, doi:10.1029/96JC03547, 1997.

Figures

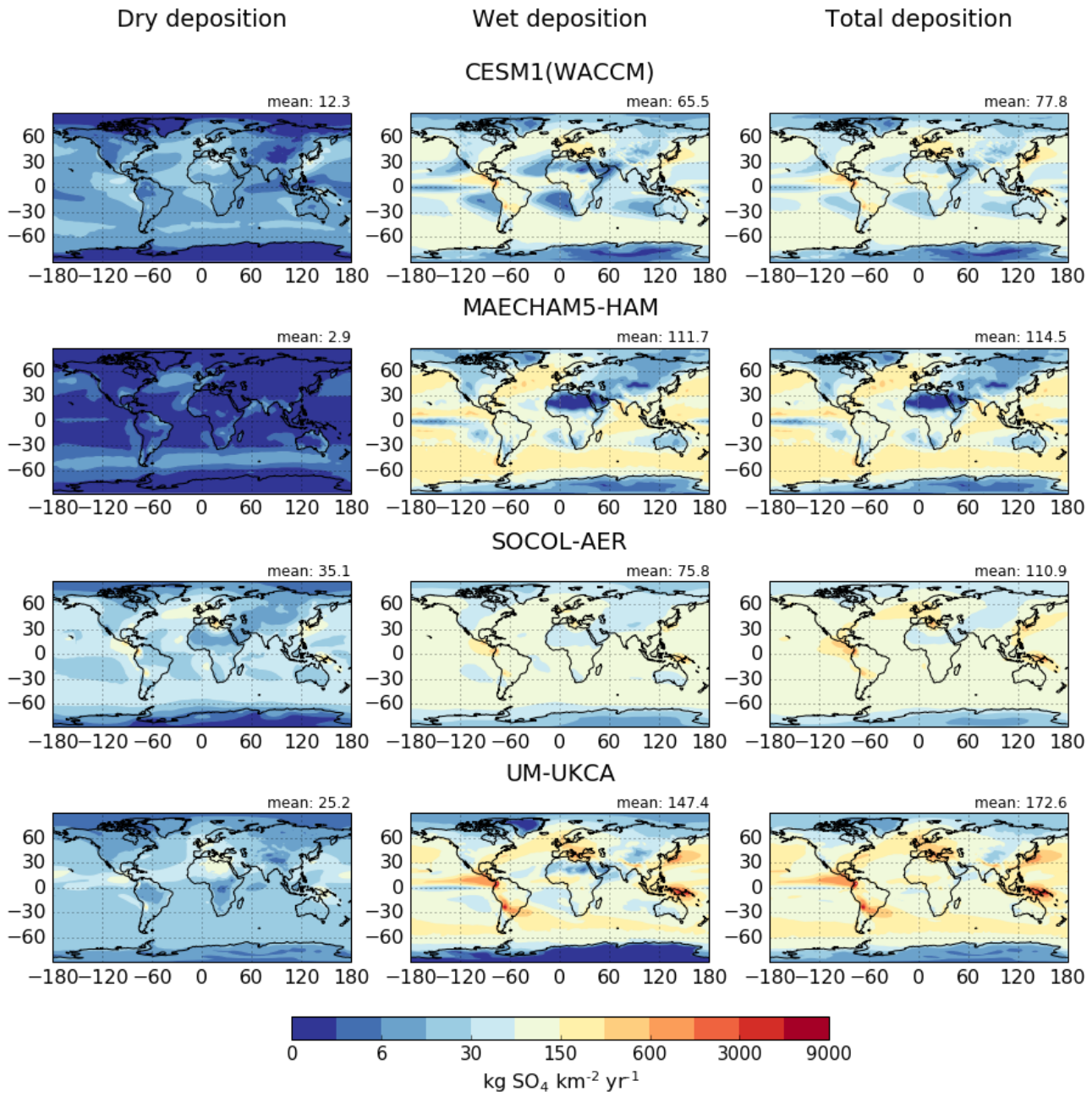

Figure 1: Pre-industrial background annual ~~mean~~ dry, wet and total (wet + dry) sulfate deposition fluxes [kg SO_4 km^{-2} yr^{-1}] (left to right) for each model (top to bottom). ~~UKCA is an average of 4 years; WACCM, MAECHAM and SOCOL WACCM are averages of 5 years and SOCOL is an average of 1 year due to the length of available simulation.~~ The value shown in the top right-hand corner of each plot refers to the global mean sulfate deposition flux. Background fluxes are averages of the annual deposition from 5 control simulations each with 4 years of data for UM-UKCA, 3 controls each with 5 years of data for CESM1(WACCM), and 1 control with 5 years of data for MAECHAM5-HAM and SOCOL-AER.

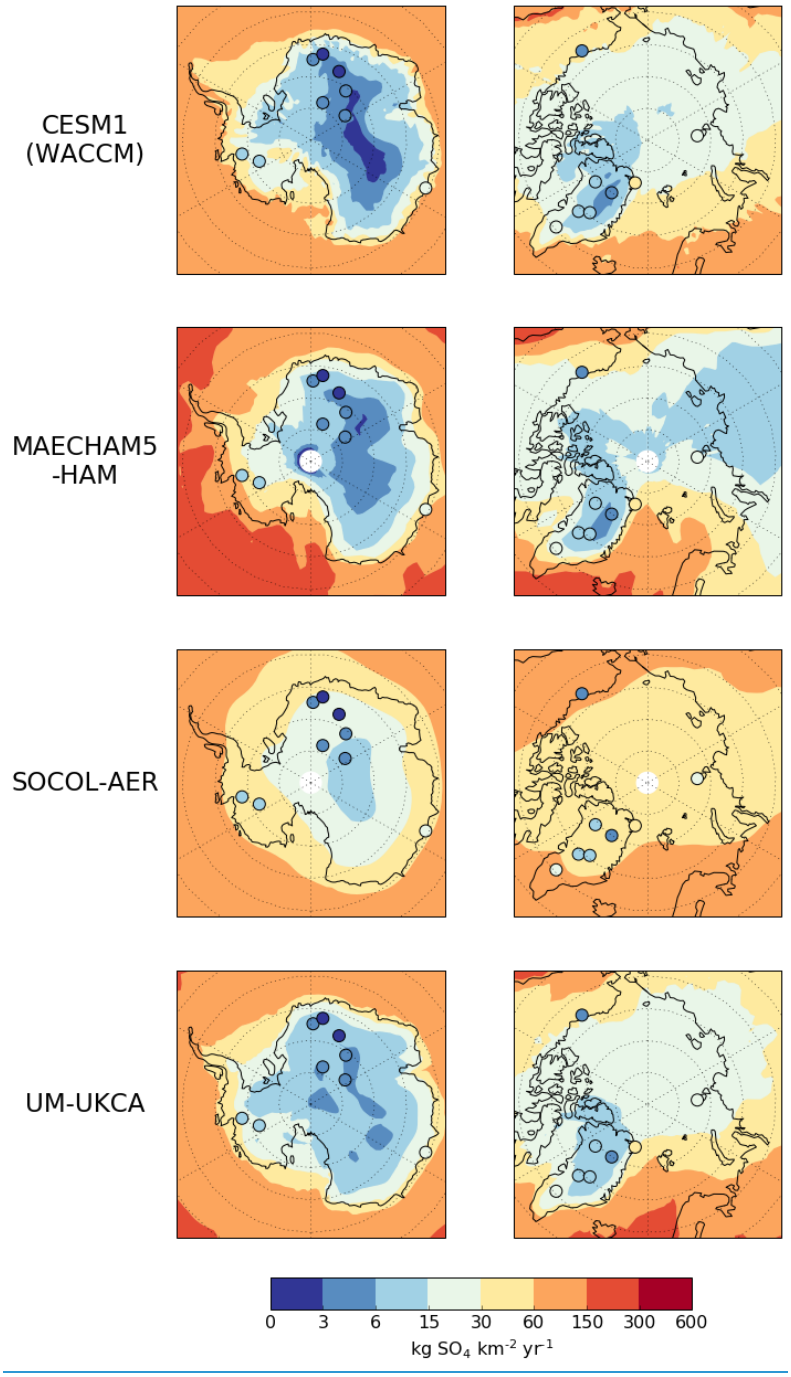

Figure 2: ~~Annual mean~~ Total (wet + dry) sulfate deposition fluxes [kg SO_4 km^{-2} yr^{-1}] for Antarctica (left) and the Arctic (right) for the pre-industrial background from the control simulations (shading) compared to pre-industrial ice core sulfate fluxes (filled circles), averaged for 1850 to 1860.

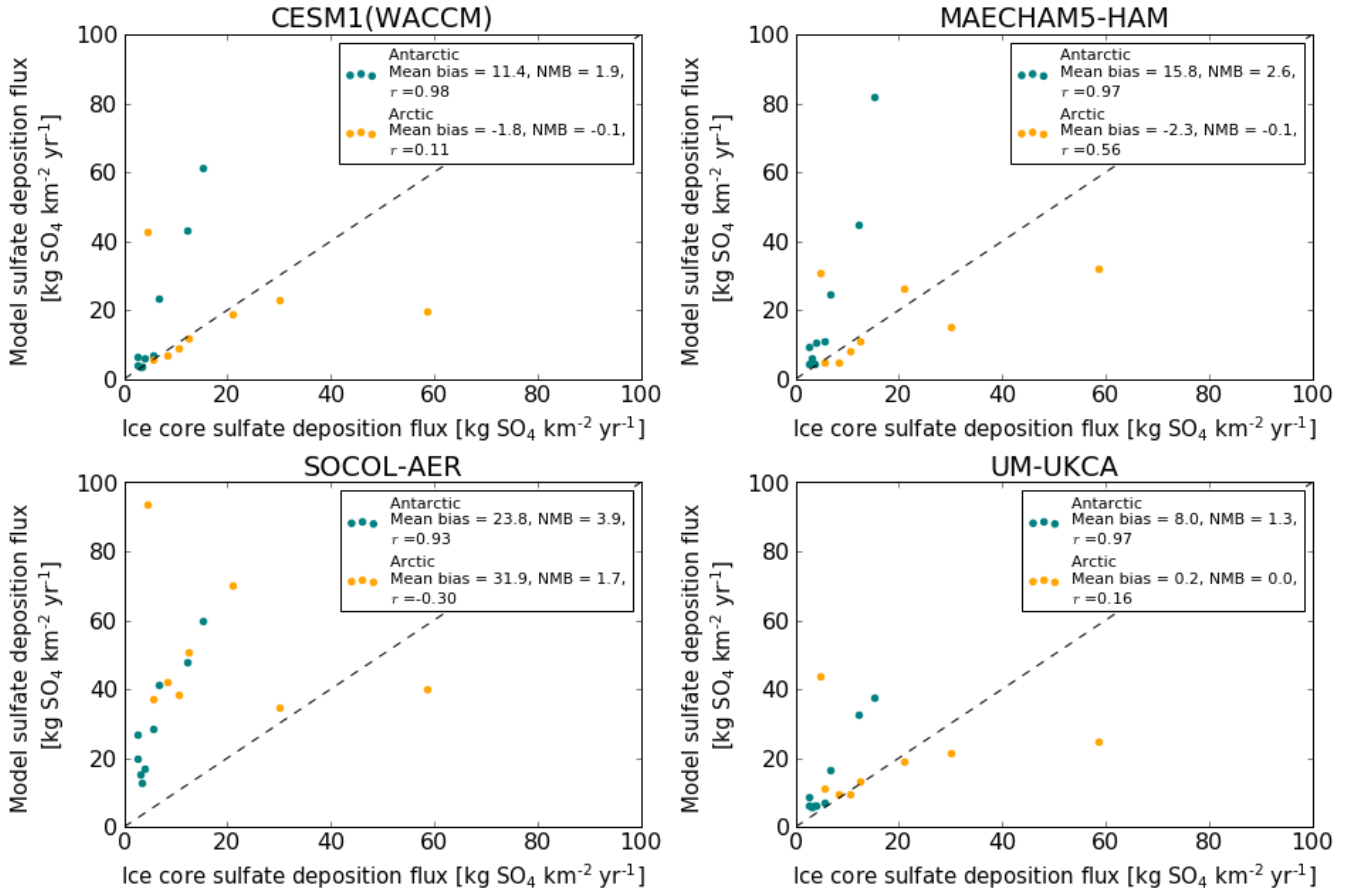

Figure 3: Scatter plots of pre-industrial background ice core sulfate deposition fluxes vs. simulated pre-industrial sulfate fluxes [kg SO_4 km^{-2} yr^{-1}] in the Antarctic (teal points) and in the Arctic (orange points) for each model. Simulated values represent the grid box value where each ice core is located. The dashed line marks the 1:1 line. Included in the legends are the mean bias, Normalized Mean Bias (NMB) and the correlation coefficient (*r*) for the Antarctic and Arctic.

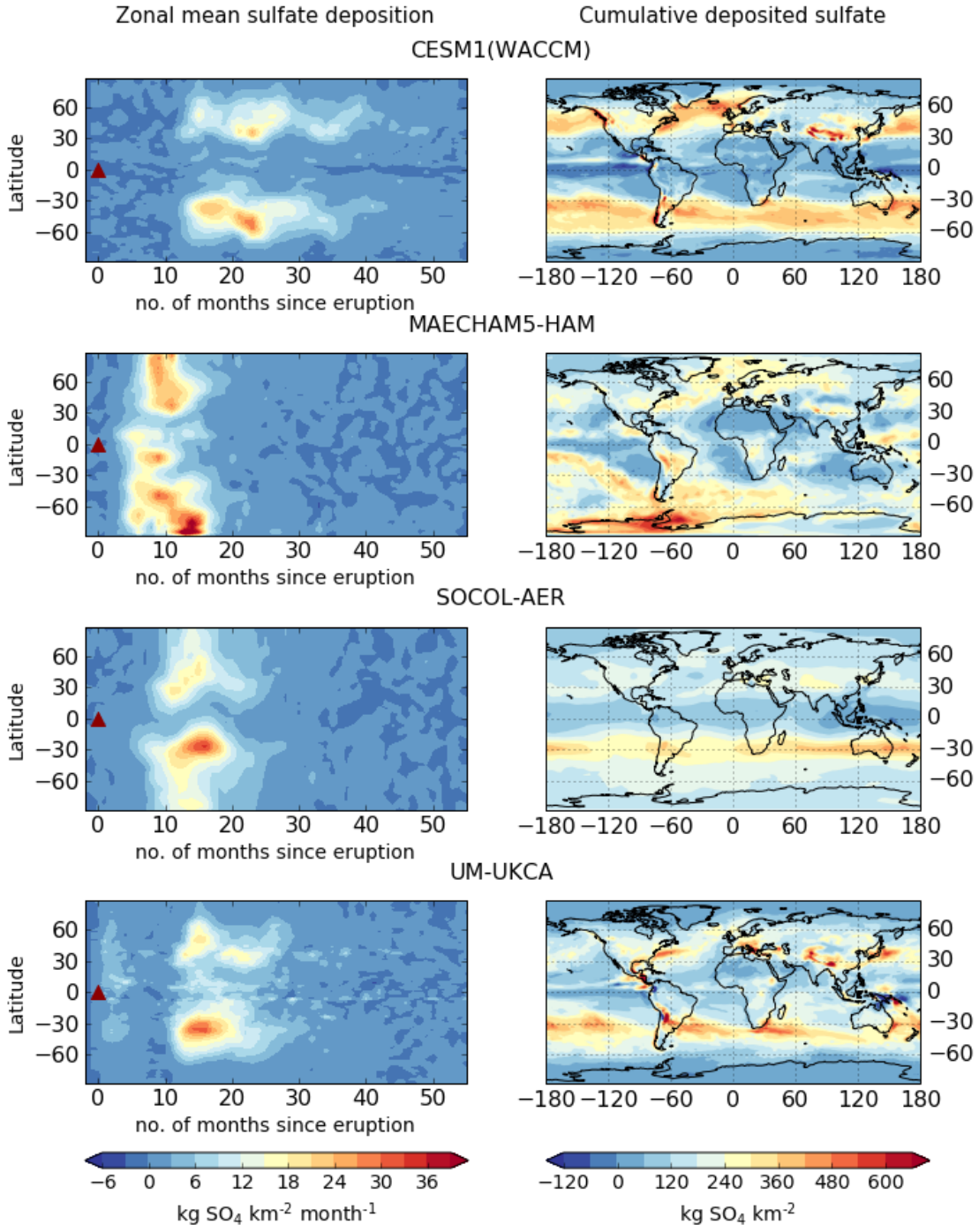

Figure ~~34~~: Zonal mean ~~monthly-mean~~ volcanic sulfate deposition [kg SO_4 km^{-2} $month^{-1}$] (left) and cumulative deposited sulfate [kg SO_4 km^{-2}] (right) for each model (ensemble mean). The red triangle marks the start of the eruption (1 April 1815). Volcanic sulfate deposition is calculated as the difference in total sulfate deposition (wet + dry) between the perturbed and control simulations and this anomaly is summed over the ~5 years of simulation to produce the cumulative sulfate deposition maps (right column).

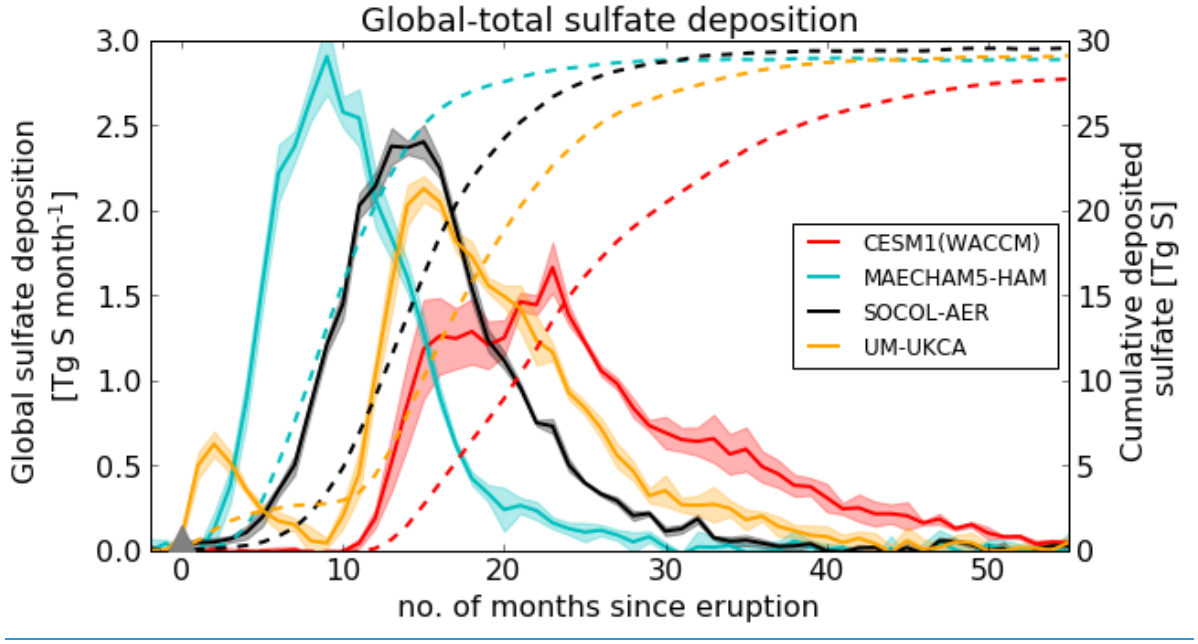

Figure 45: Global -total volcanic sulfate deposition [Tg S month^{-1}] (solid lines – left axis) and global -total cumulative deposited sulfate [Tg S] (dashed lines - right axis) for each model (colours). Ensemble mean is shown by the solid line; shading marks one standard deviation. The grey triangle marks the start of the eruption (1 April 1815).

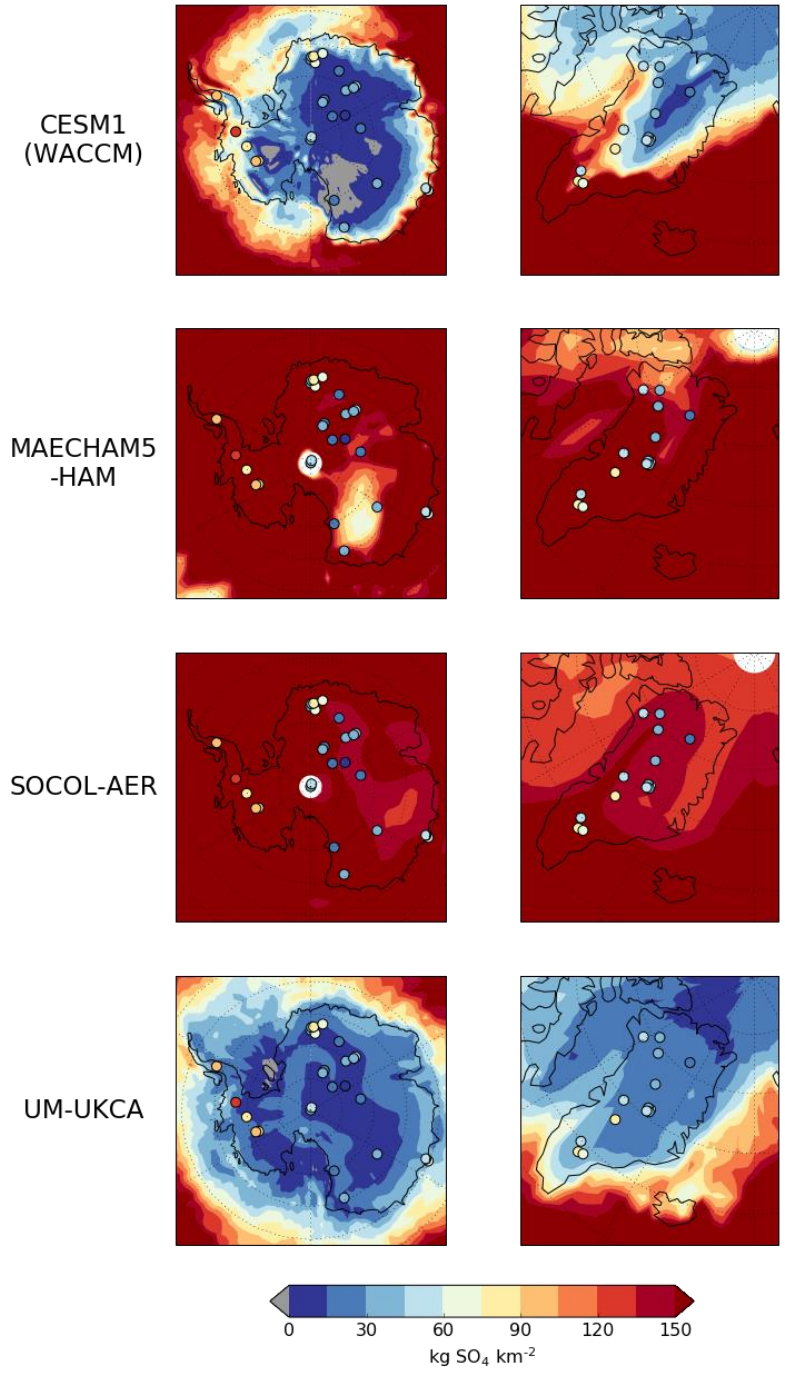

Figure 6: Cumulative deposited sulfate [kg SO_4 km^{-2}] integrated over the whole duration of model simulation (~5 years) on Antarctica (left) and Greenland (right) for each model (ensemble mean). Ice core cumulative deposited sulfate values are plotted as coloured circles. Ice cores from adjacent sites or in close proximity (Table S1) have been slightly relocated to avoid cores completely overlapping. Scaled versions for MAECHAM5-HAM and SOCOL-AER are included in the supplementary information (Fig. S45).

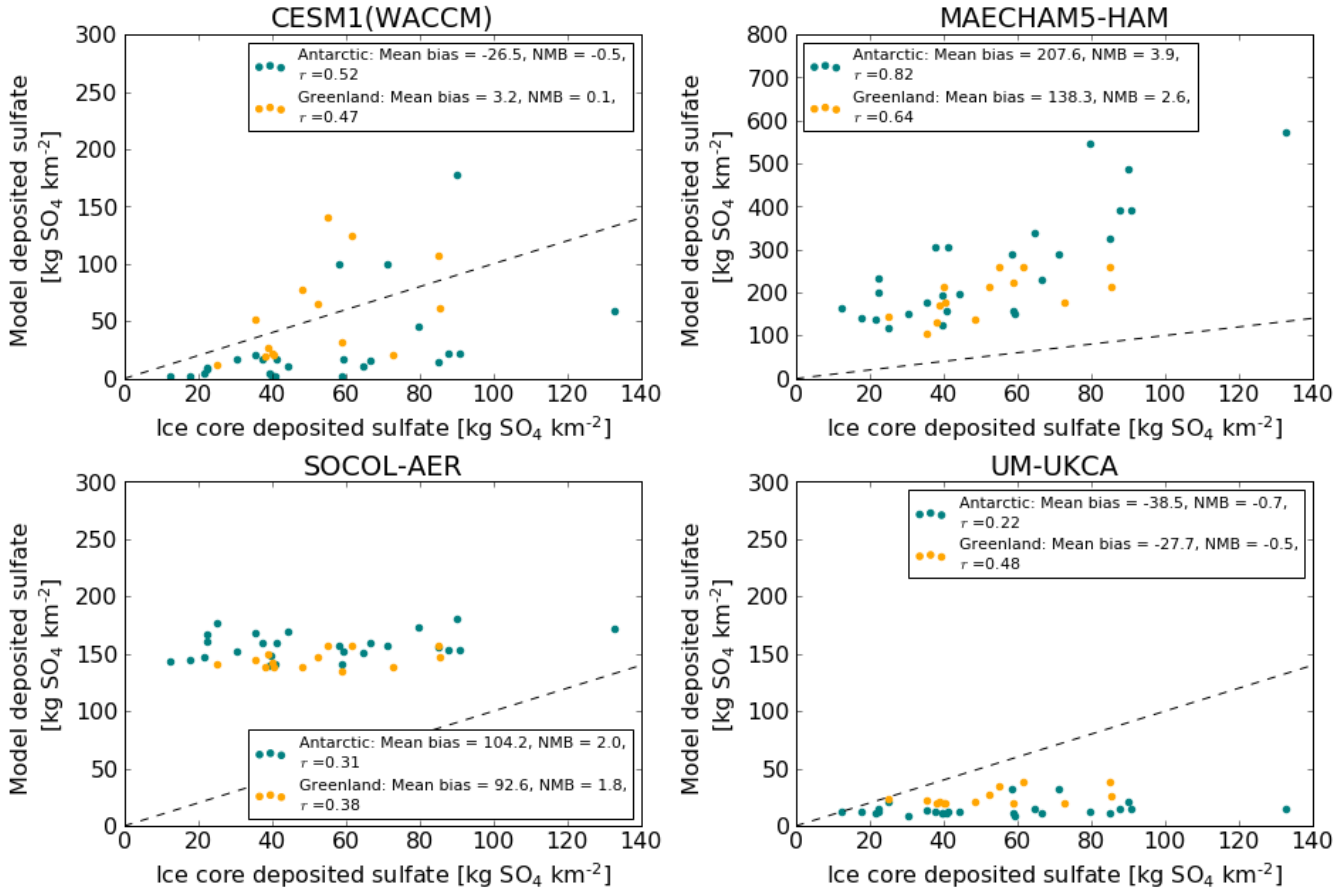

Figure 7: Scatter plots of cumulative deposited sulfate [kg SO_4 km^{-2}] due to the eruption of Mt. Tambora recorded in ice cores vs. that simulated by each model (ensemble mean) in Antarctica (teal points) and Greenland (orange points). Simulated values represent the grid box value where each ice core is located. The dashed line marks the 1:1 line. For each model and for Greenland and Antarctica the mean bias, normalized mean bias (NMB) and correlation coefficient (*r*) between the simulated deposited sulfate and ice core values are shown in the legend. ~~Note the~~There is an increased y-axis scale for MAECHAM5-HAM.

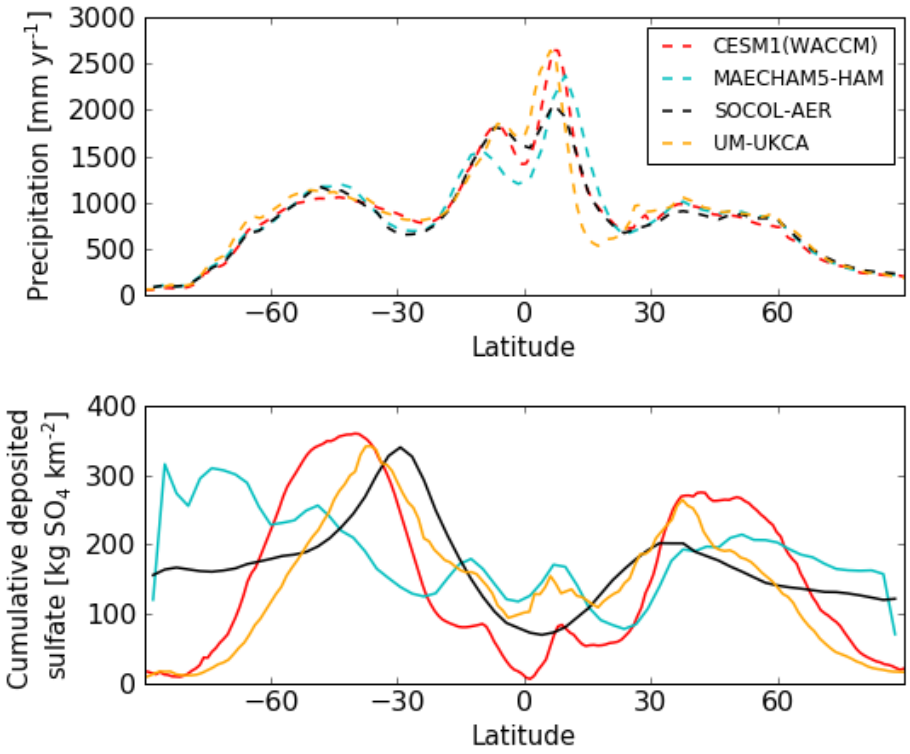

Figure 8: Zonal mean precipitation [mm yr^{-1}] averaged over the first 4 years after the eruption (top panel, dashed lines) and zonal mean cumulative deposited sulfate [kg SO_4 km^{-2}] in the first 4 years after the eruption (bottom panel, solid lines) for the ensemble mean in each model (colours).

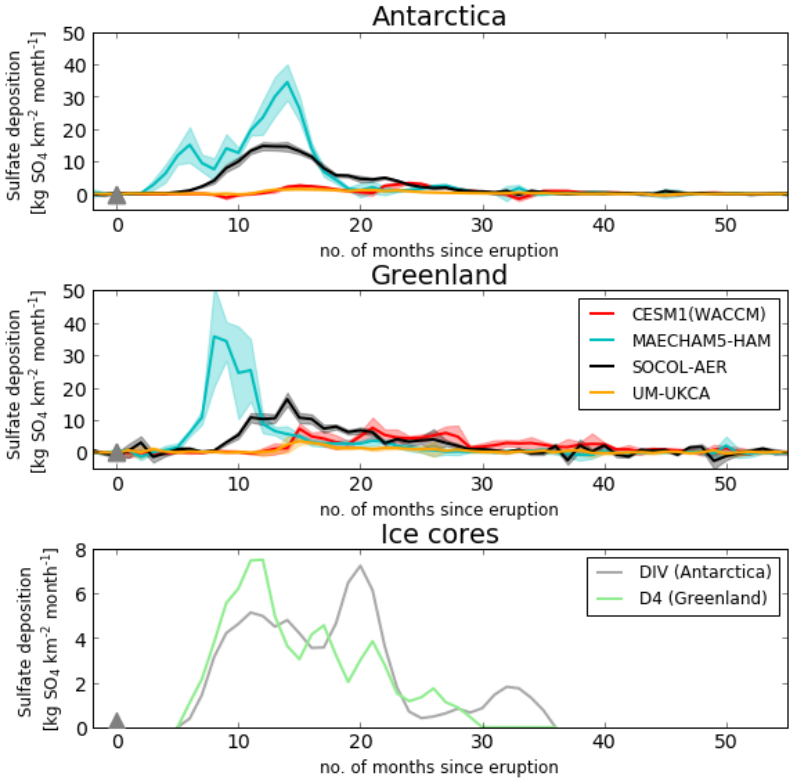

Figure 9: Simulated area-mean volcanic sulfate deposition [kg SO_4 km^{-2} $month^{-1}$] to the Antarctic ice sheet (top panel) and Greenland ice sheet (middle panel) for each model (colours). Each ice sheet mean is defined by taking an area-weighted mean of the grid boxes in the appropriate regions once a land-sea mask has been applied. Solid lines mark the ensemble mean and shading is one standard deviation. In the bottom panel are deposition fluxes from two monthly resolved ice cores (DIV from Antarctica and D4 from Greenland). ~~Note the~~The scale is ~~-~~reduced ~~in~~scale ~~for~~ the bottom panel. The grey triangles mark the start of the eruption.

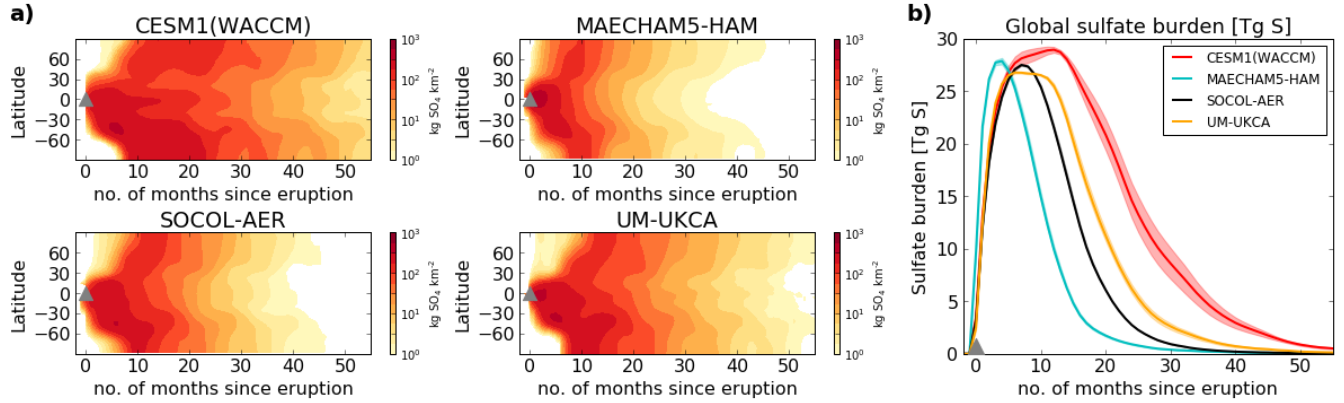

Figure 10: a) Zonal -mean ~~monthly-mean~~ atmospheric sulfate burdens -in each model [kg SO_4 km^{-2}] b) Global -total atmospheric sulfate burdens -[Tg S] in each model (colours). Ensemble means are shown by the coloured lines; shadings mark one standard deviation. Sulfate burdens are monthly mean anomalies. The grey triangle marks the start of the eruption (1 April 1815).

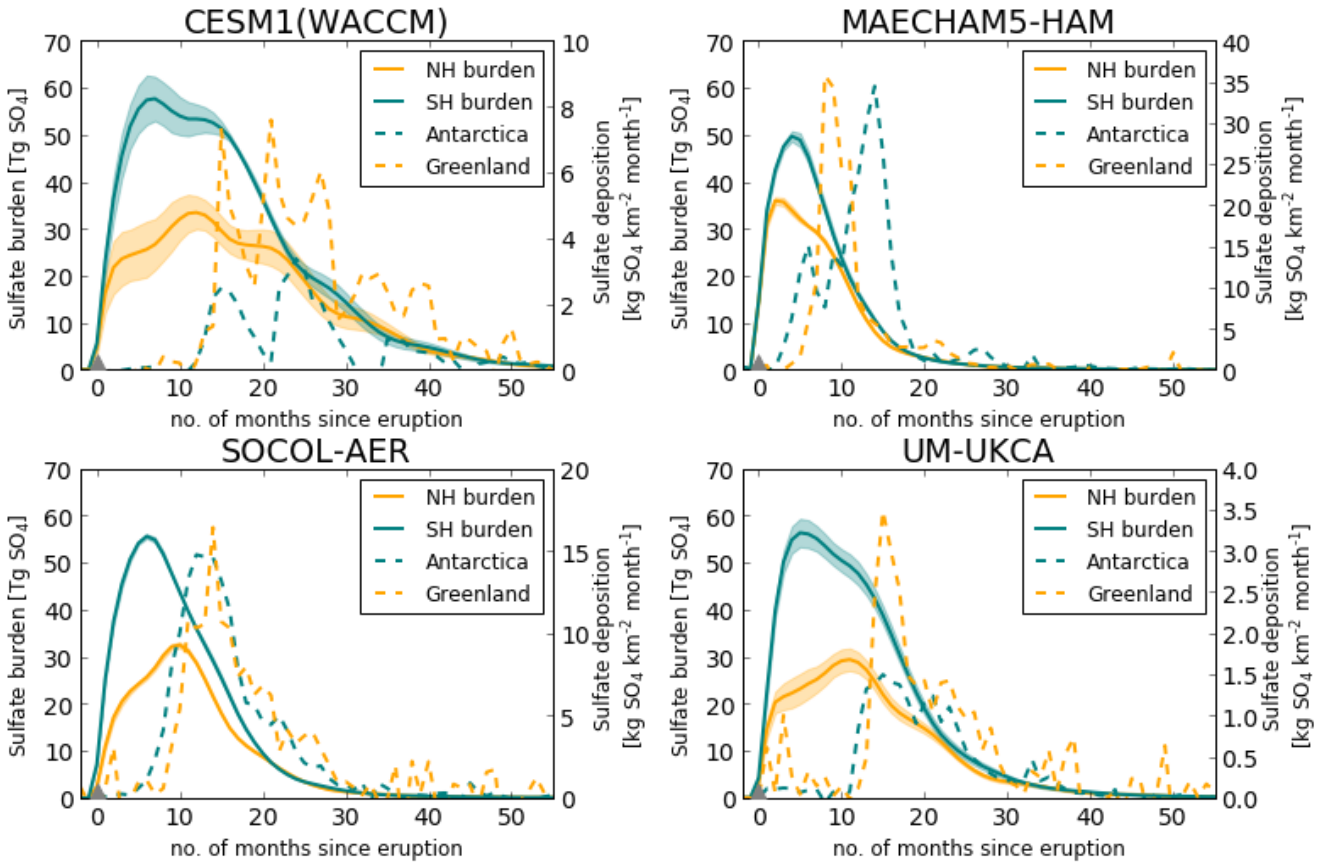

Figure 1~~0~~1: Hemispheric atmospheric sulfate burdens [Tg SO_4] (solid lines show the ~~—~~ensemble mean and ~~,~~shading is one standard deviation) and area-mean ice sheet volcanic sulfate deposition as in Fig. 9 (dashed lines) [kg SO_4 km^{-2} $month^{-1}$] for each model. The grey triangles mark the start of the eruption. There are different scales on each secondary y axis for ice sheet deposition.

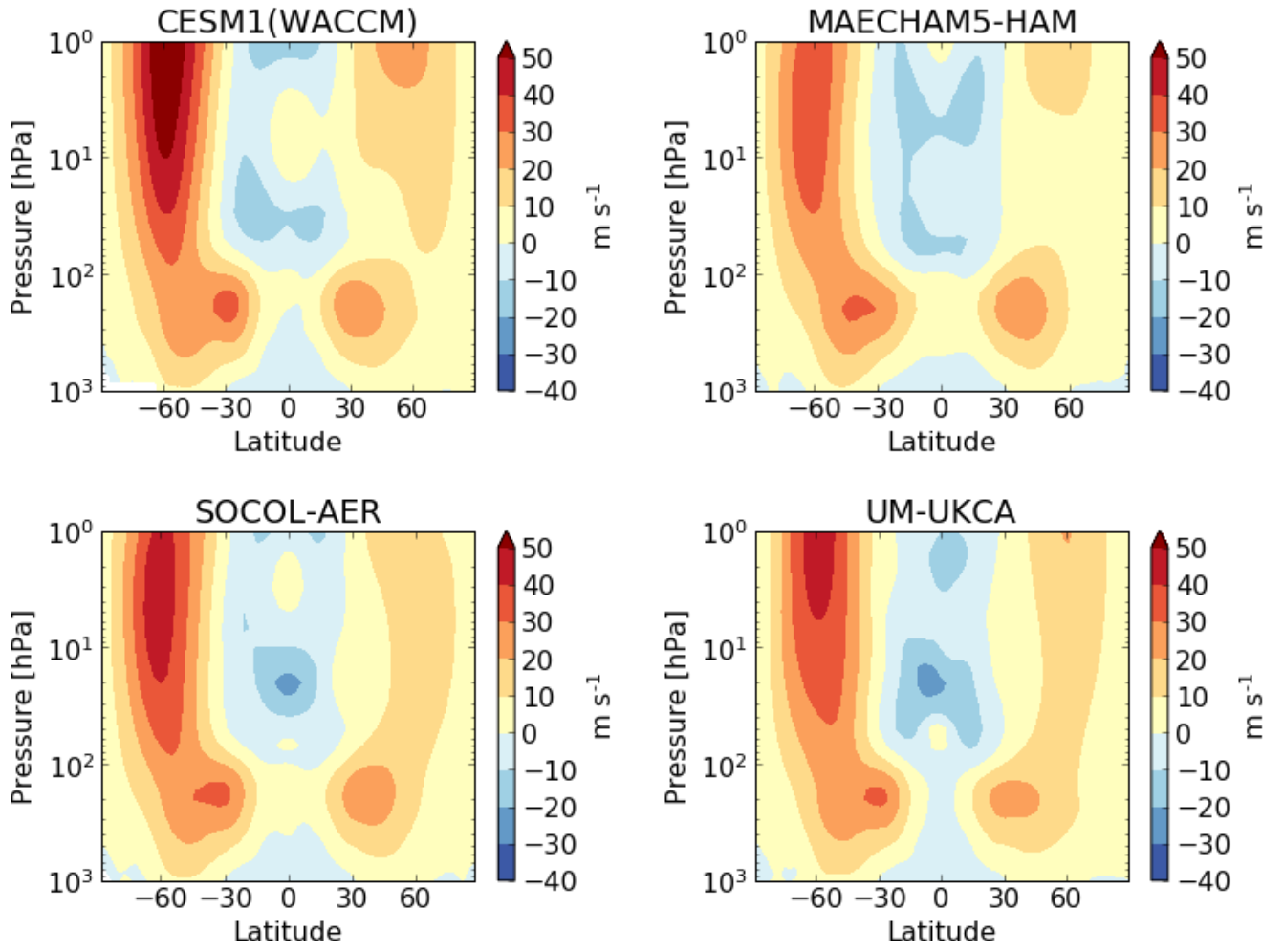

Figure 1~~1~~2: Zonal mean zonal wind [m s^{-1}] averaged over the first year after the eruption (April 1815 – April 1816) in each model simulation (ensemble mean). Zonal wind is output on 36 pressure levels in UM-UKCA, 33 pressure levels in MAECHAM5-HAM and 32 pressure levels in SOCOL-AER. Zonal wind in CESM1(WACCM) is output on an atmosphere hybrid sigma pressure coordinate and has been interpolated to the pressure levels used in UM-UKCA.

Tables

Table 1: Description of models. ~~In the remaining text model names have been truncated to the section in bold.~~ Modal vs. sectional aerosol size distributions are described in the text.

Model	Horizontal resolution	Model top, model levels	Aerosol size distribution	Stratospheric compounds	Het. Chem.[a]	OH	Sulfur source species	QBO	Reference
CESM1(WACCM)	0.94°×1.25°	4.5×10^{-6} hPa, 70 levels	modal, 3 modes	sulfate, PSC, organics	Y	Interactive	OCS (337 pptv), DMS, anthrop[b]. SO_2, volcanic[c] SO_2	Nudged	Mills et al., 2016; 2017
MAECHAM5-HAM	2.8°×2.8° (T42)	0.01 hPa, 39 levels	modal, 7 modes	sulfate	N	Prescribed	OCS (~500 pptv), DMS	NA	Stier et al., 2005; Niemeier et al., 2009
SOCOL-AER	2.8°×2.8° (T42)	0.01hPa, 39 levels	sectional, 40 size bins	sulfate, PSC	Y	Interactive	OCS (337 pptv), DMS, CS2, anthrop. SO_2, volcanic SO_2	Nudged	Sheng et al., 2015a
UM-UKCA	1.25°×1.875° (N96)	84 km, 85 levels	modal, 7 modes	sulfate, PSC, organics, meteoric dust	Y	Interactive	OCS (~500 pptv), DMS, anthrop. SO_2, volcanic SO_2	Internally generated	Dhomse et al., 2014; Brooke et al., 2017

[a]Heterogeneous Chemistry

[b]Pre-industrial anthropogenic SO_2

[c]Volcanic SO_2 indicates SO_2 from passively degassing volcanoes.

Table 2: Model parameters used for the Tambora simulations.

Parameter	Value in this study
SO_2 emission	60 Tg SO_2
Eruption length	24 hours
Eruption date	1 April
Latitude	Equator[a]
QBO phase	Easterly
SO_2 injection height	22-26 km~~*~~[b]

[a]SO_2 was emitted at 0ºN in CESM1(WACCM), MAECHAM5-HAM and UM-UKCA and at 8ºS in SOCOL-AER and at a longitude of 118ºE

[b]~~*~~The altitude distribution for the SO_2 emission ~~Injection height differed slightly~~varied slightly between the models: SOCOL-AER's SO_2 emission flux was between 22-26 km, increasing linearly with height from zero at 22 km to max at 24 km, and then decreasing linearly to zero at 26 km. MAECHAM5-HAM injected at a single model level at 30 hPa (~24 km). UM-UKCA and CESM1(WACCM) used a uniform injection between 22-26 km. ~~WACCM also emitted SO_2 uniformly between 22-26 km, however the overlap of model levels with the emission altitude range resulted in emission fluxes peaking in the centre of the plume, similar to SOCOL.~~but as the models are not on regular grids and their vertical resolutions differ, the distribution of the emission over the model grid boxes cannot be exactly the same. As a result, the injection profiles differed slightly between the models.

Table 3: Annual global -total deposition fluxes of SO_2, SO_4 and SO_X (SO_2 + SO_4) for dry deposition, wet deposition and total dry + wet (Tg S yr^{-1}) in the pre-industrial controls and in the ACCMIP multi-model mean (see text).

Model	Dry SO_2	Wet SO_2	Total SO_2	Dry SO_4	Wet SO_4	Total SO_4	Dry SO_X	Wet SO_X	Total SO_X
CESM1(WACCM)	5	11	16	2	11	13	7	22	29
MAECHAM5-HAM	2	2	4	0.5	19	19	3	21	24
SOCOL-AER	120	130	250	65	139	194	185	2619	4434
UM-UKCA	7	5	12	4	25	29	11	30	41
ACCMIP multi-model mean	-	-	-	-	-	-	11	23	34

Table 4: Global -total cumulative deposited sulfate [Tg S] from dry and wet processes for each model (ensemble mean).

Model	Dry deposition	Wet deposition	Total deposition
CESM1(WACCM)	2.4	25.4	27.8
MAECHAM5-HAM	0.2	28.~~6~~7	28.9~~8~~
SOCOL-AER	1.0~~0.8~~	28.5~~0~~	2~~8~~9.5~~9~~
UM-UKCA	3.7	25.4	29.1

Table 5: Greenland and Antarctica ice sheet mean cumulative deposited sulfate and ratio (Antarctica deposition/Greenland deposition) and peak NH and SH sulfate burdens (total atmospheric column burden anomaly) and ratio (SH burden/NH burden) for each model (ensemble mean). Also included is the equivalent mean deposited sulfate on each ice sheet calculated from ice cores (Gao et al., 2007; Sigl et al., 2015).

Model	Mean Antarctica deposited sulfate [kg SO_4 km^{-2}]	Mean Greenland deposited sulfate [kg SO_4 km^{-2}]	Antarctica/ Greenland deposition ratio	Peak SH sulfate burden [Tg SO_4]	Peak NH sulfate burden [Tg SO_4]	SH/NH burden ratio
CESM1(WACCM)	36	109	0.3	58	34	1.7
MAECHAM5-HAM	264	194	1.4	50	36	1.4
SOCOL-AER	16370	14815	1.51	56	321	1.8
UM-UKCA	19	31	0.6	56	29	1.9
Sigl et al., 2015	46	40	1.2	-	-	-
Gao et al., 2007*	51	59	0.9	-	-	-

*aerosol reported 75:25% H_2SO_4:H_2O

Table 6: Burden-To-Deposition (BTD) factors [* 10^9 km^{-2}] between the hemispheric peak sulfate burden [Tg SO_4] (total atmospheric column burden anomaly) and the mean ice sheet cumulative deposited sulfate [kg SO_4 km^{-2}] for the four models and from Gao et al. (2007). Included are the values for the ensemble mean factor and the range from individual ensemble members.

	NH_BTD [10^9 km^{-2}]		SH_BTD [10^9 km^{-2}]	
Model	**Ensemble mean**	Ensemble range	**Ensemble mean**	Ensemble range
CESM1(WACCM)	0.31	0.29-0.34	1.63	1.44-1.96
MAECHAM5-HAM	0.19	0.14-0.24	0.19	0.17-0.20
SOCOL-AER	0.227	0.205-0.249	0.343	0.321-0.35
UM-UKCA	0.97	0.74-1.14	2.91	2.67-3.30
Multi-model mean	**0.424**	-	**1.276**	-
Gao et al., (2007)	1	-	1	-