We have modified the manuscript addressing all the comments from both of the referees. Below are the replies to Referees #1 and #2 followed by the revised manuscript which includes all the changes marked in blue.

**Reply to Referee #1**
We thank the referee for the comprehensive review of our manuscript. We will improve the manuscript following the points made by the referee. Below are the replies to Referee #1. Original comments are in italic.

1. *Performance of SALSA2.0 is compared to the ECHAM-HAMMOZ default (modal) aerosol module M7. Consequently, many of the differences in this work are attributed to the difference in the numerical treatment of the aerosol size distribution in modal or sectional schemes. Because of this, the manuscript does not really give an indication of how well SALSA2.0 performs compared to other sectional modules.*

Comparing SALSA2.0 to other sectional models would be extremely interesting. However, to have a meaningful comparison of two aerosol modules, they should be compared within the framework of the same global atmospheric model. If the host atmospheric models are different, it is extremely challenging to distinguish if the differences in the simulated aerosol fields are caused by the differences in the aerosol modules or the atmospheric models. In another study (Saponaro et al. (2018) in preparation), we compare the ability of three aerosol models to simulate aerosol-cloud interactions, two of which (HAM-M7 and HAM-SALSA) share the same atmospheric model ECHAM, while the third aerosol module is in a different global model. In that study, we show that aerosol properties are much more similar for the aerosol modules which are in the same atmospheric model than for the aerosol module in a different atmospheric model. This conclusion is also supported by the study by Mann et al. (2014) where 12 aerosol models of different complexity were compared. Thus, conducting a meaningful comparison between SALSA and another sectional model would require excessive work and therefore it would not be possible within this study. We will mention this in the revised manuscript. On the other hand, SALSA has been (e.g., Mann et al., 2014; Tsigeridis et al., 2014; Kipling et al., 2016) and will be included in the international AEROCOM project model experiments, where models are evaluated against each other and against observations.

2. *The paper describes the difference between SALSA1 and SALSA2.0 well, but does not discuss the reasons why certain changes are made to the SALSA module. Please explain what the main problems with SALSA1 are and how these changes contribute to the improvement of the aerosol module.*
The major change between SALSA1 and SALSA2.0 was how the concentrations of different chemical species in different sized particles are treated. In SALSA1, particles were separated into three subregions depending on the particle size. In the subregion for the largest particles (larger than 700 nm), only the particle number and sea salt, mineral dust, and a lumped “water soluble species” mass concentrations were tracked. This choice was done to keep the number of tracer variables to the minimum. However, this had two unwanted implications. First, the total amount of sulfate, organic aerosol, and black carbon was lost when the particles grew to sizes larger than 700 nm. This was a practical problem in AEROCOM experiments where the information of individual species was required (e.g. Kipling et al., 2016). Second, although in the troposphere microphysics have very little influence on the size of particles in the third subregion, when simulating volcanic eruptions or stratospheric solar radiation management, condensation can grow the largest particles. This problem caused the model to have problems in simulating the growth of particles in a volcano plume since the third region particles did not grow which resulted in underestimating the effective radius of the volcano plume (Kokkola et al., 2009). In SALSA2.0 we extended the second subregion to cover also the size classes of subregion three. This way we can track the concentrations of all chemical compounds and the growth by microphysical processing is not limited in the largest size range. Figure 1 demonstrates how SALSA2.0 can simulate the growth of large particles while SALSA1 underestimates their growth. We will add discussion on these aspects in the revised manuscript.

Another major change was changing the size distribution structure from the moving
Figure 2: Relative difference [%] between in AOD the hybrid bin and moving center method (Figure adapted from the presentation “Representing the Evolution of Aerosol Size Distribution in Atmospheric Models” by Harri Kokkola in the 2015 International Aerosol Modeling Algorithms Conference, UC Davis, CA).

center method (Jacobson, 2005) to the hybrid bin method (Young, 1974; Chen and Lamb, 1994). As explained in the manuscript, the moving center method causes numerical artefacts (Mohs and Bowman, 2011). We also have seen in our simulations that moving center consistently produces lower number concentrations. For example, in Bergman et al. (2012) study, SALSA1 significantly underestimated the number size distributions measured at EUSAARI sites whereas in this study we did not experience such underestimation. This underestimation of particle numbers in SALSA1 was analyzed to be partly caused by the moving center method. In Figure 2 we show the relative difference AOD over the US between a model setups using the hybrid bin method and the moving center method. From the figure we can see that AOD is always lower when using the moving center method. We will add a brief discussion of this in the revised manuscript.

3. Previous work (also referred to in this article) has shown that modal aerosol representations do not perform well in simulating stratospheric aerosol caused by volcanic eruptions and that sectional approaches yield far better results. It is therefore not surprising that SALSA2.0 performs better than M7, but how does it compare to e.g. SALSA1? Overall, the discussion of the Pinatubo simulation is thin and mainly addresses issues in M7. As it is now, it might be better to remove this section from the paper as the remainder is already a very comprehensive comparison to observations.
The intention of this comparison is not only to evaluate if SALSA2.0 performs better than M7 but also to evaluate how well SALSA2.0 can simulate the evolution of the stratospheric aerosol that originates from the Mt Pinatubo plume. This comparison is also justified because the volcano plume is simulated with a sectional model and a modal model coupled to the same atmospheric model. The main result from this evaluation is that SALSA2.0 can simultaneously simulate both tropospheric and stratospheric aerosol without changing the setup. It has to also be noted that M7 and SALSA can be considered to be used for different types of studies. As a sectional model, SALSA would be a preferable choice for simulating stratospheric aerosol, volcano eruptions, and geoengineering. However, we will add simulation results from SALSA1 and will also add discussion on the differences between SALSA1 and SALSA2.0 simulated volcanic aerosol.

4. The authors state that a size-resolved wet deposition scheme for SALSA2.0 is still under development. To make a fair comparison to M7 an older removal scheme is also used for those simulations. In my opinion, the quality of this paper would greatly improve if the new wet deposition scheme is included in this work. In the comparison to aircraft measurements in the results section it is also explicitly mentioned how the new wet deposition scheme would improve the results of the simulations with SALSA2.0.

In this paper, the focus in comparing M7 and SALSA is in comparing the modal and sectional approach and thus we kept the two model configurations as close as possible. The new wet deposition scheme of M7 is used in the upcoming paper by Tegen et al. (2018) which will provide a detailed evaluation of ECHAM-HAMMOZ with M7. The inclusion of the new wet deposition scheme would in SALSA also be a valuable addition, however, here we present the release version ECHAM6.3.0-HAM2.3-MOZ1.0 which does not include the new wet deposition scheme. The work on the new wet deposition scheme is underway, but it is by no means a trivial task to include such a new scheme in the model. Simulating wet deposition properly in a global scale model is a major challenge for the whole aerosol modelling community. The improved wet deposition schemes in M7 are for a modal scheme and cannot be applied to SALSA2.0. In order to implement an improved sectional wet deposition scheme in SALSA2.0 will require extensive model development, testing, and analysis and could be a topic of an independent article (see e.g. Croft et al. (2009, 2010)). Consequently, it is beyond what can be done within this manuscript.

5. Title / Page 8, line 9-13 Judging from the text, the MOZ module is not used in this work and HAMMOZ is reduced to HAM only. This is a bit misleading and causes confusion in the text. It is not clear what value the combination of HAM and MOZ has in this work. It should be clarified better what MOZ does in the model simulations for this work, otherwise it might be better to remove MOZ from the title.

We will clarify in the revised manuscript that MOZ is not used in these simulations.
However, the model licence requires that when publishing results with the model, the full name ECHAM-HAMMOZ is used and the exact release reference is stated.

6. **Page 2, line 31** 100 size classes is a bit of an exaggeration, in global models the number of size classes is usually (much) lower.

Here we refer to the study by Mann et al. (2014) where the GEOS-Chem-APM model is reportedly using 100 size classes.

7. **Page 5, line 21-23** Two subranges of SALSA1 are combined into one in SALSA2.0, what is the reason for this simplification and what are expected changes in the simulated aerosol size distribution?

See our reply to the second referee comment.

8. **Page 5, line 25-26** The moving center method is replaced by the hybrid bin method. What are the downsides to this method as is was not used in SALSA1 before?

The moving center method was not chosen for SALSA1 because of downsides of the hybrid bin method but because it seemed to be suitable for global scale models. Like all size distribution methods, hybrid bin method suffers from numerical diffusion (e.g. Dinh and Durran, 2012). However, in our model framework it improved the correspondence between the simulated and the observed aerosol size distributions.

9. **Page 7, line 8-9** In view of the importance of meteorology for e.g. dust emissions, what is the nudging time interval used in the simulations?

The relaxation times for the nudging of the surface pressure, vorticity, and divergence are 24 h, 6 h, and 48 h, respectively. We will add this information in the revised manuscript.

10. **Page 8, line 14-34** The different parameterisations of SALSA2.0 and M7 are explained extensively in Section 2.3, but several of these are changed for the simulations. This is very confusing and makes large parts of previous section irrelevant. It would be better to describe what is actually used. Also, this means that this work presents results from a suboptimal model run and the full potential of the SALSA2.0 module in the ECHAM-HAMMOZ model is not shown.

With respect to SALSA, the only parameterization that was changed was secondary organic aerosol (SOA) formation using volatility basis set (VBS). We will remove the row describing SOA for both SALSA and M7. This setup is the default for the model release. It may be suboptimal for simulations of SOA formation. However, using VBS increases computational time and significantly increases model output to the extent that it may not be practical to use it for simulations that do not focus on organic aerosol.

11. **Page 9, line 26** How is this ensemble constructed? What is the difference between the 5 members?
We have perturbed offline anthropogenic aerosol emissions by values of the order of $10^{-6}$ which is an insignificant number for emission strengths but due to the chaotic nature of the atmospheric model, changes the model dynamics sufficiently. This will be clarified in the revised manuscript.

12. **Page 13, line 6** Judging from this work, the implementation of the new Long et al. (2012) sea salt emission decreases the model performance, why was it introduced?

This is only true regarding aerosol mass. An upcoming study by Tegen et al. (2018) will show that the comparison between the measured and modelled number size distribution improves going from Guelle et al. (2001) to Long et al. (2011) parameterization. In addition, they show that aerosol optical depth compares better with MODIS retrieved AOD’s when using the Long et al. (2011) parameterization.

13. **Page 15, line 4** Here, the low AOD bias is (almost) completely attributed to the low SS emissions. Although this assumption is acceptable for the SH, there is also a strong bias in the NH high latitudes. Here, the low bias over the land masses cannot be attributed to sea salt only.

This is a very good point. The low bias at high latitudes can also be attributed to issues in aerosol transport. Bourgeois and Bey (2011) have shown that ECHAM-HAMMOZ underestimates the aerosol load over the high latitudes. We will add discussion on this in the revised manuscript.

14. **Page 15, line 13-14** How did you arrive to this conclusion?

We came to this conclusion comparing modelled and observed time series. However, we will remove this line from the revised manuscript since these results are not shown in the manuscript.

15. **Page 15, line 25** Why is this not mentioned in Section 4.1.1?

We will add discussion on MODIS biases to Section 4.1.1

16. **Section 4.2** Restructure section. Multiple species of multiple model runs are compared to multiple measurement networks. This already makes the discussion hard to read. I suggest a fixed format/structure in discussing the different species to help the reader.

We will follow the suggestion of the referee and change the structure to separate different species.

17. **Page 18, line 1-22** Include comparison results of M7 in discussion of SU/BC.

We will add M7 results for sulfate and black carbon in the revised manuscript.

18. **Page 19, line 9** Aerosol load over oceans is not low over SH subtropics.

We will rephrase this to state that AOD is low south of latitude 40°S.
19. **Page 19, line 15** If periods in observations and simulations are different, how are they collocated?

The monthly model values are constructed by averaging daily means only for days where an observation is available. Moreover each model monthly mean is spatially collocated to the location of the observation station (by bi-linear interpolation). We will clarify this in the revised manuscript.

20. **Page 22, line 20** Why are monthly mean model values used here? Having 3 hourly output, collocation can be greatly improved. Also, comparison for SU and OA is based on daily mean output. What is the reason for this inconsistent approach?

This was done to enable an easier comparison between earlier model experiments. For black carbon, we reproduced the AEROCOM experiment by Koch et al. (2009) which used monthly mean values. For sulfate and organic aerosol, we reproduced the model evaluation by Heald et al. (2011) where daily values were used.

21. **Page 25, line 5-6** This conclusion is too strong and drawn too quickly. It would be the case for the ARCTAS Spring and ARCPAC campaigns, but for the ARCTAS Summer, the wet deposition scheme barely influences the results for the lower part of the atmosphere where observations are available. Also, for the source regions, an increase due to the wet deposition scheme, would increase the already high bias in SALSA.

This is a good point from the referee. This is purely speculation and we will remove this sentence from the revised manuscript.

22. **Page 26, line 1** Difference was fairly small”. Can this statement be quantified?

We have quantified this for the revised manuscript as follows. “For most campaigns and height levels, the relative difference in black carbon mass concentration is less than 50% and the shape of the vertical profiles are very similar, mostly overlapping each other. The largest difference is for black carbon in the CR-AVE campaign where the relative difference in the mass concentration is ~83% in the lowest layer.

23. **Section 6** In this section, it is explained why the section approach of M7 does not perform well in simulation the stratospheric aerosol burden resulting from a volcanic eruption. There is even a reference to a solution for this problem. Yet you dont incorporate this in your model and compare the performance of SALSA2.0 mainly to the simulation with the unadjusted M7 scheme. As a result, it is difficult to really judge how well SALSA2.0 performs in these simulations. It would be more interesting to see the performance of SALSA2.0 to other sectional aerosol modules or modal schemes that were properly adjusted.

Implementing a modal structure with varying mode width would require rewriting major part of HAM and M7. In addition, it would increase the amount of tracer variables which would be undesired for M7 in which the number of these variables is kept to the
minimum. As it is said in the manuscript, the mode widths of M7 have been optimized for the troposphere and were not designed to represent stratospheric aerosol.

Regarding comparison to another sectional model, a meaningful comparison to another sectional model would have been too challenging to do within this study (as was explained above).

24. **Page 26, line 1-2 How are the model values and observations collocated?**
   
   In the simulations, where we used only year 2010 values, for black carbon, we used the modelled monthly values from the flight path for the corresponding month of the year. For sulfate and organic aerosol, we used the modelled daily values from the flight path for the day of the year which corresponded to the observations. For the whole period, we used the same values, but also for the corresponding year of the observations. We will clarify this in the revised manuscript.

25. **Section 7. Conclusions** The structure of the conclusion section is unnecessarily confusing. Follow same order as discussion in Results section.
   
   We will restructure Conclusions as suggested

26. **Page 29, line 6 What are recommendations for optimizing?**
   
   The details will be explained in an upcoming study by Tegen et al. (2018)

27. **Page 30, line 1-2 Underestimation of particle number in SALSAI not mentioned in Section 4.1.1.**
   
   Here we refer to a study by Bergman et al. (2012) and will add the reference in the revised version

28. **3 Technical comments** Page 4, line 12 Sentence is not clear, please rephrase.
   
   We will rephrase this as: "In its default setup, HAM uses the modal approach together with the model aerosol microphysics module M7 (Vignati et al., 2004).

29. **Page 5, line 13 “using the volume ratio” → “using volume ratio”**
   
   This will be corrected in the revised manuscript

30. **Page 7, line 10 Add full name of PCMDI**
   
   We will add the full name in the revised manuscript.

31. **Page 10, line 21-23 Why is India omitted from this list?**
   
   We will add India to the list.
32. **Page 11, Figure 2** Add statistics (e.g. corresponding global mean AOD to a,b,c and correlation coefficients and NMBs to d,e) to the figure for a good overview between model configurations.
   We will do as suggested.

33. **Page 12, line 2** Equation straightforward, can be removed and explained in words.
   We will remove the equation.

34. **Page 13, line 1** Fig 4. → Fig. 5
   We will correct this.

35. **Page 14, line 12** Reference to current section.
   We will remove the reference.

36. **Page 15, line 17** Add minus sign to 0.05.
   We will correct this.

37. **Page 16, Figure 6** Change colours of Asia and North America. These are the two regions discussed in the text but hardly distinguishable from each other. Also, adding a regional mean values would provide a good overview of model performance.
   We will add regional values to the revised manuscript. However, to us the colors look sufficiently different and since the values for the two regions are completely different, they are easy to separate. We do not see the need for changing the colors.

38. **Page 17, Figure 7** Remove additional abbreviations of species in lower left corner of each panel. Add names of network to panels.
   We will do as suggested.

39. **Page 20, Figure 9** Observed and simulated values in year 2010? Please add to caption.
   This is correct and will be added to the figure caption.

40. **Page 22, line 32** Vertical profiles of AMMA, ARCTAS Spring and OP3 are not captured well either.
   We will mention this in the revised manuscript.

41. **Page 26, line 1-2** Add references for the HIRS and lidar observations.
   We will add the references.

42. **Page 29, Fig. 14** Add errorbars to observed values.
   We will add the uncertainty estimates of the observations in the figure caption.

43. **Page 30, line 23** configurations → configurations
   We will correct this.
References


Reply to Referee #2

We thank the referee for a comprehensive review of our manuscript. Addressing the points made by the referee will improve the paper. Below are the replies to Referee #2. Original comments are in italic.

1. However, some of the principal statements made in the Abstract are not supported by the results presented in the paper, and much more care needs to be taken in the statements interpreting potential reasons for differences between the sectional and modal schemes.

In particular, the authors claim that the size distribution comparisons shown in Figure 4, shown for locations where AOD difference is largest, are indicative of different microphysical processing in the modal and sectional schemes, and that this is the reason why the two aerosol schemes predict different extinction/AOD in these regions. But from careful inspection of Figure 4, it's clear that is not the case. The authors already identify the two locations (in China and in Russia) as regions where the observed AOD is very high, and clarify that the Russian site is in a region where biomass burning emissions are high. The China site is in a region of strong anthropogenic emissions. The size distribution of the black carbon (the black bars in the stacked bar chart) are very different between the SALSA and M7 simulations at both locations and this clearly indicates that there is a systematic difference in the size at which primary carbonaceous aerosol particles are emitted, which is a much more likely explanation of the reason for the difference.

At the China site, the M7 run has about half of the BC in particles larger than 200nm, whereas for the SALSA run this is only about 10%. This indicates that there is a systematic difference between the two schemes in the sizes assumed for primary carbonaceous emissions. This is an important issue, because if such large differences in AOD could indeed be attributed to the simpler modal scheme having inadequate representation of microphysical processing compared to the bin scheme, then this could be cited in the literature extensively as a reason to justify the additional expense required for sectional aerosol schemes.

It is noticeable that whereas Table 1 explains in detail the size segregation of the sea-salt and dust emissions, there is no information given about the assumed size at which the carbonaceous particles are emitted (yet these are the dominant primary aerosol in polluted regions).

I am sure this is just an oversight in the writing of the paper, and that there was no intention to omit this information, or to make a statement that is not supported by the results.

1) Abstract, page 1, lines 5-6 As per my main comments above about Figure 4, this sentence is not supported by the results and needs to be removed or reworded. If the authors can repeat either the M7 or the SALSA simulations with the emissions size distribution for emitted carbonaceous particles identical in the two schemes then it may be possible to make some statement about this, but the different BC-size-distribution in M7 clearly
indicates there is a substantial difference in the “emissions size distribution” applied for carbonaceous particles in the two aerosol schemes, which is much more likely reason for the difference in AOD between the two schemes. In any case the locations shown in Figure 4 (in Russia and China) are regions of very strong primary emissions. Lee et al. (2011) apply a perturbed parameter uncertainty analysis to show how (at least for the global aerosol microphysics scheme applied there) the regions where aerosol properties are most dominated by uncertainties in microphysical processes are away from such “emissions hot spots”. So even if one of the models was re-run with the same “primary emissions size distribution” as for the other, one might expect any difference from microphysical processes to have most impact in a different region than the two locations shown.

It is true and a good point that differences in emission sizes could explain the differences over areas with high anthropogenic emissions. However, in our simulations this is not the case since for offline anthropogenic emissions, we use identical emission size distributions for SALSA and M7 (see Page 7 in the manuscript). There will be some difference resulting from remapping modal emissions to SALSA size classes, however the emission masses and numbers and their size distributions are identical for M7 and SALSA. The significance of microphysics calculation over the chosen areas can be demonstrated by two SALSA runs: one where condensing organics are treated either assuming them to be non-volatile and one where they are assumed to be semi-volatile, however so that the resulting secondary organic aerosol yield is approximately the same. In Figure 1 we show how the simulated yearly mean aerosol optical depth (AOD) compare against MODIS (Figure 1a) over China. Panel b shows the AOD with non-volatile OA and c with semi-volatile OA. From the figure, we can see that although everything else except for the microphysical treatment of OA is exactly the same, there is a large difference in simulate AOD’s. Other size dependent processes such as wet removal do also contribute to the difference, but the changes are initiated by differences in microphysics.

Figure 1: Retrieved aerosol optical depth by a) MODIS, b) SALSA2.0 with non-volatile organic aerosol, and c) SALSA2.0 with semi-volatile organic aerosol.
Lee et al. (2011) studied the sensitivity of CCN on different microphysical processes. The conclusions from that study do not necessarily apply for AOD. It may also be that the sensitivities of a modal model such as GLOMAP-mode (which was used by Lee et al. (2011)) and a sectional model can be different. We have shown previously that mode merging causes some damping effect on the sensitivity of CCN sized particles (Korhola et al., 2014)

We will add a more detailed description of the emission sizes in the revised manuscript. However, we don’t see a reason to change our conclusions on this matter.

2. *It is also clear from Figure 2 that, for many regions, the M7 scheme could actually be argued to perform better (compared to the MODIS AOD) than the SALSA scheme. SALSA seems to have substantial bias over North Africa and over marine regions in the Southern tropics, for example.*

This is true and we will add regional values for the comparison between the modeled and AERONET AOD for both SALSA and M7 for a more detailed comparison and discussion.

3. *One final general comment was that it needs to be stated somewhere early on in the text the difference between the acronyms “HAM” and “M7”. My understanding is that “HAM” is the overall modal aerosol scheme and that M7 is a component of HAM, basically the modal microphysical routines.*

The reason I ask for this clarification is that I was expecting then the SALSA to not just be an alternative to M7, but an alternative to HAM, and that perhaps the correct naming should then be ECHAM6.3.0-SALSA2.0-MOZ1.0 when the SALSA scheme is applied.

Please can the authors clarify I am understanding this correctly.

However, perhaps that is not quite right and the implementation of SALSA into the model has in fact only implemented the microphysics routines within SALSA (or indeed that SALSA has always only been the microphysics routines). I realise that within HAM there is a separate acronym for the microphysics routines (M7) than the overall modal framework, which is known as HAM. By contrast many other aerosol schemes do not have this distinction and there is only one acronym for the overall aerosol module including both the microphysics routines and the other aspects (primary emissions, dry deposition, scavenging). The naming convention of the different parts of the model are important in this case as it helps the reader to appreciate which aspects of the HAM scheme have been retained in the implementation of SALSA.

I realise different groups will have different ways of naming their modules and I'm not necessarily suggesting the SALSA group come up with a new acronym for the microphysics elements of SALSA. However I do think it needs to be stated somewhere in the section 2.2 description exactly what constitutes the Hamburg Aerosol Model and what are the SALSA aspects. See also my first specific comment about the wording of the title.
You have understood the difference between HAM, M7 and SALSA correct. As explained in Section 2.2, HAM is the aerosol model which calculates aerosol emissions, removal, hydration, and radiative properties. M7 and SALSA are the aerosol microphysics modules of HAM which uses either modal or sectional approach depending on which aerosol microphysical module is used. Most of the actual HAM code is shared with both M7 and SALSA. This is why the model is called ECHAM-HAMMOZ. In addition, the model licence requires that when publishing results with the model, the full name ECHAM-HAMMOZ is used and the exact release reference is stated. We will further clarify this in Section 2.2.

2) Abstract, page 1, lines 9-11 This difference in the modal and sectional aerosol microphysics predictions for the microphysical evolution and global dispersion of the Pinatubo volcanic cloud is interesting, but, as the authors point out, the standard mode widths for M7 are not intended to be applied to the stratospheric aerosol evolution. Figure 14 shows that actually, provided the model is applied with the “stratospheric-enabling adjustment” to the accumulation mode and coarse mode widths, the modal scheme compares well to the sectional scheme. As I understand it, the Hamburg stratospheric aerosol modelling group would not apply the model without this adjustment to the mode widths, so the emphasis really needs to be changed in how this is worded in the Abstract and in the discussion of the results. I think it is very important, to minimise the chance of an incorrect inference from the reader, to present the results having that “M7mod” essentially as the default (or even “validated”?) configuration for when the model is applied for simulating interactive stratospheric aerosol. Indeed I would strongly recommend to change the “branding” of that model run to “M7-strat” rather than “M7-mod”. As I understand it, the adjustment to the mode widths is a pre-requisite for simulating stratospheric aerosol for that scheme, so the authors of the manuscript need to change the current wording of the results to be clearer that it is essentially “the stratospheric configuration of M7” or so. One could consider it in some ways equivalent to a tropospheric or stratospheric chemistry scheme. One would not apply a tropospheric chemistry scheme to simulate the chemistry of the stratosphere.

This is a very good point and we will add an explanation that the modified mode width used in the Pinatubo comparison is the stratospheric setup of M7, which is untested for the troposphere. We will also add references where it has been used: Niemeier et al. (2009, 2011); Niemeier and Timmreck (2015); Niemeier and Schmidt (2017). We will also change the model run name from M7mod to M7-strat following the suggestion of the referee. On the other hand, we need to emphasize that this setup is not a feature in this model release and using it requires code level changes.

List of minor revisions

Referee comments 3, 5, 6, 8, 9, 10, 12, 14-21, 24-29, 32, 36, 39-41, 49, and 50 suggest rewording and changes in the terms used. We will make the changes as the referee suggests in the revised manuscript.
1) Title, page 1, The way the title is currently worded suggests the SALSA2.0 module is as a new sub-model within the Hamburg Aerosol Model (HAM). Is that the case then that HAM includes both M7 and SALSA as alternative aerosol microphysics modules? Or is the SALSA module an alternative to “the overall HAM” or so?

SALSA2.0 is a new submodel and an alternative to M7 within HAM and a component of the release version of ECHAM-HAMMOZ. We will clarify this in the revision.

2) Abstract, page 1, 1st sentence Related to point 1) is ECHAM-HAMMOZ still ECHAM-HAMMOZ when SALSA is applied or should it then be referred to as ECHAM-SALSA-MOZ or so?

As explained above, the model name remains ECHAM-HAMMOZ even with SALSA turned on.

4) Abstract, page 1, 3rd sentence insert ”within ECHAM” or ”within ECHAM-HAMMOZ” between ”implementation” and ”is evaluated” to be clear it is this particular implementation that is evaluated (one could imagine it potentially being implemented in another framework at some point in the future).

We will insert “within ECHAM-HAMMOZ” in this sentence.

7) Abstract, page 2, lines 1-2 as per major comment 1-2, this sentence needs to be changed since (as I understand it) the M7 microphysics would not be applied for stratospheric aerosol applications unless the mode widths for the accumulation and coarse soluble modes were reduced to 1.2 in this way. In this sentence and the results of the Pinatubo comparisons, I this should be referred to as ”the stratospheric aerosol configuration of M7” or similar.

We will refer to the M7 setup with modified mode widths as “the stratospheric aerosol configuration of M7” in the revised manuscript.

11) Introduction, page 2, lines 8-9 The words ”at the lower end of the size spectrum of nanometer size in diameter” somehow seemed a strange wording. The term ”lower end” seemed odd suggest to replace that text above with something more linked to their formation process, replacing ”at the lower...” with ”freshly nucleated particles are observed at nanometer sizes...” then the rest of the sentence can continue with ”as they grow...”. Then similarly instead of ”upper end of the spectrum” suggest ”coarse part of the spectrum”.

It is true that the original wording is too complicated. We will rephrase this to “For example, when the nanometer sized smallest particles grow in size, they contribute to the number of aerosol particles which can form cloud droplets (Kulmala and Kerminen, 2008) while the largest particles of micrometer size affect rain formation (Jensen and Lee, 2008).”

13) Introduction, page 2, lines 11-12 change the start of this setence to be more specific about the size effect you are explaining in simple terms it can be understood simply as
particles only interacting effectively with the radiation once theyre above a certain size. I'd suggest to re-word the sentence to something like "There is a steep size dependence for how effectively aerosol particles interact with radiation (Chung et al., 2005) and clouds (Lohmann and Feichter, 2005)." Suggest also to cite the chapters 7 and 8 of the 2013 IPCC AR5 report rather than the 2005 references given there i.e. Myhre et al. (2013) and Boucher et al. (2013).

The size dependence of particles on radiation and cloud formation is so complex that we do not want to go into more details as it would require a lot of text to explain it comprehensively. This sentence was meant to only briefly mention the size dependency of these effects. We will add the suggested references.

22) Introduction, page 2, line 28 there is also the Piecewise Lognormal Approximation (von Salzen, 2006) which has each size section represented as a log-normal distribution. Please add that as another approach here.

We will add mention this approach together with the reference in the revised version of the manuscript.

23) Introduction, page 3, line 1 need to be more careful with this explanation here. Suggest to re-word the end of this sentence instead to say "the application of sectional models in global 3-D simulations often involves a trade-off with horizontal or vertical resolution" or similar.

Here we mean to say that in global 3D models so many other processes than aerosol microphysics affect the atmospheric aerosol properties that the improvement due to a higher aerosol size resolution is not evident. We will clarify this in the revised manuscript. With SALSA we do not have a trade-off with horizontal or vertical resolution compared since they are the same for SALSA and M7.

30) Section 2.2. lines 11-13 The paper has not quite explained what is the distinction between HAM and M7. Until reading this I thought they were the same thing but I think I now understand that "HAM" is the overall aerosol module (including emissions, dry deposition, scavenging etc.) whereas M7 is just the aerosol microphysical routines. Am I understanding that correctly? If so this needs to be stated explicitly somewhere here in so-doing it will help ensure the community apply the acronyms correctly and consistently in future.

You have understood it correctly. It seems to be a common misunderstanding that M7 and HAM are synonymous. We will clarify the differences in the revised manuscripts.

31) Page 5, section 2.2 line 20 suggest to replace "represents several real-life compounds" with "represents several specific single-species compounds" if that is what is intended?

We will replace “real-life compounds” with “individual chemical compounds”.

6
33) Page 6, Table 1 need to add additional entries to the "emissions" section for primary carbonaceous and give the different size assumptions for emitted primary carbonaceous particles from biomass burning, bio-fuel and fossil-fuel sectors. As per my major comment 1, I think this is the primary reason why there is the AOD different in those strong emissions regions. You can see that the BC size distribution is at different sizes in the sectional and modal scheme, and I think this can simply be explained by a different size assumption. I would be very suprised if that was caused by microphysical processing.

We will describe more in detail the emission size distributions in the revised manuscript.

34) Page 7, line 7 be clear what you mean by "coupled" you mean "radiatively-coupled" right? Need to add an extra sentence briefly explaining how that's done here for aerosol-radiation interactions and aerosol-cloud interactions in the sectional scheme (and how that differs from the radiative coupling when the modal scheme is used).

Here we mean that we run ECHAM-HAMMOZ with SALSA2.0 aerosol microphysics. We will rephrase this accordingly in the revised manuscript.

35) Page 7, line 10 you write "we used the climatologies" but I don't think you mean climatologies here do you? What is the time-variation of the specified SST and sea-ice distributions?

We do use sea surface temperature and sea ice cover climatologies. They are monthly fields which we will clarify in the revised manuscript.

37) Page 8, line 16 you write "For most of the processes the difference is only in the numerical treatment" but that's not quite right the nucleation processes are different as shown in Table 1 please change this wording.

We will rephrase this part as follows: “In the default setups of M7 and SALSA2.0, wet deposition and secondary organic aerosol (SOA) formation are the only processes (in addition to the calculation of aerosol microphysics) that use different methods for solving the physics of the process. For the rest of the processes the difference is only in the numerical treatment.”.

38) Page 8, line 25 you write "more detailed size-dependent scavenging rates" but you need to add a few qualifying words so the reader knows what you mean by "more detailed" here. The reader might expect the sectional SALSA scheme to have more detailed scavenging than the modal scheme or maybe you don't mean detailed in a size-resolved way do you mean the way the scavenging applies different scavenging efficiency for the different types of precipitating cloud?

We will modify this part to indicate that the wet deposition scheme is more physically based instead of using the ambiguous term “detailed”.

42) Page 9, line 9 As per my major comment 2, it is not fair to refer to the initial settings of the scheme as "The default settings". They are indeed the default settings for tropospheric aerosol simulations, but they are not the default settings for stratospheric
aerosol simulations. As per my major comment 2 please change the branding of these two M7 simulations from "M7" and M7mod" to "M7-trop" and "M7-strat". They are alternative configurations of M7 specifically for those applications. It's fine to show that simulation with the tropospheric configuration of M7 in fact that will show why it's important to only apply the scheme in the stratosphere with the stratospheric configuration (M7-strat). But you need to change the wording so that it's clear that this is only default for tropospheric aerosol application of the model.

This is true and we will modify this part as explained in the reply to the major comments.

43) Page 9, line 11 The authors write "This is because the high concentration of sulfur produces a bi-modal aerosol population". Is this statement referring to the Laramie balloon-borne OPC measurements (Deshler et al., 2003) which show the bimodal size distribution after Pinatubo? If so please give that reference here.

Here we refer to the model study (Kokkola et al., 2009), but we will also include the Deshler reference.

44) Page 9, line 12 the narrowing of the width again I think you are referring to what is observed from the measurements right? That is the case that the accumulation mode is observed to have a narrower size distribution cite Deshler et al. (2003) or Deshler (2008).

Here we refer to the detailed aerosol microphysics model which was used in the Kokkola et al. (2009) paper.

45) Page 9, line 14-16 you're referring to the box model simulations here, right? Its not so clear how the effect plays out in 3D simulations, and more so when you consider the trade-off in the better stratospheric circulation that can be afforded (by resolving more vertical levels for example) with a computationally faster aerosol scheme. So you need to be clear that you're referring here to what is seen in a box model. For a balanced discussion of this, you also need to add a qualifying sentence explaining this trade-off between the cost of the aerosol scheme and the cost of the atmosphere model.

Yes, we are referring to the box-model study Kokkola et al. (2009). However, the same effect can be seen in our global simulations where we can see that the largest particles are removed faster in the tropospheric setup M7 and is evident in the manuscript Figure 14b. The difference in the computational speed between M7 and SALSA is not so large that they would use different grid resolution. We will however mention the increase in the computational burden when using SALSA.

46) Page 9, line 15 you need to re-word "grows too fast and the particles are sedimented too fast". The box model shows that in those simulations the growth proceeds faster, but you do need to qualify the 2nd part with "which would result in particles sedimenting faster" or something like this. Since it has not really been demonstrated in global models you need to tone down the way that is described.
We will modify this sentence as suggested. However, we investigated the difference between the growth of particles between the two M7 simulations. With the default mode widths, the coarse particle burden grows quicker right after the eruption and these particles are consequently removed much faster than in the case of using narrower modes.

47) Page 9, line 17 It is not appropriate to refer to this as "A work-around solution". The "code-owners" of the M7 scheme are clear in their publications that when the scheme is applied for stratospheric aerosol applications, the modal settings need to be configured differently than for tropospheric aerosol applications. That's not correct to refer to that as a work-around. Effectively the scheme is only "licensed" to be applied in the stratosphere if it has this adjustment to the modal settings. As per my major comment 2 this section needs to be re-worded to make this clear in my strong opinion, for the reasons above, you should refer to the tropospheric aerosol and stratospheric aerosol configurations of M7, and label them as "M7-trop" and "M7-strat". That is then consistent with the way the owners of the adjusted scheme have re-configured the model to be applicable for the stratosphere.

We will change the label to M7-strat. However, it has to be noted that the model version we present here requires code level changes to start using the alternative stratospheric aerosol configuration. It also has to be noted that there is no release version of ECHAM-HAMMOZ, which would support easy switch to this configuration. We will also remove the term "work-around".

48) Page 9, line 22 the Guo et al. (2004) has the SO2 emissions range as 14 to 23 Tg of SO2 you need to give that range (and any widening of that to include values from other publications).

We will give this range in the revised manuscript.

References


Lee, L. A., Carslaw, K. S., Pringle, K. J., Mann, G. W., and Spracklen, D. V.: Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters, Atmo-


SALSA2.0: The sectional aerosol module of the aerosol-chemistry-climate model ECHAM6.3.0-HAM2.3-MOZ1.0

Harri Kokkola¹, Thomas Kühn¹,², Anton Laakso¹,³, Tommi Bergman⁴, Kari E. J. Lehtinen¹,², Tero Mielonen¹, Antti Arola¹, Scarlet Stadtler⁵, Hannele Korhonen⁶, Sylvaine Ferrachat⁷, Ulrike Lohmann⁷, David Neubauer⁷, Ina Tegen⁸, Colombe Siegenthaler-Le Drian⁹, Martin G. Schultz⁵,¹⁰, Isabelle Bey⁹,¹¹, Philip Stier¹², Nikos Daskalakis¹³, Colette L. Heald¹⁴, and Sami Romakkaniemi¹

¹Atmospheric Research Centre of Eastern Finland, Finnish Meteorological Institute, P.O.Box 1627, FI-70211 Kuopio, Finland.
²Aerosol Physics Research Group, University of Eastern Finland, P.O.Box 1627, FI-70211 Kuopio, Finland.
³Department of Soil, Water and Climate, University of Minnesota, Twin Cities, St. Paul, MN 55108, USA
⁴Weather and Climate Models, Royal Netherlands Meteorological Institute, P.O.Box 201, 3730AE De Bilt, the Netherlands
⁵Institut für Energie- und Klimaforschung, IEK-8, Forschungszentrum Jülich, Germany
⁶Climate Research, Finnish Meteorological Institute, Helsinki, FI-00100, Finland
⁷Institute for Atmospheric and Climate Science, ETH Zurich, 8092 Zurich, Switzerland
⁸Modeling of Atmospheric Processes, Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany
⁹Centre for Climate Systems Modeling (C2SM), ETH Zürich, Switzerland
¹⁰Jülich Supercomputing Centre, JSC, Forschungszentrum Jülich, Germany
¹¹Centre météorologique de Genève, Office fédéral de météorologie et de climatologie MétéoSuisse, av. de la Paix 7bis, CH-1211 Genève 2, Switzerland
¹²Department of Physics, University of Oxford, Parks Road, OX1 3PU, UK
¹³Laboratory for Modeling and Observation of the Earth System (LAMOS), Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany
¹⁴Department of Civil and Environmental Engineering, Department of Earth, Atmospheric, and Planetary Science, Massachusetts Institute of Technology, Cambridge, MA 02139

Correspondence to: Harri Kokkola (harri.kokkola@fmi.fi)

Abstract. In this paper, we present the implementation and evaluation of the aerosol microphysics module SALSA2.0 in the framework of the aerosol-chemistry-climate model ECHAM-HAMMOZ. It is an alternative microphysics module to the default modal microphysics scheme M7 in ECHAM-HAMMOZ. The SALSA2.0 implementation within ECHAM-HAMMOZ is evaluated against the observations of aerosol optical properties, aerosol mass, and size distributions, comparing also to the skill of the M7 implementation. The largest differences between SALSA2.0 and M7 are evident over regions where the aerosol size distribution is heavily modified by the microphysical processing of aerosol particles. Such regions are, for example, highly polluted regions and regions strongly affected by biomass burning. In addition, in a simulation of the 1991 Mt Pinatubo eruption in which a stratospheric sulfate plume was formed, the global burden and the effective radii of the stratospheric aerosol are very different in SALSA2.0 and M7. While SALSA2.0 was able to reproduce the observed time evolution of the global burden of sulfate and the effective radii of stratospheric aerosol, M7 strongly overestimates the removal of coarse stratospheric particles and thus underestimates the effective radius of stratospheric aerosol. As the mode widths of M7 have been optimized for the
troposphere and were not designed to represent stratospheric aerosol the ability of M7 to simulate the volcano plume was improved by modifying the mode widths decreasing the standard deviations of the accumulation and coarse modes from 1.59 and 2.0, respectively, to 1.2 similar to what was observed after the Mt Pinatubo eruption. Overall, SALSA2.0 shows promise in improving the aerosol description of ECHAM-HAMMOZ and can be further improved by implementing methods for aerosol processes that are more suitable for the sectional method, e.g size dependent emissions for aerosol species and size resolved wet deposition.

1 Introduction

Describing the global physical and chemical properties of the atmospheric aerosol in atmospheric models is challenging due to their large spatial and temporal variability in these properties. The diameter of the particles span several orders of magnitude and the chemical composition can include hundreds of compounds (e.g., Colbeck et al., 2014). For example, when the nanometer sized smallest particles grow in size, they contribute to the number of aerosol particles which can form cloud droplets (Kulmala and Kerminen, 2008) while the largest particles of micrometer size can also affect rain formation (Jensen and Lee, 2008). For example, at the lower end of the size spectrum, particles of nanometer size in diameter, as they grow in size, contribute to the number of aerosol particles which can form cloud droplets (Kulmala and Kerminen, 2008) while at the upper end of the spectrum, particles larger than one micrometer in diameter affect rain formation (Jensen and Lee, 2008). Particles of different sizes affect both atmospheric radiation (Chung et al., 2005) and cloud processes (Lohmann and Feichter, 2005) in different ways (Boucher et al., 2013; Myhre et al., 2013). Therefore, in order to accurately simulate the effects of aerosol on the global climate, the entire aerosol particle size spectrum must be represented. In addition to the particle size, the chemical composition of particles, i.e. in particular the absorption (Dubovik et al., 2002) and solubility/hygroscopicity (Che et al., 2016) vary strongly between different aerosol constituents, influencing their ability to affect the aerosol constituents, also influence radiation and cloud interactions in the atmosphere. In order to properly simulate these aerosol effects, the composition should also be adequately represented in the models.

This multitude of variability in the physical and chemical properties of aerosols poses a challenge for global modellers to describe aerosol particles the atmospheric aerosol in a computationally efficient way. Simulating the aerosol size distribution at high resolution including size resolved detailed chemical composition within hundreds of thousands of grid boxes is computationally challenging. However, solving the size-resolved evolution of atmospheric particles computationally efficiently is not a new challenge as such simulations were made in the early years of computational atmospheric physics (e.g., Young, 1974). Currently, most of the global models which describe the evolution of the aerosol size distribution resort to using either modal or sectional approaches or a mix of these two (e.g., Mann et al., 2014). The application of sectional models in global 3-D simulations can involve a trade-off with horizontal or vertical resolution because sectional models are computationally more expensive.

Essentially, modal and sectional approaches can be considered as two variants of the same method, as both approaches divide the aerosol size distribution into size classes. The modal approach assumes individual size classes (modes) to be log-
normally distributed and the total aerosol size distribution to be a superposition of these modes (e.g., Vignati et al., 2004; Stier et al., 2005). In the sectional approach, the size classes are either assumed to be monodisperse (Zaveri et al., 2008) or they are assumed to have a linear size distribution within a section (Young, 1974; Stevens et al., 1996) or a piecewise log-normal approximation within a section is used (von Salzen, 2006). The modal setup is usually computationally more efficient since the number of size classes needed to represent typically observed size distributions is much smaller than in the sectional approach. Typically modal models use seven or fewer modes while sectional models use up to 100 size classes (Mann et al., 2014; Yu and Luo, 2009). On the other hand, sectional models allow for more flexibility in e.g. the shape of the size distribution and volume distribution of chemical compounds (Kokkola et al., 2009). Although sectional models have been shown to perform significantly better than modal models in 0-D and 2-D frameworks (Weisenstein et al., 2007; Kokkola et al., 2009; Korhola et al., 2014) the benefits of sectional models in global 3-D simulations are less evident (Mann et al., 2012, 2014). It is also difficult to quantify the benefit of the sectional approach because this is mainly because the comparison between modal and sectional models are, in most cases, not done within the same model framework and the structural differences in the models cause such a large difference in the modeled aerosol that the contribution to the differences from the choice of the size distribution scheme cannot be identified (Mann et al., 2014). Another reason is that the evaluation of the skill of global aerosol models against observations is extremely challenging as the model value for a given observable may not represent the measured value at a particular monitoring site, a given parameter may not represent the observed value e.g. at a particular measurement site (Schutgens et al., 2016). This discrepancy can for example be caused by the fact that the global model value represents the mean for a grid box ∼200 km × 200 km in size. Aerosol properties can exhibit large variations within that area and the measurement site may not represent the mean conditions within that grid box.

Here we present the implementation of the sectional aerosol microphysics module SALSA (Kokkola et al., 2008) in the aerosol-chemistry-climate model ECHAM-HAMMOZ (echam6.3-ham2.3-moz1.0) which also includes the modal aerosol microphysics module M7 (Vignati et al., 2004). The paper is structured as follows. In Section 2 we present the details of individual model components, especially the methods for solving aerosol processes. In Section 3 we briefly present the model to be analyzed simulations that were made with the different models/configurations different model configurations. In Section 4 we present the evaluation of the model against observations. The performance of the model is evaluated using retrievals of aerosol optical properties from both satellite and ground based remote sensing instruments. We also compare the model with in-situ observations, including vertical profiles of aerosol composition and mass from aircraft measurements in situ observations as well as aircraft observations of aerosol composition and mass. Finally, we compare the sectional model results with those obtained from ECHAM-HAMMOZ in modal aerosol configuration. The ECHAM-HAMMOZ model framework allows for running simulations in an otherwise very similar global model setup, but only switching between the modal and sectional aerosol representations. This comparison provides insights on the impacts of the representation of the aerosol size distribution on the simulated aerosol properties, and thus on the simulated climate and climate effects.
2 Model description.

2.1 ECHAM

The host atmospheric model in ECHAM-HAMMOZ is the sixth generation atmospheric general circulation model ECHAM6. The details of the model have been described by Stevens et al. (2013). It is the atmospheric component of the Max Planck Institute for Meteorology Earth System Model (MPI-ESM) and was originally based on the European Centre for Medium Range Weather Forecasts (ECMWF) weather prediction model (Simmons et al., 1989). The dynamical core applies the spectral method for calculating the atmospheric circulation and flux form semi-Lagrangian transport scheme. In our model configuration, we use the T63 spectral truncation for the horizontal grid, with 47 flexible vertical levels which follow the terrain and use the hybrid vertical coordinate representation described in detail by Roeckner et al. (2003).

In atmosphere-only simulations, ECHAM6 uses prescribed sea surface temperatures (SST) and sea ice cover (SIC). The land processes are calculated using the JSBACH -model (Raddatz et al., 2007) which is integrated in ECHAM6. The aerosol processes are simulated by the HAMMOZ aerosol-chemistry model (Schultz et al., 2017).

2.2 HAMMOZ

The aerosol-chemistry model HAMMOZ combines the Hamburg Aerosol Model (HAM) and the Model for OZone And Related chemical Tracers (MOZART) which simulates the chemical reactions of 300 species through 650 reactions in the troposphere and stratosphere. A more detailed description of MOZ and its implementation in ECHAM-HAMMOZ is given in the accompanying paper by Schultz et al. (2017). Please note that in the simulations made for this paper, we did not use MOZ in any of the simulations. Instead, sulfate chemistry is calculated in the more simplified scheme of HAM (Zhang et al., 2012).

HAM will also be presented in detail in another accompanying paper by Tegen et al. (2018). However, as SALSA is integrated within HAM and as SALSA incorporates many of the model design characteristics of HAM, we briefly introduce the aerosol related features of HAMMOZ and detail the coupling between HAMMOZ and SALSA.

The HAM aerosol model has been designed to simulate all tropospherically relevant aerosol processes, the interactions between aerosol and radiation, and the interactions between aerosol and clouds (Stier et al., 2005; Zhang et al., 2012). It includes two options for the calculation of microphysics, the modal microphysics module M7 and the sectional microphysics module SALSA2.0 which was implemented in this study. The model design has been optimized for computational efficiency together with solving aerosol processes accurately. In its default setup, HAM uses the modal approach together with the model aerosol microphysics module M7 (Vignati et al., 2004). HAM applies the modal scheme which has been developed together with the development of the aerosol microphysics module M7 Vignati et al., 2004. In this modal setup, the aerosol size distribution is described by a superposition of seven log-normal modes. Chemical components incorporated in each mode are chosen so that only those compounds which are relevant in the real atmosphere for each size range of each mode are included in those modes. The external mixing of aerosol is considered such that the soluble and insoluble compounds are emitted in separate parallel modes and as the insoluble modes are aged (i.e. soluble compounds accumulate on insoluble modes) insoluble modes are merged into the soluble modes. The chemical compounds in HAM can be considered as compound classes in the sense...
that they group certain types of aerosols to model compounds. These compounds are “sulfate” (SU), “organic aerosol” (OA), “sea salt” (SS), “black carbon” (BC), and “mineral dust” (DU). In practice, each individual model compound represents several individual real-life compounds and especially OA represents hundreds of different organic compounds (Kanakidou et al., 2005). Using lumped component classes is a fairly standardized practice in global aerosol models and the model components are usually the same in most models (Mann et al., 2014). The exception are organic compounds which are often separated e.g. based on their formation mechanisms, i.e. primary and secondary organic aerosol (Tsigaridis et al., 2014).

Processes and properties related to the aerosol particles which are simulated by HAMMOZ are: emissions, dry deposition, wet deposition, sulfur chemistry, sedimentation, radiative properties, microphysical processes, and relative humidity in the cloud-free part of the grid cells. HAMMOZ simulated aerosol are also coupled to the ECHAM-HAMMOZ cloud scheme and affect liquid cloud droplet formation and ice crystal formation (see Lohmann et al. (2007)).

In the default model configuration, all of these processes are calculated using the modal approach and the microphysics are calculated using the M7 module. Thus, the implementation of SALSA requires also the modification of HAM routines to follow the sectional representation and allow for consistent representation of these processes for modal and sectional approaches.

2.3 SALSA

The aerosol microphysical model SALSA is designed to be applicable to different scales of aerosol modeling starting from 0-dimensional simulations of laboratory or chamber experiments (Kokkola et al., 2014). It has also been implemented in the large eddy simulations (LES) model UCLALES (Tonttila et al., 2017) for 1-, 2-, 3-dimensional simulations. SALSA has also been implemented in the chemical transport model MATCH (Andersson et al., 2015) which in turn has been coupled to the regional climate model RCA4 (Thomas et al., 2015). This scalability and usage of one model across different scales allows for the easy parameterization of small scale aerosol processes up to the global scale. On the global scale, SALSA has previously been implemented in ECHAM5-HAM (Bergman et al., 2012). Here we present the configuration of SALSA which has been implemented in ECHAM-HAMMOZ and builds upon the implementation of SALSA in ECHAM5-HAM. For clarity, in this section, the ECHAM5-HAM implementation is called SALSA1 and the one implemented in ECHAM6-HAMMOZ is called SALSA2.0.

SALSA represents the aerosol size distribution using the sectional approach. The size distribution is divided into 10 size classes using the volume ratio discretization (Jacobson, 2005). However, the width of the size classes vary over three size ranges: Subrange 1 for particles with diameters $D_p = 3\,\text{nm} - 50\,\text{nm}$, Subrange 2 for $D_p = 50\,\text{nm} - 700\,\text{nm}$, and Subrange 3 for $D_p = 700\,\text{nm} - 10\,\mu\text{m}$. This separation was done so that the size resolution is highest in the accumulation mode sizes which increases the accuracy of the cloud activation calculations. For each size class the tracer variables are: the number of particles and the concentration of individual species.

In SALSA1, Subrange 1 assumed internal mixing for all sizes, Subrange 2 included two externally mixed size classes (soluble and insoluble), and Subrange 3 included three externally mixed size classes (soluble, fresh insoluble and aged insoluble). In addition, the number of chemical compounds varied between the three size ranges. In SALSA2.0, the width of the size bins
remains unchanged from SALSA1. However, subranges 2 and 3 are now treated as one so that the combined size range includes two externally mixed size classes; one where the insoluble compounds are emitted and one where the soluble compounds are emitted. These subregions are visualized in Figure 1. The change in how the chemical compounds are treated was first of all due to practical reasons. In SALSA1, the information of individual species was lost when the particles grew to sizes larger than 700 nm in diameter. This caused problems in studies where the information of individual species in all particle sizes was required (e.g. Kipling et al., 2016). Second, although in the troposphere microphysics have very little influence on the size of particles in the third subregion, when simulating volcanic eruptions or stratospheric solar radiation management, condensation can grow the largest particles. This caused the model to have problems in simulating the growth of particles in a volcano plume since the third region particles did not grow. This in turn resulted in underestimation of the effective radius of the volcano plume (Kokkola et al., 2009). In 0-dimensional model tests (not shown here), SALSA2.0 did not exhibit such problems.

Another significant change between SALSA1 and SALSA2.0 has been the modification of the aerosol size distribution update routine. In SALSA1, the moving center method (Jacobson and Turco, 1995) was used for Subranges 1 and 2, and the

Figure 1. Schematic of the number (\(N\)) size distribution representation as a function of diameter \(D_p\) in SALSA1 (upper size distribution) and SALSA2.0 (lower size distribution). The color of each size class indicates which compounds are included in the size class.
fixed sectional method (Gelbard et al., 1980) for Subrange 3. In SALSA2.0 the hybrid bin method (Young, 1974; Chen and Lamb, 1994) is used for all size sections. This is because the moving center method has been shown to introduce numerical artefacts in zero dimensional box model simulations (Mohs and Bowman, 2011) and when simulating aerosol formation and growth in high sulfur concentration conditions typical for e.g. large volcanic eruptions and simulations of stratospheric solar radiation management (e.g., Kokkola et al., 2008, Figure 2). In addition, in the study by Bergman et al. (2012), SALSA1 underestimated aerosol number concentrations observed at ground stations when using the moving center method. Switching to the hybrid bin method decreased the low bias.

The implementation of SALSA2.0 in ECHAM-HAMMOZ was designed such that it shares the routines with the modal scheme of M7 wherever possible. In the sense of model processes, the biggest difference is in the aerosol microphysical calculations which are treated using methods that are designed for the respective size distribution description. The microphysical processes and other aerosol processes that are treated differently between the two model configurations are listed in Table 1. A comprehensive review of the relative importance of these processes within the ECHAM framework has been given previously by Schutgens and Stier (2014).

For offline emissions of SU, OA, and BC, SALSA2.0 uses the emission size distributions of M7 which are remapped to SALSA2.0 size sections. The details of these emission size distributions for different chemical compounds and emission sectors are given in Zhang et al. (2012) (Table 2). Emission size distributions for offline emissions of SU, OA, and BC follow the emission size distribution used in the modal representation (Zhang et al., 2012) remapped to SALSA2.0 size sections. Online emissions for SS, and DU are calculated online according to Long et al. (2011) and Cheng et al. (2008), respectively.

### 3 Model simulations

As the base simulation, we run ECHAM-HAMMOZ with coupled SALSA2.0 for a ten year period (2003-2012) which was preceded by a one year spin-up period. The large scale meteorology (vorticity, divergence, and surface pressure) was nudged towards the ECMWF (European Centre for Medium-Range Weather Forecasts) reanalysis data ERA Interim (Berrisford et al., 2011). The relaxation times for the nudging of the surface pressure, vorticity, and divergence are 24 h, 6 h, and 48 h, respectively. For sea surface temperatures and sea ice distributions we used the monthly mean climatologies from the PCMDI’s Atmospheric Model Intercomparison Project (AMIP) of the Program for Climate Model Diagnosis & Intercomparison (PCMDI) http://www-pcmdi.llnl.gov/projects/amip/ (Taylor et al., 2012). The mass emission fluxes of each aerosol species from anthropogenic sources are based on Aerocom II - ACCMIP emissions (Lamarque et al., 2010) which for the period 2000-2100 have been linearly interpolated to the Representative Concentration Pathways (RCP) projection RCP4.5 (van Vuuren et al., 2011). For the mass emission fluxes of individual species from biomass burning we used the GFASv1 database multiplied by a factor of 3.4 following the recommendation by Kaiser et al. (2012). Emissions of OA from biogenic sources, were based on the Aerocom I monoterpene emissions (Dentener et al., 2006) of which 15% was assigned to the particle phase OA mass. For the terrestrial emissions of dimethylsulfide (DMS) we used the Pham et al. (1995) emission dataset and the oceanic DMS emissions were calculated online according to Kloster et al. (2006).
Table 1. Overview of the treatment of different aerosol processes in the sectional approach (SALSA) and the modal approach (M7) when using the default setup. Processes marked with * have been modified for the sake of a unified evaluation between SALSA2.0 and M7 and are detailed in the body text.

<table>
<thead>
<tr>
<th>Process</th>
<th>SALSA2.0</th>
<th>M7</th>
</tr>
</thead>
<tbody>
<tr>
<td>microphysical process:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nucleation</td>
<td>activation type nucleation (Sihto et al., 2006)</td>
<td>neutral and charged nucleation of H$_2$SO$_4$ and H$_2$O (Kazil and Lovejoy, 2007)</td>
</tr>
<tr>
<td>condensation of H$_2$SO$_4$</td>
<td>analytical predictor of condensation solved simultaneously with nucleation (Jacobson, 2005)</td>
<td>Two-step operator splitting scheme with an analytical solution for production and condensation (Kokkola et al., 2009)</td>
</tr>
<tr>
<td>coagulation</td>
<td>semi-implicit method (Jacobson and Turco, 1995)</td>
<td>implicit method (Vignati et al., 2004)</td>
</tr>
<tr>
<td>hydration</td>
<td>ZSR method (Stokes and Robinson, 1966)</td>
<td>$\kappa$-Köhler (Petters and Kreidenweis, 2007)</td>
</tr>
<tr>
<td>SOA formation</td>
<td>volatility basis set (Donahue et al., 2006) with three volatility classes, partitioning calculated solving the condensation equation according to the analytical predictor of condensation (Jacobson, 2005). SOA partitioning routine in SALSA has been described by Kokkola et al., 2014 *</td>
<td>thermodynamical-equilibrium (Zhang et al., 2012) *</td>
</tr>
<tr>
<td>emissions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sea Salt</td>
<td>Size segregated sea salt emissions from Long et al. (2011) parameterization mapped to the soluble size sections in subrange 2 following the M7 mode parameters for accumulation and coarse modes.</td>
<td>Size segregated sea salt emissions from Long et al. (2011) parameterization mapped to the soluble accumulation and coarse modes</td>
</tr>
<tr>
<td>Mineral Dust</td>
<td>Size segregated mineral dust emissions from Cheng et al. (2008) parameterization mapped to the insoluble size sections in subrange 2 following the M7 mode parameters for accumulation and coarse modes.</td>
<td>Size segregated mineral dust emissions from Cheng et al. (2008) parameterization mapped to insoluble accumulation and coarse modes</td>
</tr>
<tr>
<td>radiative effects</td>
<td>Look-up-tables which are based on mie calculations for the extinction cross section, asymmetry factor, and single scattering albedo as a function of Mie size parameter and refractive index. Size sections are assumed to have a “flat top” size distribution within bins</td>
<td>Look-up-tables which are based on mie calculations for the extinction cross section, asymmetry factor, and single scattering albedo as a function of Mie size parameter and refractive index. Look-up-tables have been pre-calculated separately for modes with geometric standard deviation of 1.59 and 2.0</td>
</tr>
<tr>
<td>below-and in-cloud scavenging</td>
<td>Prescribed scavenging coefficients for each size section according to Bergman et al. (2012)</td>
<td>Prescribed impaction scavenging coefficients for each mode according to (Stier et al., 2005) or size dependent scavenging rates according to (Croft et al., 2009, 2010).</td>
</tr>
</tbody>
</table>
The model output consisted of instantaneous values at 3 h interval. Although ECHAM-HAMMOZ includes the explicit chemistry model MOZ, it was not used in these simulations. Instead, we used the simplified sulfur chemistry scheme of HAM (Feichter et al., 1996; Zhang et al., 2012). The module calculates the oxidation of DMS and SO$_2$ by OH, H$_2$O$_2$, NO$_2$, and O$_3$ in the gas and the aqueous phase. The oxidant concentrations are prescribed using monthly mean 3-dimensional fields from the MOZART chemistry model simulation (Horowitz et al., 2003).

In order to evaluate how the sectional approach performs against the modal approach within the same atmospheric model, we repeated the simulations for the year 2010 using M7 as the aerosol microphysical module with a setup as similar as possible. In the default setups of M7 and SALSA2.0, wet deposition and secondary organic aerosol (SOA) formation are the only processes (in addition to the calculation of aerosol microphysics) that use different methods for solving the physics of the process. For the rest of the processes the difference is only in the numerical treatment. For most of the processes the difference is only in the numerical treatment. In the default setups of M7 and SALSA2.0, wet deposition and secondary organic aerosol (SOA) formation are the only processes (in addition to the calculation of aerosol microphysics) that use different methods for solving the physics of the process. To minimize the differences between simulations done with the sectional and modal versions, the wet deposition scheme for M7 was changed to use the same prescribed wet scavenging coefficients as were used for SALSA2.0 (see Table 1). These coefficients have also been used in M7 in previous versions of ECHAM-HAMMOZ (Stier et al., 2005; Zhang et al., 2012). The implementation of the M7 scavenging coefficients for SALSA1 size sections has been presented by (Bergman et al., 2012). The reason for using the older approach is that the implementation of an improved wet scavenging scheme in SALSA2.0 is still under development. However, in order to compare the significance of microphysical processing and wet deposition on the modeled aerosol, we ran one additional simulation for the year 2010 with M7 using the more physically based detailed size dependent scavenging rates (Croft et al., 2009), i.e. the default configuration of ECHAM-HAMMOZ. On the other hand, it should be noted that a comprehensive evaluation of the default version of ECHAM-HAMMOZ with M7 will be given in a separate paper by Tegen et al. (2018) and thus we do not do a full evaluation of it here.

In addition to the wet deposition scheme, we also turned off the SOA formation routine to keep the model configurations similar in the evaluation. The SOA schemes are very different in their approach, as M7 assumes equilibrium partitioning for SOA while SALSA2.0 calculates SOA partitioning kinetically, solving size resolved condensation equations. The SOA scheme will be presented in detail by a companion paper by Kuhn et al. (in preparation). Instead of the detailed SOA schemes detailed in Table 1, for both SALSA2.0 and M7 we used the Aerocom I monoterpane emissions (Dentener et al., 2006) of which 15% was irreversibly assigned to the particle phase OA mass.

As the sectional method requires more tracer variables for representing aerosol size dependence, SALSA2.0 is computationally slower than M7. The computation time depends very much on the time interval and the number of output variables. With Cray XC 30 architecture using 120 CPU cores, the evaluation simulations of SALSA2.0 took approximately double the time of M7.
3.1 Pinatubo experiment

Previous 2D (Herzog et al., 2004; Weisenstein et al., 2007) and box model (Kokkola et al., 2009) studies have shown that the modal approach, especially when the mode width is prescribed, can not reproduce aerosol growth when the concentration of condensing species is very high (Weisenstein et al., 2007; Kokkola et al., 2009). This can be the case in simulating stratospheric sulfur solar radiation management by injecting sulfur in the stratosphere or in the case of strong volcanoes which emit high concentrations of sulfur into the stratosphere. Using the default mode width of M7 in high sulfur concentrations the growth of the aerosol effective radius is too rapid and leads excessive removal of stratospheric aerosol by sedimentation (Kokkola et al., 2009). This is because the high concentration of sulfur produces a bi-modal aerosol population seen in model simulations (Kokkola et al., 2009) and observations after the Mt Pinatubo eruption (Deshler et al., 2003; Deshler, 2008). The width of the aerosol size distribution is narrow because the smaller the particles are the faster they grow by condensation as the surface-to-volume ratio increases with decreasing particle size (Turco and Yu, 1999). Such size distributions were also observed after the Pinatubo eruption (Deshler et al., 1997). If prescribed widths are used for the modes, the volume mean diameter, i.e. the diameter that dictates the sedimentation velocity of the modes, grows fast resulting in particles sedimenting fast grows too fast and the particles are sedimented too fast from the stratosphere (Kokkola et al., 2009).

An alternative approach for M7 in simulations of high stratospheric sulfur load is to change the geometric standard deviation to 1.2 in the accumulation mode and remove the coarse mode. This modal setup has been shown to improve the ability of the model to reproduce the aerosol growth in high sulfur stratospheric conditions (Kokkola et al., 2009) and has been used in several studies related to stratospheric aerosol (Niemeier et al., 2009, 2011; Niemeier and Timmreck, 2015; Niemeier and Schmidt, 2017). However, we have to emphasize that such a setup is not a feature in the release version ECHAM6.3.0-HAM2.3-MOZ1.0 and using such a setup would require code level changes and obtaining suitable look-up-tables for the radiation calculations.

One commonly used test case (see e.g., English et al., 2013; Laakso et al., 2016; Timmreck et al., 2018) to evaluate how models perform in simulating high sulfur conditions is the eruption of Mt Pinatubo (15.14° N, 120.35° E) in 1991. It has been estimated that the volcano emitted approximately 14 to 23 Tg S of SO2 at 24 km altitude (Read et al., 1993; Guo et al., 2004). Here we used the mean of this range (8.5 Tg S). The oxidation of emitted SO2 and the consequent new particle formation and growth of sulfate particles perturbed the stratospheric aerosol layer for over 3 years produced stratospheric aerosol that persisted in the stratosphere for over 3 years (Read et al., 1993; Guo et al., 2004).

To investigate how our model reproduces the aerosol properties of the Mt Pinatubo eruption, we ran three sets of transient (no nudging) simulation ensembles (5 ensemble members per set) using: SALSA2.0, M7, and M7 with 1.2 geometric standard deviation for the accumulation and coarse mode (denoted as M7\text{stratmod}). This modification applies also to the tropospheric aerosol. For each model configuration, the ensemble consisted of five 30 month simulations that were preceded by a one year spin up. In each ensemble run, we have perturbed offline anthropogenic aerosol emissions by values of the order of $10^{-6}$ which
is an insignificant number for emission strengths but due to the chaotic nature of the atmospheric model, changes the model dynamics.

The emission settings are identical to those used by Niemeier et al. (2009) and Laakso et al. (2016). In addition, to see how much the current model differs from the previous generation model, we also included simulated aerosol properties from a MAECHAM5-SALSA simulation (Laakso et al., 2016) where the name MAECHAM5 refers to the middle-atmosphere configuration of ECHAM5. In this model setup, SALSA1 was modified so that subregion 2 was extended to cover subregion 3, similarly to SALSA2.0 in order to properly simulate the growth of the particles in high sulfur conditions (see Section 2.3).

4 Results

4.1 Aerosol optical properties

Satellite observations provide the best global coverage of aerosol optical properties and thus comparing the model with satellite retrievals gives a good indication of how the models perform in reproducing regional aerosol characteristics. Here we compared simulated aerosol optical depths (AOD) with those retrieved from the moderate-resolution imaging spectroradiometer (MODIS) instrument on board of both Aqua and Terra satellites (King et al., 1999).

The ground based sun photometers also provide good coverage of observations of aerosol optical properties. Although they are column measurements covering a much smaller area than satellites, they are often considered as the “ground truth” of aerosol properties as they are less affected by the uncertain surface reflectance. Here we used the AOD retrievals from the sun photometer network AERONET (AErosol RObotic NETwork (Holben et al., 1998)) to evaluate the modeled aerosol optical properties.

4.1.1 Evaluation against MODIS observations

The model versus MODIS evaluation was made for the year 2010. From MODIS, we used the level 2.0 combined product of Deep Blue and Dark Target retrievals for 550 nm wavelength aerosol optical depth (Sayer et al., 2014). It has been shown that, in order to get a representative comparison between model data and satellite observations, model data should be sampled at the time and the location of the satellite observations they are compared to (Schutgens et al., 2016). For this purpose, we used the Community Intercomparison Suite (CIS) tool (Watson-Parris et al., 2016) which was applied to collocate the model AOD with the observations.

From the figure, we can see that the overall comparison between both models and satellite data is generally good. For the yearly mean values, the correlation coefficient R between MODIS AOD and SALSA AOD is 0.74 and for M7 it is 0.75. The normalized mean bias (NMB) for SALSA is -0.13 while for M7 it is -0.26. Areas that exhibit the largest differences between the models and observations are 1) the Saharas, 2) highly polluted areas over India and southeast Asia, 3) regions with high AOD due to biomass burning over Russia, Canada, Central Africa and South America. These are regions which are strongly
Figure 2. Yearly mean Aerosol Optical Depth for the year 2010 retrieved by a) MODIS (Aqua and Terra combined), and modelled by b) SALSA2.0 (model data collocated with Aqua and Terra retrievals), c) M7 (model data collocated with Aqua and Terra retrievals). Absolute differences between d) MODIS and SALSA2.0, e) MODIS and M7.
affected by primary emissions. However, over these areas also SALSA2.0 and M7 have noticeable differences in the simulated AODs which means that the aerosol representation has a significant effect on the modelled AOD. Over the Sahara, the most significant contribution to the AOD comes from mineral dust. Since dust emissions in ECHAM-HAMMOZ are very sensitive to small changes in 10-meter wind speed, changes in the wind speed can cause large changes in dust emissions even if the model meteorology is nudged (Bergman et al., 2012). This is because the nudging does not strictly force the model meteorology to reanalysis data. Consequently, difference in the model dynamics which result in changes in DU emissions explains the difference in AOD values between the two model configurations especially in the northwest regions of Africa where DU mass emissions in some of the grid boxes are more than 3 times higher in SALSA2.0 than in M7. This can be seen in Figure 3 which shows the relative change \( \Delta E_{DU} \) between SALSA and M7 mass emission strengths for DU. calculated from equation

\[
\Delta E_{DU} = \frac{E_{DU, SALSA} - E_{DU, M7}}{E_{DU, M7}}
\]

where \( E_{DU, SALSA} \) and \( E_{DU, M7} \) are the dust mass emissions strengths of SALSA and M7, respectively.

However, over southeast Asia and biomass burning regions, simulated aerosol load is mostly dictated by offline emissions which are, in mass, identical for both model setups. Thus, differences over these areas predominantly come from the differences in the representation of the size distribution, the microphysical processing of aerosols, and sink processes. This can be seen when comparing the simulated composition and extinction distributions at two sites where the simulated AOD is mainly driven by aerosol compounds from offline emissions but where the AOD in SALSA2.0 significantly differs from those in MODIS and M7. Figure 4 shows the 2010 yearly mean mass and extinction size distributions for SALSA2.0 and M7 over China at a location (30.775° N, 114.375° E) where the simulated AODs are extremely high (Fig. 4a) and Russia at a location (55.025° N, 39.375° E) where biomass burning emissions are high (Fig. 4b). To make the visual comparison easier, the M7 size distributions were remapped to SALSA2.0 size classes. At the Chinese site, AODs from SALSA, MODIS and M7 are 0.47, 0.87, and 1.13, respectively. At the Russian site, AODs from SALSA, MODIS and M7 are 0.70, 0.42, and 0.44, respectively.
When analyzing the aerosol mass size distributions, it is evident that over these locations the aerosol extinction is strongly affected by microphysical processing. In the upper panels of Figure 4 we show the mass size distribution for SALSA2.0 and M7. In each class, the mass fraction of each compound is indicated by a color. Figure 4 shows that the largest difference in the composition distribution comes from SU, which is the only condensable species in this model configuration. Compared to M7, SU in SALSA2.0 is more evenly spread among all sizes, and there is a relatively higher amount of sulfate in the largest sizes. This difference is very likely due to the numerical limitations of the modal scheme. The modal scheme has been shown to overestimate the condensational growth of the accumulation mode thus underestimating the amount of condensable species in the largest particles (Zhang et al., 1999). In addition, in the modal approach the mass distribution of all compounds follow the shape of the mode restricting the mass distribution of individual compounds. It has to also be noted that the emission size distributions are not optimal for M7 as the emissions in each mode are assumed to have a fixed radius. The same applies to SALSA2.0 since the emission size distribution assumed the same shape as M7.

The extinction at 550 nm wavelength for different sized particles at 70% relative humidity are shown in the lower panel of Figure 4. The aerosol extinction is a quantity which is highly nonlinearly dependent on the aerosol size, aerosol hygroscopicity, and relative humidity. Thus although the differences in SALSA2.0 and M7 simulated aerosol are caused by similar processes, differences in the simulated extinctions can switch signs at different sites. At the Chinese site the resulting shape of the size distribution of M7 yields a higher aerosol extinction than SALSA2.0 while at the Russian site, it is the opposite. At the Russian site the composition distributions of both OA and SU are significantly different between the two model versions. This is because OA is not included in the insoluble accumulation mode in M7 while in SALSA2.0, both soluble and insoluble size classes include OA. Wet removal is faster for soluble particles which results in faster removal of OA accumulation sized particles in M7. The overestimation of AOD with both model setups at the Russian site indicates that biomass burning emissions are overestimated.

It should be noted that over China MODIS has been shown to have a high bias in AOD when compared to AERONET observations (Lipponen et al., 2017). Especially over the highly polluted areas in China this high bias is likely to increase the discrepancy between the SALSA2.0 simulated AODs and MODIS AODs.

Figure 5 shows the zonal mean AOD for MODIS together with SALSA2.0 and M7 model data. To visualize how the wet deposition scheme affects the zonal AOD we also included the zonal mean AOD from the simulation with the aerosol size-dependent wet deposition scheme, which is used as the default scheme in ECHAM-HAMMOZ (denoted M7default in Fig. 4).

As can be seen from the figure, the modelled AOD decreases faster when moving from the equator towards the poles in comparison to the satellite observations. This is the case for both M7 and SALSA2.0 and has been apparent also in previous model versions (Stier et al., 2005; Bergman et al., 2012). Compared to the previous model versions, the decrease in AOD towards the South Pole has been further amplified due to the new Long et al. (2011) sea salt emission parameterization. This is because sea salt mass emissions decrease significantly in ECHAM-HAMMOZ when using Long et al. (2011) in comparison to the previously used Guelle et al. (2001) implementation.
Figure 4. SALSA2.0 and M7 simulated mass ($dn/d\log D_p$) and extinction size distribution in a) China (114.375° E, 30.775° N) and b) Russia (39.375° E, 55.025° N). The height of the bars in the upper row represents the number concentration of particles $dN/d\log D_p$. The colorbars represent the mass fraction of each chemical compound in each size class. In the bottom row the height of the bars denote the extinction of the size classes.

Figure 5. Zonal mean of aerosol optical depth in year 2010 observed by MODIS (Aqua and Terra combined), SALSA2.0 (red curve), M7 (blue curve), and M7 with default wet deposition scheme (green). AOD values from MODIS at high latitudes were excluded due to the larger retrieval uncertainty at high latitudes.

Overall, the zonal average of SALSA2.0 is in a better agreement with the observations than M7 except between the latitudes 10°S -35°S. Over these latitudes, the AOD is overestimated compared to MODIS. This is caused by biomass burning aerosol for which the emissions are likely overestimated. Similar to biomass burning regions in Russia (see Section 4.1.1), SALSA2.0
produces higher AOD than M7 over biomass burning influenced regions over Africa and South America affecting also AOD over the oceans in this latitude band 10°S -35°S.

Over the Northern Hemisphere, the magnitude of the zonal gradient of AOD in ECHAM-HAMMOZ is strongly dependent on the wet deposition scheme (Bourgeois and Bey, 2011). From Figure 5, it can be seen that compared to M7 the improved wet deposition scheme (M7default) increases the AOD towards the Arctic improving the comparison between the model and MODIS. The improved wet deposition scheme affects the AOD gradient to a similar degree as the choice of the aerosol microphysics scheme. For example, SALSA2.0 and M7default AOD values overlap in the sub-Artic and the Arctic region and, on average, the difference between M7 and M7default is smaller than the difference between M7 and SALSA.

The global mean AOD is also underestimated with both model setups although the bias in SALSA2.0 is smaller than in either of the M7 setups. Especially, the tropical maximum is better captured with SALSA. The observed global mean AOD from MODIS (Aqua and Terra combined) is 0.170 while the modelled values for MODIS collocated AOD are for SALSA2.0: 0.145, for M7 0.122, and for M7default 0.136. Figure 5 indicates that the low bias near in the high latitudes can partly be explained by low SS emissions, especially in the Southern Hemisphere. On the other hand, it has been previously shown that insufficient aerosol transfer in ECHAM-HAMMOZ can also partly explain low aerosol mass over the high latitudes (Bourgeois and Bey, 2011; Kristiansen et al., 2016). It is clear from Figure 5 that the low bias is the result low SS emissions. In addition, it has to be noted that except for South Africa and Oceania, MODIS overestimates AOD compared to AERONET observations (Lipponen et al., 2017).

4.1.2 Evaluation against AERONET observations

For comparing the model data with the AERONET sun photometer observations, we used the whole simulated (2003-2013) period for SALSA2.0 and the simulated year 2010 for M7. The level 2.0 daily AOD data from AERONET was collected for all available 984 stations. Simulated daily means were sampled for the days where AERONET observations are available and they were also spatially collocated to the location of the AERONET station. Afterwards, a yearly average of both observed and simulated daily means were computed.

Figure 6 shows the scatterplots of SALSA2.0 modelled AOD against AERONET observed AOD. Fig 6a illustrates that the model AOD correlates well with the observations for the years 2003 – 2012. This is also reflected in the statistical values of the comparison as the correlation coefficient R is 0.79 and the normalized mean bias is -0.09. SALSA2.0 also captures the temporal evolution of AOD very well for the majority of stations.

In the year 2010 comparison (see Fig 6c), the correlation coefficient decreases slightly to 0.73 and the NMB reduces to a value -0.03. M7 (see Fig 6d) shows also a very good correlation with the AERONET observations with correlation coefficient of 0.71 and bias of -0.05.

In Fig 6, different regions are separated by color. From this separation, we can see that although statistical values are comparable between M7 and SALSA2.0, similar to the comparison with MODIS, there are regional differences. Regional AOD values together with their correlation coefficient values are listed in Table 2. From these values, we can see that AOD in both model setups is biased low compared to AERONET AOD. For example, SALSA2.0 underestimates AOD in 7 out of
Figure 6. Scatterplots of yearly mean of daily AERONET AOD value against yearly mean of collocated simulated daily mean AOD. Panel a) represents the comparison between AERONET and SALSA2.0 for the period 2003-2012. Colors in the scatterplots denote different regions shown in the map in Panel b). Panel c) shows the comparison between AERONET and SALSA2.0 for the year 2010, d) comparison between AERONET and M7 for the year 2010. The given statistical values are the following: Root Mean Square RMS (normalised RMS), absolute bias (normalised bias), correlation coefficient R (R on log scale) and the ratio between simulated and observed standard deviation (sigma).
Table 2. Yearly means of daily AOD values from AERONET, SALSA, and M7 and the corresponding correlation coefficient values for the models.

<table>
<thead>
<tr>
<th>Region</th>
<th>AERONET</th>
<th>SALSA</th>
<th>M7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AOD</td>
<td>AOD</td>
<td>R</td>
</tr>
<tr>
<td>North America</td>
<td>0.143</td>
<td>0.135</td>
<td>0.97</td>
</tr>
<tr>
<td>South America</td>
<td>0.343</td>
<td>0.338</td>
<td>0.92</td>
</tr>
<tr>
<td>Europe</td>
<td>0.149</td>
<td>0.157</td>
<td>0.57</td>
</tr>
<tr>
<td>Northern Africa/Middle East</td>
<td>0.366</td>
<td>0.321</td>
<td>0.66</td>
</tr>
<tr>
<td>Central/Southern Africa</td>
<td>0.298</td>
<td>0.216</td>
<td>0.78</td>
</tr>
<tr>
<td>Asia</td>
<td>0.427</td>
<td>0.264</td>
<td>0.47</td>
</tr>
<tr>
<td>Siberia</td>
<td>0.113</td>
<td>0.067</td>
<td>0.86</td>
</tr>
<tr>
<td>Australia</td>
<td>0.052</td>
<td>0.076</td>
<td>1.00</td>
</tr>
<tr>
<td>Oceans</td>
<td>0.075</td>
<td>0.120</td>
<td>0.79</td>
</tr>
<tr>
<td>Elsewhere</td>
<td>0.139</td>
<td>0.127</td>
<td>0.73</td>
</tr>
</tbody>
</table>

10 regions. However, the correlation coefficient values are high for both models. The exceptions are Europe and Asia where the correlation coefficient values R are 0.57 or less for both SALSA2.0 and M7. In 3 out of 10 regions, SALSA has a higher correlation coefficient than M7 while the number of regions where each model has a lower bias is evenly divided. Asia is the only region where the AOD in M7 is higher than in SALSA. For example, M7 underestimates the AOD over the US (red markers), more than SALSA2.0. Over Asia, SALSA significantly underestimates the AOD both model versions differ from AERONET observations (shown by dark red markers in Figure 6) which was also the case in the evaluation against MODIS data. Again, SALSA2.0 simulated AOD over Asia is lower than in M7 and is generally underestimated compared to AERONET. As was shown in Section 4.1.1 microphysics can significantly affect the number and composition of aerosol over highly polluted regions causing differences in the modeled AOD between the sectional and modal setups. However, the differences between the simulated and AERONET AOD are not as evident as in the MODIS evaluation. One reason for this is that MODIS AOD is biased high over Asia, especially over highly populated regions of China (Lipponen et al., 2017).

4.2 Aerosol mass concentrations at the surface

Section 4.2 has been restructured so that different chemical compounds are discussed in subsections. Unless otherwise noted by blue color, the results and conclusions remain the same as in the originally submitted version of the manuscript.

To evaluate the simulated aerosol mass concentrations at the surface, we compared the model data with those measured by the European Monitoring and Evaluation programme (EMEP; http://www.emep.int) and the United States Interagency Monitoring of Protected Visual Environment (IMPROVE; http://vista.cira.colostate.edu/improve/). Both of these observation
networks provide data for the mass concentrations of individual chemical components of the aerosol and the data are freely available from both sources. From the European EMEP programme and the USA based IMPROVE monitoring sites we used the PM2.5 aerosol mass concentration data for sulfate and elemental carbon. Additionally, from IMPROVE we used the data for organic carbon. In total, data from 530 stations were used in the comparison. The comparison between SALSA2.0 and the surface observations was done for the period 2003-2012. From the model, we used the daily mean data sampled according to the days when there were observations at each station. To evaluate the difference between SALSA and M7, we also compared the simulated data for mass concentrations of SU, BC, and OA for the year 2010 against EMEP and IMPROVE observations.

In order to evaluate the simulated DU and SS mass concentrations, i.e. compounds whose emissions are wind driven, we compared the simulated masses against two sets of observations. Simulated dust masses were compared with the observations which were used in the Aerocom experiment by Huneeus et al. (2011) where 15 global models were compared to observations related to desert dust aerosols. Surface mass concentrations of DU were provided for the Pacific Ocean sites from the sea/air exchange SEAREX program (Prospero et al., 1989) and for the northern Atlantic sites from the Atmosphere-Ocean chemistry experiment AEROCE (Arimoto et al., 1995). The AEROCE observations include also data for SS surface mass concentrations which were used in evaluating the simulated SS mass concentrations.

4.2.1 Sulfate

Figure 7. Scatterplots of yearly mean aerosol mass concentrations observed at EMEP (left column) and IMPROVE (right column) stations versus those from SALSA2.0 simulations for SU. The given statistical values are the same as in Figure 6
Figure 7 shows the scatterplot of observed and modelled yearly mean PM2.5 concentrations of SU. The left column of the figure shows the data for EMEP stations and the right column shows the data for IMPROVE stations.

Similar to the comparison to the AERONET AOD, SU mass concentrations from SALSA2.0 simulations correlate well with the observed surface concentrations. The correlation coefficient for SU for EMEP sites is 0.72 and for IMPROVE sites it is 0.89. SALSA2.0 tends to overestimate SU for both EMEP (NMB of 0.25) and IMPROVE (NMB of 0.33) stations. The high bias of aerosol mass concentration of SU over both Europe and the US is in contrast to the underestimation of AOD by the model in these regions when compared to MODIS and AERONET AOD. This highlights the sensitivity of AOD to the shape of the aerosol size distribution. Aerosol water has also a significant contribution to AOD and simulated relative humidity and aerosol hygroscopicity can cause differences between the simulated and observed AOD. In addition, in these regions, nitrate is a significant source of aerosol mass (Bauer et al., 2007) and as it is missing in our model it may also be a cause for the differences between model and observations, although the representation of nitrate in coarse resolution models is not without complications (Weigum et al., 2016).

The evaluation was repeated for the year 2010 in order to include M7 in the comparison. For this simulation year, the correlation coefficient values for SALSA2.0 simulated SU mass concentrations were 0.60 for EMEP stations and 0.93 for IMPROVE stations. SALSA2.0 simulated SU mass concentrations in year 2010 had higher positive bias than those for the whole simulation period for both EMEP (NMB of 0.40) and IMPROVE (NMB of 0.42). For M7, the corresponding correlation coefficient values were 0.62 for EMEP and 0.93 for IMPROVE stations. Although, M7 had also high biases for EMEP stations (NMB of 0.23) and IMPROVE stations (NMB of 0.27), they were lower than for SALSA2.0.

4.2.2 Black Carbon

Figure 8 shows the scatterplots of observed and modelled yearly mean PM2.5 concentrations of BC for the whole simulation period. The left column of the figure shows the comparison for EMEP stations and the right column shows the comparison for IMPROVE stations.

Compared to sulfate, for BC, the correlation is slightly lower with the correlation coefficient being 0.62 for EMEP and 0.56 for IMPROVE sites. In contrast to SU, BC mass concentrations are underestimated while underestimating BC for both EMEP (NMB of -0.50) and IMPROVE (NMB of -0.20).

Similar to the evaluation of SU mass concentrations, the evaluation was repeated for the year 2010 including M7 in the comparison. For this simulation year, the correlation coefficient values for SALSA2.0 simulated BC mass concentrations were 0.42 for EMEP stations and 0.65 for IMPROVE stations. SALSA2.0 simulated mass concentrations in year 2010 were biased low, similarly to the whole simulation period, for both EMEP (NMB of -0.48) and IMPROVE (NMB of -0.21). For M7, the corresponding correlation coefficient values were 0.34 for EMEP and 0.64 for IMPROVE stations. M7 was also biased low for both EMEP stations (NMB of -0.53) and IMPROVE stations (NMB of -0.31).
Figure 8. Scatterplots of yearly mean aerosol mass concentrations observed at EMEP (left column) and IMPROVE (right column) stations versus those from SALSA2.0 simulations for BC for years 2003-2012. The given statistical values are the same as in Figure 6

### 4.2.3 Organic Aerosol

Surface mass concentrations of OA were compared to the IMPROVE observations. The data was available only for years until 2004 so here we compared the simulated year 2010 to observations for the year 2004 in order to get a better comparison between SALSA2.0 and M7. Figure 9 shows the scatterplots of observed OA surface mass against simulated values. Both models are biased low with NMB values of -0.56 and -0.59 and correlation coefficients of 0.40 and 0.42 for SALSA2.0 and M7, respectively. A more detailed evaluation of organic carbon will be carried out in a companion paper by Kuhn et al. (in preparation).

### 4.2.4 Mineral Dust

SEAREX was a 10-year (1977-1986) program and AEROCE a 5-year (1990-1995) program and thus outside of our simulation period, we compared the simulated data for the year 2010 to DU and SS climatologies. The monthly model values were constructed by averaging daily means only for days where an observation is available. Moreover, each model monthly mean was spatially collocated to the location of the observation station (by bi-linear interpolation).

Figure 10 shows the scatterplots of monthly mean observed DU and SS surface concentrations against those simulated using SALSA2.0 and M7. DU mass concentrations from both SALSA2.0 and M7 show a moderate agreement against observations but underestimate the low values. The correlation coefficients for SALSA2.0 and M7 are 0.66 and 0.47, respectively. Both
Figure 9. Scatterplots of yearly mean OA aerosol mass concentrations simulated and observed at IMPROVE sites (year 2004) for a) SALSA2.0 (simulated year 2010) and b) M7 (simulated year 2010). The given statistical values are the same as in Figure 6.

SALSA2.0 and M7 exhibit low NMB with values of -0.33 and -0.26, respectively. It has to be noted that, due to different periods in observations and simulations, DU and SS mass concentrations are not strictly comparable because they are very sensitive to the 10 meter wind speed.

4.2.5 Sea Salt

For evaluating mass concentrations SS we also used the data from SEAREX and AEROCE programs which were compared to the simulated SS mass concentrations for the year 2010. Figure 11 shows the scatterplots of observed monthly mean SS surface mass concentrations against simulated monthly mean surface mass concentrations from SALSA2.0 and M7. Collocation of the data was done identically to DU evaluation, described in the previous subsection. As was seen in the comparison between the models and MODIS retrievals, aerosol load over oceans south of latitude 40°S, especially in the Southern Hemisphere seems to be low in both model versions. This is also reflected in low SS mass concentrations in simulations when compared to the observations; in very few cases the values exceed the observed values. This indicates that the sea salt emissions are significantly underestimated in this model setup. The NMB for SALSA2.0 and M7 were -0.68 and -0.64 while the correlation coefficients were 0.19 and 0.18, respectively. This may also explain the discrepancies between the model and satellite AODs over the oceans as sea salt strongly affects the aerosol size distribution over the oceans.

Since DU and SS emissions are calculated online, they vary annually. In order to evaluate, how much the choice of the year affects these results, we repeated the analysis for DU and SS for each year using the 10 year SALSA2.0 simulation.
analysis showed that the main characteristics in the comparison between modeled and observed mass concentrations remain similar each year, i.e. the model has low bias in both DU and SS mass concentrations and the low model bias increases with decreasing mass concentration (for both DU and SS). For DU, the annual variability of the modeled mass concentration is fairly large with NMB ranging between -0.35 and -0.09. For SS the variability is low and the NMB varies between -0.74 and -0.70. The correlation between modeled and observed mass concentrations varies very little annually. For DU, the logarithmic scale correlation coefficient varies between 0.67 – 0.74 for DU and 0.58 – 0.67 for SS.

In addition, SS measurements are mostly at coastal sites where global models may have large biases in sea salt surface concentrations as SS emission parameterizations assume open ocean conditions (Spada et al., 2015). It has been suggested that caution should be taken when evaluating global models against coastal observations (Spada et al., 2015).

4.2.6 Summary

Table 3 summarizes the biases of simulated surface mass concentrations of SALSA2.0 and M7. In addition, as a reference, the table shows the same values from the previous model version ECHAM5-HAM-SALSA1 for the year 2008 which used the emissions for the year 2000 (Bergman et al., 2012). The table also shows the global burdens for these compounds for the same three model versions together with values reported by Liu et al. (2005) and Textor et al. (2006). Liu et al. (2005) has made a synthesis of model data and Textor et al. (2006) provides the analysis of global aerosol properties in Aerocom Phase I models for the year 2000.
From the table we can see that surface concentrations of sulfate and its global burden are significantly larger in SALSA2.0 than in the previous generation model and they are at the upper end of the estimate of (Liu et al., 2005). Although our simulation period is not for the same period as for ECHAM5-HAM, by Liu et al. (2005), and Textor et al. (2006), global sulfate emissions have been suggested to be fairly constant through 2000-2010 (Granier et al., 2011). Even larger increases between the two model generations are evident for the BC and OA burdens which are approximately 3 times higher in ECHAM-HAMMOZ-SALSA2.0. Despite these higher burdens the simulated BC and OA surface concentrations are biased low when compared to the observations from the IMPROVE network (see Figures 7, 8, and 9). The largest decrease in the burden can be seen for SS, which in SALSA2.0 has decreased to approximately 1/3 of the SS burden in ECHAM5-HAM supporting the conclusions of too low sea salt emissions in this model configuration. The DU burden has slightly increased between the two model generations with the DU burden being near the values of the Aerocom I mean.

4.3 Evaluation against aircraft observations

The previous evaluations showed how well the model reproduces surface concentrations and column quantities of aerosol. To get an indication how well the model reproduces the vertical properties of different aerosol compounds, we repeat the model evaluation of Koch et al. (2009) where Aerocom models were compared against observed BC concentrations from several aircraft measurement campaigns shown in Figure 12. Data from the following campaigns were used: ARCPAC (Brock
Table 3. Comparison of mean NMB in ECHAM5-HAM (with SALSA1), ECHAM-HAMMOZ (with SALSA2.0), and ECHAM HAMMOZ (with M7) for individual compounds at IMPROVE sites, the global burdens (Tg) of all compounds together with those reported by Liu et al. (2005) and the mean of Aerocom I models analyzed by Textor et al. (2006).

<table>
<thead>
<tr>
<th></th>
<th>ECHAM5-HAM-SALSA1</th>
<th>ECHAM-HAMMOZ-SALSA2.0</th>
<th>ECHAM-HAMMOZ-M7</th>
<th>(Liu et al., 2005)</th>
<th>Aerocom I</th>
</tr>
</thead>
<tbody>
<tr>
<td>SU</td>
<td>0.19</td>
<td>0.49</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>-0.24</td>
<td>-0.21 -0.47</td>
<td>-0.31 -0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OA</td>
<td>0.25</td>
<td>-0.37</td>
<td>-0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global burden (Tg)</td>
<td></td>
<td></td>
<td></td>
<td>(Liu et al., 2005)</td>
<td>Aerocom I</td>
</tr>
<tr>
<td>SU (Tg S)</td>
<td>0.64</td>
<td>0.96</td>
<td>0.74</td>
<td>0.53-1.07</td>
<td>0.66</td>
</tr>
<tr>
<td>BC</td>
<td>0.07</td>
<td>0.26</td>
<td>0.20</td>
<td>0.12-0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>OA</td>
<td>0.96</td>
<td>2.68</td>
<td>1.77</td>
<td>0.95-1.8</td>
<td>1.70</td>
</tr>
<tr>
<td>SS</td>
<td>11.73</td>
<td>3.53</td>
<td>4.21</td>
<td>3.41-12.0</td>
<td>7.52</td>
</tr>
<tr>
<td>DU</td>
<td>13.11</td>
<td>18.26</td>
<td>15.14</td>
<td>4.3-35.9</td>
<td>19.20</td>
</tr>
</tbody>
</table>

et al., 2011), ARCTAS (Jacob et al., 2010), ARCTAS-CARB (Jacob et al., 2010), TC4 (Toon et al., 2010), CR-AVE (https://espo.nasa.gov/ave-costarica2/), and AVE-Houston (https://espo.nasa.gov/ave-houston).

In addition, we evaluated the modeled mass concentrations of SU and OA measured in 17 different aircraft campaigns which have been compiled by Heald et al. (2011) shown in Figures 13 and 14. We also repeated the evaluation for the M7 and M7default setups. Data from the following campaigns were used: ACE-Asia (Huebert et al., 2003; Maria et al., 2003; Gilardoni et al., 2007), ADIENT (Morgan et al., 2010), ADRIEX (Highwood et al., 2007; Crosier et al., 2007), AMMA (Redelsperger et al., 2006; Capes et al., 2009), ARCTAS (Jacob et al., 2010; Cubison et al., 2011), DABEX (Haywood et al., 2008; Capes et al., 2008), DODO (Capes et al., 2008), EUCAARI (Kulmala et al., 2009; Morgan et al., 2010), IMPEX (Dunlea et al., 2009), ITCT-2K4 (Heald et al., 2006; Sullivan et al., 2006), ITOP (Fehsenfeld et al., 2006; Lewis et al., 2007), OP3 (Hewitt et al., 2010; Robinson et al., 2011), TexAQS (Parrish et al., 2009; Bahreini et al., 2009), TROMPEX (Heald et al., 2011), VOCALS-UK (Wood et al., 2011; Allen et al., 2011).

Figure 12 shows the vertical profiles of BC concentration (black curve) measured using the Single Particle Soot Photometer (SP2, Droplet Measurement Technologies, Inc., Boulder, CO) on board of aircrafts. In this comparison, we used only the model data for the year 2010.

The red curves represent the monthly mean BC concentrations sampled along the flight path from the SALSA2.0 simulations. The monthly means were calculated for the year 2010 for the month during which each aircraft campaign was performed. The BC aircraft campaigns can be divided between campaigns in the tropics and midlatitudes (AVE Houston, CR-AVE, TC4, and
CARB) and those performed at high latitudes (ARCTAS, ARCPAC). More details of these campaigns and their locations are given by Koch et al. (2009).

From Fig 12 we can see that near the source areas (tropics and midlatitudes) SALSA2.0 tends to overestimate BC concentrations quite significantly with the exception of the CARB campaign, where SALSA2.0 simulated BC concentrations are slightly lower than the observed mean and fall within the standard deviation of the data. Overestimation near the source areas can partly be attributed to the multiplication of biomass burning emissions by the factor of 3.4. In contrast, over high latitudes, SALSA2.0 simulated BC concentrations always fall below the observed mean. This is in line with many of the Aerocom models analyzed in the study by Koch et al. (2009).

Modelled SU and OA profiles showed significantly better comparison with the observations than BC. Especially the vertical profiles of SU in ACE-Asia, ADRIEX, TexAQS, EUCAARI, ARCTAS Summer, ITOP, and VOCALS-UK campaigns are captured very well by the model with the exception of TROMPEX, DODO and IMPEX campaigns. The SU profiles for the campaigns are shown in Figure 13 and OA profiles in Fig 14. The coloured lines represent the average of model daily means sampled along the flight tracks and the corresponding days of the flights. For BC, the difference between the observations and the model was more than one order of magnitude, whereas for SU and OA the difference is in most cases significantly smaller.
Figure 13. Observed and modelled mean vertical profiles of SU in aircraft measurement campaigns. Black curves represent the measured SU concentrations and the grey whiskers show the variability of measurements.

In many cases, modelled BC concentrations exceeded the limits of the variability of observations (grey whiskers in Figs 12, 13, and 14). However, modelled SU and OA concentrations fall within the variability of the observations in most campaigns. Note also that in Figure 12 concentrations are shown on a logarithmic scale, while in Figures 13 and 14 the scale is linear.

From these figures we can see that also for M7 the comparison between the model and the observations is clearly better for SU and OA than for BC. Similar to SALSA2.0, M7 tends to overestimate BC concentrations near the source regions while underestimating them at high latitudes. It is noteworthy that the simulated BC mass in SALSA2.0 and M7 generally agrees better near the surface and near the source regions than aloft and in the remote regions. At higher altitudes, above the 200 hPa pressure level, SALSA2.0 has always higher BC mass compared to M7. For the ARCPAC and Spring ARCTAS campaigns SALSA2.0 also simulates higher BC mass through the vertical column than M7. These differences indicate that in SALSA2.0,
Figure 14. Observed and modelled mean vertical profiles of organic aerosol in aircraft measurement campaigns. Black curves represent the measured OA concentrations and the grey whiskers show the variability of measurements.

Microphysical aging of BC is slower, which means that it takes a longer time for BC particles to obtain enough condensed material to be transferred to the soluble size classes in which they would be more efficiently removed.

Since SU and OA masses are less sensitive to microphysical processing than BC, similar systematic differences are not seen between SALSA2.0 and M7 simulated profiles of SU and OA. On the contrary, SALSA2.0 and M7 profiles are very similar for most of the campaigns and in most regions SALSA and M7 differ much less with each other than both with the observations. Although the microphysical processing of SU was shown to produce different mass size distributions of SU between SALSA and M7 in Figure 4, this does not translate to differences in mass as it is not very sensitive to aerosol microphysics.

The new wet deposition scheme improves noticeably the comparison between the model and the observations from the Arctic campaigns. Comparing M7 to M7default, the differences are larger for the BC profiles than for SU and OA profiles.
which are very similar for all three model setups. Especially for the ARCPAC and Spring ARCTAS campaigns the difference in BC concentration profiles between the two M7 setups becomes extremely large, with the difference being approximately two orders of magnitude near the ground level. This comparison is a clear indication that in order to simulate the vertical profiles of BC realistically, especially in the remote regions, an accurate description of both microphysics and wet deposition is required. This was also shown by Bourgeois and Bey (2011) who evaluated the effect of scavenging rates on the simulated Arctic BC concentrations and by Kipling et al. (2016) who compared the contribution of different aerosol processes on vertical profiles. Thus, with a more physically based size segregated wet deposition scheme, the SALSA2.0 simulated vertical profiles would likely improve with respect to observations.

It has to also be noted that the model data was for different years than the observations. To see how much this affects the results, we did an additional comparison, where we used the exact years from the SALSA2.0 simulation. In this comparison, we used model values for the year 2010. For black carbon, we used the modelled monthly values from the flight path for the month of the year which corresponded to the observations. For sulfate and organic aerosol, we used the modelled daily values from the flight path for the day of the year which corresponded to the observations. However, the difference between using the whole simulation period of 2003-2012 and using the actual days of flights as opposed to using only one year was fairly small. For most campaigns and height levels, the relative difference in black carbon mass concentration is less than 50% and the shape of the vertical profiles are very similar, mostly overlapping each other. The largest difference is for black carbon in the CR-AVE campaign where the relative difference in the mass concentration is ~83% in the lowest layer. However, the main characteristics of vertical profiles for all compounds were similar and would not change our conclusions.

5 Aerosol size distributions

Since the choice of the aerosol microphysical module will affect the particle properties that are most sensitive to microphysical processing, i.e. the number and the composition of fine particles, we evaluated the simulated size distributions. This was done by comparing the size distributions from SALSA2.0 and M7 simulations against those measured at the EUSAAR sites (Asmi et al., 2011). Figure 15 shows the median number and mass size distributions for four selected sites: Hyytiälä (boreal region), Mace Head (marine), Zeppelin (Arctic), and Kosetice (industrialized). The figure is separated into four panels, each of which includes four subplots. In each panel, the upper row shows the yearly median number size distribution together with the EUSAAR observed number size distribution shown (blue solid curve) and the lower row shows the mass size distributions (bar plots). In each panel, the left column is for SALSA2.0 and the right is for M7. In order to make the comparison clearer for the reader, we remapped the M7 modes to SALSA2.0 size classes. All size classes also show the relative mass contribution of individual model compounds using colored bars. The EUSAAR measurements were made using either Differential Mobility Particle Sizers (DMPS) or Scanning Mobility Particle Sizer (SMPS) for which the measured size range corresponds roughly to the size range of SALSA2.0 (~3 nm–10 µm). For the details about the measurements, see Asmi et al. (2011).

From Figure 15 we can see that both models reproduce the observed size distributions fairly well except for the Zeppelin station. The observed size distribution at Zeppelin exhibits a distinct mode with mean diameter of ~0.2 µm. This mode is
Figure 15. Observed and modelled yearly median number size distributions ($dn/d\log D_p$), and modelled mass ($dm/d\log D_p$) size distributions for four different EUSAAR stations: a) Hyytiälä, b) Mace Head, c) Zeppelin, and d) Kosetice. Observed values are represented by the solid blue curves and the observations are represented by the bar plots. The relative mass contribution of individual chemical compounds in each size class are denoted by a color.

not seen in either of the model setups. An overall difference between SALSA2.0 and M7 can be seen in the accumulation mode which peaks (both mass and number) at smaller sizes in SALSA2.0. Similarly to what was shown earlier in the MODIS comparison, this is likely because in the modal method, condensing species accumulate more the accumulation mode than in the sectional method (Zhang et al., 1999). In all four cases, the sulfate mass peaks in M7 in particles with diameters between 0.2 and 0.4 µm. In SALSA2.0 the peak of sulfate has more station to station variation, but it also peaks in particles with diameters
between 0.2 and 0.7 µm. In general, the differences between the two model approaches are largest for sizes smaller than 0.2 µm, i.e. the sizes that are more sensitive to microphysical processing.

6 Evaluation of the Pinatubo simulation

The simulations of the stratospheric aerosol formation and growth following the Mt Pinatubo eruption were compared against the High-resolution Infrared Radiation Sounder (HIRS) (Baran and Foot, 1994) and Raman lidar observations (Ansmann et al., 1997). Figure 16a shows the time evolution of the global burden of SU and SO₂ retrieved from HIRS, SALSA2.0, and the two M7 setups: one using the standard mode widths (M7), one using the mode widths recommended by Kokkola et al. (2009) (M7stratmod), and additionally one simulated using MAECHAM5-SALSA (denoted as SALSA1). The colored solid lines show the mean of the model ensemble and the shading the variability of different ensemble members. It has to be noted that the standard mode setup of M7 was optimized for describing tropospheric aerosol and was not intended to be used in the stratosphere. However, it is possible to modify the mode properties so that the model can simulate both tropospheric and stratospheric aerosol and has been done successfully in e.g. the GLOMAP-MODE global aerosol model (Dhomse et al., 2014).

From the figure we can see that here the difference between the sectional and modal setups becomes large. The simulated sulfur burden in the SALSA2.0 simulation is more than two times higher than in the M7 simulation at approximately 10 months from the start of the eruption. Provided that the estimate of the emissions strength of 8.5 Tg of sulfur for the Mt Pinatubo eruption is realistic SALSA2.0, SALSA1, and M7strat overestimate the sulfur burden in the early part of the simulation On the other hand, M7 in its standard mode setup significantly underestimates the sulfur burden for most of the duration of the simulation period. This is because the effective radius of the particles in the standard M7 grows to larger sizes than in SALSA and the growth enhances the removal of stratospheric sulfate particles by sedimentation. In the simulated volcano plume, growing sulfate particles form a mode of a very narrow width with a diameter of approx 1 µm (Kokkola et al., 2008) which is not well represented by the standard M7 coarse mode which has the geometric mean deviation of 2. The stratospheric aerosol configuration of M7 modified mode setup (M7mod) brings the M7 values close to SALSA2.0 values, however this mode setup is then only valid for the volcanic plume and very likely decreases the models ability to simulate the tropospheric aerosol. Especially the simulated tropospheric coarse mode particles are very different between the standard M7 and M7mod.

For example, the annual global burden of mineral dust is 6 Tg in the M7mod simulation, which is less than 40% of the values for the standard M7 simulation values shown in Table 3 being near the lower limit of the estimate by (Liu et al., 2005) and are likely to be significantly underestimated.

The evolution of the effective radii of the stratospheric aerosol after the eruption was also evaluated against Raman lidar retrievals from balloon-borne observations at Laramie, Wyoming (41° N, 141° W) and ground based observations at Geesthacht, Germany (53° N, 10° E). Figure 16b shows the mean effective radius in the 12 km to 20 km layer observed by the Raman lidars as well as those simulated using SALSA2.0, M7 in its standard mode configuration, and M7 with the modified mode setup.
Figure 16. a) The simulated global burden of SO$_2$ (dashed curves) and sulfate (solid curves), simulated by SALSA2.0 (red curves), M7 (blue curves), and M7stratmod (green curves), and SALSA1 (purple curves). The black curve shows the global sulfate burden retrieved from HIRS observations. The shading around the solid curves represents the variability of the model ensembles. b) The mean effective radius of stratospheric aerosol in the 12 km to 20 km layer observed by Raman lidar (diamond markers represent the observations at Laramie, circles represents those at Geesthacht, dashed lines show the 5 month running means), simulated by SALSA2.0 (red curves), M7 (blue curves), and M7stratmod (green curves) and SALSA1 (purple curves). The shading around the solid curves represents the variability of the model ensembles.

From Fig 16b we can see that SALSA2.0 reproduces the retrieved values of the effective radii, which do have a very large variability. The effective radii in the SALSA2.0 simulations follow the mean of the retrieved values as well as the time evolution of the retrieved effective radii. In M7, the coarse mode is effectively removed from the stratosphere and thus the effective radius is underestimated and the M7 values are at the low end of the retrieved values. M7mod setup shows a much better comparison following the average of the lidar measurements. The effective radius in M7mod reaches a larger maximum value and decreases faster than in SALSA2.0. Out of the two, SALSA2.0 corresponds slightly better to the observed time evolution of the effective radius. SALSA1 simulation results are fairly similar to both SALSA2.0 and M7strat results. Especially, the sulfur burden of SALSA1 follows closely the values in the M7strat simulation. The effective radius in SALSA1 peaks earlier and higher than in SALSA2.0 and M7strat simulations.

7 Conclusions

Section Conclusions has been restructured so that different chemical compounds are discussed in subsections. Unless otherwise noted by blue color, the results and conclusions remain the same as in the originally submitted version of the manuscript.
We coupled the sectional aerosol module SALSA2.0 to the aerosol-chemistry-climate model ECHAM-HAMMOZ. During the coupling, also HAM (the aerosol model of HAMMOZ) was modified to implement the sectional aerosol model SALSA2.0 as alternative to the default modal microphysics module M7. ECHAM-HAM coupled with SALSA2.0 was evaluated using a 10 year simulation period for the years 2003-2012, preceded by a one year spin up. Using 3-hourly output, SALSA2.0 required double the calculation time of M7 with Cray XC-30 architecture when 120 cores were used for the simulation.

Simulated aerosol optical depths were evaluated against those retrieved from satellite-based MODIS instruments and ground-based AERONET sun photometers. Aerosol mass concentrations of individual compounds were evaluated against EMEP and IMPROVE networks of ground-based particulate mass concentration observations and vertical profiles from several different aircraft campaigns.

The aerosol optical depths simulated with ECHAM-HAMMOZ-SALSA2.0 were biased slightly low compared to both MODIS and AERONET retrievals. Local differences were the highest over Southern Hemisphere oceans, deserts, southeast Asia, and regions affected strongly by biomass burning. Over the oceans and deserts, these differences are very likely caused by emissions of natural aerosols. In desert regions, dust emissions are very sensitive to the model meteorology as they are driven by the simulated 10 meter wind speed (Bergman et al., 2012). Currently, the dust source strenghts in ECHAM-HAMMOZ, are optimized for M7 and thus a better match between the observations and SALSA2.0 could be achieved by optimizing the source strengths for SALSA2.0. Over the Southern Hemisphere oceans, the newly introduced sea salt emission scheme (Long et al., 2011) is likely the main cause of the underestimation in AOD (in both SALSA2.0 and M7) as the simulated SS mass concentrations are much lower than with the emission parameterization of Guelle et al. (2001) which was used in the previous version of ECHAM5-HAM. The overestimation of AOD over biomass burning regions indicates that in this model configuration using the multiplier 3.4 for GFASv1 emissions produces excessive aerosol load near the sources. Over southeastern Asia, the reason for the low bias in SALSA2.0 simulated AOD against observations is likely due to the aerosol microphysical processing of the aerosol size distribution. This conclusion is backed up by the fact that M7 overestimates AOD over the same reason despite having the same emissions over that region as SALSA2.0.

When comparing the AODs simulated by SALSA2.0 and M7, the largest differences between the model versions occur in regions where SALSA2.0 also differs most significantly from the observation, i.e. deserts, southeast Asia and regions affected by biomass burning. The differences in southeast Asia and biomass burning regions are mainly caused by the different microphysics schemes as in these region, the size distribution is heavily modified by the condensation of sulfuric acid on aerosol. In addition, the choice of chemical compounds that are taken into account in different size classes and modes cause differences between SALSA2.0 and M7. In M7, the insoluble accumulation and coarse modes do not include organic compounds while in SALSA2.0, organic compounds are included in all size classes. This results in different the composition size distributions of organics in the two model configurations.

The evaluation of aerosol mass concentrations against surface measurements showed that simulated SU mass concentration on average exceeds the observed SU mass while OA and BC are slightly underestimated. This holds for both EMEP and IMPROVE stations. Simulated DU and SS were underestimated in the majority of stations. Especially, the simulated SS was significantly underestimated indicating that the Long et al. (2011) emission parameterization used in the SALSA2.0 config-
uration is biased low, a conclusion also supported by low AOD over the oceans. In the comparison between the SALSA2.0 simulated vertical profiles of SU and OA mass concentrations were in fairly good agreement with those measured in aircraft campaigns. However, the vertical profiles of BC mass concentrations in the SALSA2.0 simulations and aircraft measurements had large discrepancies especially in the Arctic with differences of more than one order of magnitude.

Comparison to M7 simulations showed that the vertical profiles of SU and OA are not very sensitive to the choice of the microphysics module. However, the simulated vertical profiles of BC mass concentrations show fairly large difference between SALSA2.0 and M7 especially at high altitudes and away from the source regions. This is likely to be caused by differences in the rate of microphysical aging of BC. However, in the current ECHAM-HAMMOZ version, SALSA2.0 and M7 simulated SU and OA vertical profiles seem to be more similar than in the previous model version (Kipling et al., 2016).

Simulated size distributions show fairly good comparison against those measured at EUSAAR sites. Especially compared to the previous generation model version ECHAM5-HAM-SALSA1 (Bergman et al., 2012), the agreement between measured and modelled size distributions has improved. ECHAM5-HAM-SALSA1 showed significant underestimation of particle numbers at most EUSAAR stations (Bergman et al., 2012), in the current model version such bias is no longer evident.

Overall, the microphysical scheme affects mainly particles in the lower end of the size spectrum, the simulated number size distributions and mass size distributions in SALSA2.0 and M7 differ especially for sizes smaller than 0.7 μm. The largest difference among different model compounds is in the accumulation size mass distribution of SU which is the only compound affected by condensation. One reason for this discrepancy is that modal models tend to overestimate the condensational growth of accumulation size particles (Zhang et al., 1999).

One simulated case where SALSA2.0 reproduces the observations considerably better than the default tropospheric setup of M7 is the simulation of the volcanic plume produced by the 1991 Mt Pinatubo eruption. In the volcano plume, microphysical processes affect strongly the aerosol size distribution leading steep gradients at particle sizes of approximately 1 μm in diameter. This is because the sectional size distribution allows for steep gradients in the size distribution. Such steep gradients are evident in volcanic plumes as the condensation of sulfuric acid “narrows” the size distribution by growing small particles faster than the largest ones.

Overall, SALSA2.0 performs slightly better than M7 in the evaluation cases where the statistical metrics were possible to calculate. Out of 9 comparisons against the observations of optical properties and mass, in 6 of them SALSA2.0 had smaller root mean square deviation, in 5 it had a smaller normalized mean bias and in 7 it had a higher value in the correlation coefficient R. On the other hand, it has to be noted that for many aerosol properties, e.g. vertical profiles of SU and OA mass concentration, SALSA2.0 and M7 show better agreement between each other than with the observations.

The result of this study indicate that SALSA2.0 is a competitive choice for modal aerosol microphysics modules in global atmospheric models. A sectional scheme, such as SALSA2.0, can capture a wide variety of possible atmospheric size distributions, including explosive volcanic eruptions and stratospheric geoengineering for which modal models often need tuning. Size-dependent anthropogenic aerosol emissions, which are starting to become available (e.g., Xausa et al., 2017), can also be easily incorporated into sectional modules further improving the ability of SALSA2.0 to realistically reproduce global aerosol size and composition.
**Code availability.** The ECHAM6-HAMMOZ model is made available to the scientific community under the HAMMOZ Software Licence Agreement, which defines the conditions under which the model can be used. The licence can be downloaded from https://redmine.hammoz.ethz.ch/attachments/download/291/License_ECHAM-HAMMOZ_June2012.pdf. The standalone zero dimensional version of SALSA2.0 is distributed under the Apache-2.0 licence and the code is available at https://github.com/UCLALES-SALSA/SALSA-standalone/releases/tag/2.0 with DOI 10.5281/zenodo.1251668

**Data availability.** The model data can be reproduced using the model revision r4098 from the repository https://redmine.hammoz.ethz.ch/projects/hammoz/repository/show/echam6-hammoz/branches/fmi/fmi_trunk. The settings for the simulation are given in the same repository, in Folder gmd-2018-47. MODIS data is available for download from Level 1 and Atmosphere Archive and Distribution System (LAADS) https://ladsweb.modaps.eosdis.nasa.gov/search/. AERONET data can be obtained using the Aerosol Robotic Network download tool https://aeronet.gsfc.nasa.gov/cgi-bin/webtool_opera_v2_new. EMEP data is available for download from the EBAS database at http://ebas.nilu.no/. IMPROVE data is available for download from the Federal Land Manager Environmental Database http://views.cira.colostate.edu/fed/DataWizard/Default.aspx. AEROCE and SEAREX data can be downloaded from http://aerocom.met.no/download/DUST_BENCHMARK_HUNEEUS2011/conc_aeroce.prn. The aircraft measurement data for BC can be downloaded from http://aerocom.met.no/download/BC_BENCHMARK_KOCH2009/. EUSAAR size distributions are available for download at https://www.atm.helsinki.fi/eusaar/. The data for the Mt. Pinatubo evaluation can be downloaded from https://redmine.hammoz.ethz.ch/projects/hammoz/repository/show/echam6-hammoz/branches/fmi/fmi_trunk/gmd-2018-47 . The aircraft data for SU on OC was received from several measurement teams who hold the ownership for the data and thus it is only provided by request from harri.kokkola@fmi.fi.

**Competing interests.** The authors declare that no competing interests are present.

**Acknowledgements.** The ECHAM-HAMMOZ model is developed by a consortium composed of ETH Zürich, Max Planck Institut für Meteorologie, Forschungszentrum Jülich, University of Oxford, the Finnish Meteorological Institute and the Leibniz Institute for Tropospheric Research, and managed by the Center for Climate Systems Modeling (C2SM) at ETH Zurich. This work was supported by the Nordic Center of Excellence, eSTICC (57001), the Academy of Finland Center of Excellence programme (grant no. 307331), the Academy of Finland research projects no. 308292, 283031, and 287440, the European Research Council (ERC) Consolidator Grant No 646857, and the European Union’s Horizon 2020 research and innovation programme under grant agreement No 641816 Coordinated Research in Earth Systems and Climate: Experiments, kNowledge, Dissemination and Outreach (CRESCEndO). Philip Stier would like to acknowledge funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) projects BACCHUS under grant agreement 603445 and the European Research Council project ACCLAIM under grant agreement FP7-280025 as well as the European Research Council project RECAP under the European Union’s Horizon 2020 research and innovation programme with grant agreement 724602. We thank Hauke Schmidt and Sebastian Rast (Max Planck Institut für Meteorologie) for their work on developing the ECHAM model and the interface between ECHAM and HAMMOZ, Maria Kanakidou (ECPL, University of Crete) for the help in compiling the EMEP and IMPROVE datasets. For the aircraft data we thank Hugh Coe (University of Manchester), Lynn Russell (Scripps Institution of Oceanography), Rodney Weber (Georgia Institute
of Technology), Jose Jimenez (University of Colorado at Boulder), Roya Bahreini (University of Colorado - CIRES, NOAA ESRL Chemical Sciences Division), Ann Middlebrook (NOAA ESRL Chemical Sciences Division), James S. McDonnell Foundation Award for 21st Century Science, NOAA grant NA17RJ1231, National Science Foundation grants ATM-0002035, ATM-0002698, and ATM04-01611, and the NERC Global Aerosol Synthesis and Science Project (GASSP) NE/J023515/1. IMPROVE is a collaborative association of state, tribal, and federal agencies, and international partners. US Environmental Protection Agency is the primary funding source, with contracting and research support from the National Park Service. The Air Quality Group at the University of California, Davis is the central analytical laboratory, with ion analysis provided by Research Triangle Institute, and carbon analysis provided by Desert Research Institute.
References


