Dear Editor,

We would like to thank you and the reviewers for their valuable input to our study. The reviewer remarks have helped us to considerably improve the manuscript. One of the main revisions was the assessment of the statistical significance of the results. Johannes Mülmenstädt of Universität Leipzig helped a lot with this new aspect, and subsequently we asked him to become a co-author of our paper. We found a problem in the NorESM SULF experiment in its treatment of the radiative effect of the sulfate in the terrestrial spectrum, which only marginally affects our conclusions. An explanation is added to the revised manuscript: " In the NorESM SULF experiment, an implementation inaccuracy leads to an overly large ra- diative effect in the terrestrial spectrum, by up to 0.5 to 1 $\mathrm{Wm}{-2}$ in the last decade of the geo- engineering. The consequence of a too high LW absorption by the aerosols in the stratosphere is moderately strong radiative warming in the stratosphere. This means a bit more SO2 was needed in order to achieve the desired effect in NorESM1-M SULF. "

Below we quote the Reviewer remarks in plain and our responses in blue italic. Please note that the sections, page and line numbers are based on the revised manuscript. In the revised manuscript changes are shown as bold italic and texts removed from the previous manuscript are stirked through.

Sincerely,

Aswathy

Reviewer 1 (S. Tilmes)

The paper by Aswathy et al. is investigating climate extremes in climate engineering simulations. The paper is well written and structured and is another important contri- bution to the discussion on how climate engineering would affect climate extremes with focus on temperature and precipitation. The impact of two different climate engineering methods are compared using three different Earth system models.

We would like to thank the Reviewer for her work in helping us improve the manuscript.

I have one major comment to the paper. One new finding of this study is that tempera- ture and precipitation extremes are more or less equally affected by climate engineering as the mean values. I am somewhat concerned about this conclusion, since this has implications for the calculation of social costs of geoengineering (as stated in the text) and I am not convinced that this statement is true. Some more analysis or at least dis- cussions would be helpful to support this conclusion. The study does only investigate annual changes and does not look into seasonal variations, in particular important for precipitation. Is this sufficient? Would an extension of the analysis to different seasons and somewhat refined regions (tropics, mid-latitudes, high latitudes) change the result?

We changed the conclusion of social costs of geoengineering. New conclusion (Section 4, Page 19, Line 623-631) is "Overall, we conclude that the climate-change driven increases in the upper extremes of temperature and precipitation are simulated to be rather well mitigated by the two SRM climate engineering methods. However, we also find that the potential to mitigate effects of climate change by means of SRM differs around the globe and seasonally. Not very well dampened are in particular the increase in the mean temperatures is in the Arctic, and especially the increase in the lower temperature percentile in the Arctic winter. At the same time, it is not easily possible to locally engineer the climate by SRM methods, as the analysis of the SALT scenario shows. These findings indicate additional conflicts of interest between regions of the world if it should come to discussions about an eventual implementation of SRM."

We have included seasonal analysis for both temperature and precipitation analysis and discussed in Section 3.5. A new table with refined regional values of (tropic, mid-latitudes and high latitudes) are shown in Table 2 and 3.

Another caveat of the findings is that the models used here do not simulate the effect of volcanic aerosols on dynamics and chemistry, which may change the results.

Thank you for pointing out the caveat, and a new text is included in the text in the Section 2, Page 5, Line 139-145. New text added now reads: "In the SULF simulation, the aerosol effects on radiation is included in the models via their optical properties (Niemeier et al., 2013). This is achieved by prescribing aerosol optical depth (AOD) and effective radius, which were calculated in previ-

ous simulations with an aerosol microphysical model ECHAM5-HAM (Niemeier et al., 2011); (Niemeier and Timmreck, 2015). This approach allows an impact of the aerosol heating on the dynamic of the ESM, while the feedback process of the dynamic on the areosols was only included in the previous simulations with ECHAM5-HAM ."

Specific Comments:

Introduction: Line 7: There are many studies following Crutzen's paper, I suggest to add "e.g.," before the citation list.

"e.g.," added before the citation list in Section 1, Page 2, Line 36.

Section 2: Line 25: It seems that all the climate models used in this study prescribe AOD and effective radius for the G3 experiment. So, these models do not inject SO2 directly. Please clarify if any of these models simulate the impact of aerosols on strato- spheric dynamics or chemistry, and if not, could this change the results?

We do not treat chemistry. However, dynamics may respond of course. As mentioned above a text is added in Section2, Page5, Line 139-145

Section 2.1 The climate extreme analysis may be misleading, since there is no separa- tion between seasons. Please comment. Also, from Table 2, P10 and P1 is not shown for precipitation. As shown in Tilmes et al., 2013, P10 and P25 indicate changes in light precipitation and a reduction is an indication for droughts, while the increase of higher intensities, like P90 or P99, indicates increasing heavy precipitation and there- fore flooding.

We now added, also in reponse to the major comment, a seasonal analysis for mean and extremes of temperature and precipitation in order to address this suggestion and discuss it in Section 3.5. We tried to also investigate P10 and P1. However, based on the results, we believe that the analysis of the consecutive dry days is a more reliable way to investigate the lower extremes of precipitation.

Page 32400, Line 5: Please clarify how maximum and minimum temperatures are defined, are these daily minimum/maximum temperatures or maximum and minimum temperatures of daily mean temperatures?

Daily minimum and maximum temperatures are used and the secentence is rephrased. Sentence now changed to "Data for daily maximum (TX) and daily minimum (TN) 5 temperature are directly provided from the model" in Section 2.1, Page 6, Line 189.

Section 3: Tables 2-4 only show multi-model mean values. Adding values from single models would be helpful to see how those vary. Also, a separation between Tropics, mid-latitudes (North and South) and global would be interesting, as well as the corre- sponding discussion in the text.

Values for the individual ensemble members are provided as supplementary *material. Area seperated ie., Tropics, mid-latitudes North and South and high latitudes North and South values are computed and shown as two sepreate tables Table 2 for temperature and Table 3 for precipitation and discussed in Section 3.2, Page 10, Line 320-331 and Page12, Line 384-392.*

Section 3.1: Page 32401, Line 15: Reference in brackets.

Changed the reference to brackets in Section 3.2, Page 9, Line 278.

Line 22: Instead or in addition to Figure 1, it would be helpful to show a PDF for example for the northern mid-latitudes over land, to easier identify the statement that there is "no shift in the tail of the temperature distribution". It seems to me that there is more warming over northern Europe and Canada in looking at P10. Again, differentiating in seasons may show a stronger signal than the annual average.

Thank you for this suggestion. Yes with the seasonal analysis it is clear that warming occurs over the northern hemisphere midlatitudes during SALT experiment. Hence we subsequently remove the statment that there is no shift in the tail of the temperature distribution and the new scentence (Section 3.2, Page 9, Line 292) now changed to "In SALT, the pattern for the upper per- centile temperature (T90) values are similar to those for the mean values in the northern hemisphere". As pointed out earlier, seasonal analysis of the percentiles are provided in Section 3.5

Line 32402, Line 4: Is this really only the case for northern high latitudes, or also mid-latitudes (30-60N)?

Thank you for the correction, to some extend it is also there in mid latitudes as well. Change is included in Section 3.2, Page 10, Line 308 as " For the Northern hemisphere high latitudes and continental regions in the Northern mid-latitudes as well as sea-ice regions in the Southern hemi- sphere mid-latitudes, a much stronger increase in the lower percentile of the temperature distribution (T10) is simulated.".

Line 13ff: This statement needs the addition, that only annual averages were considered and seasonal changes may be much larger.

Suggestion included and a new sentence added to the paragraph in Section 3.2, Page10, Line 317, "This aspect of the SRM is more detailed in the Section:season".

Line 18: Reference in brackets.

Changed the reference to brackets in Section 3.2, Page 11, Line 344.

Figures 6-8: these are too small to read. Maybe 4 rows and two column would work better?

In the revised manuscript we have combined the figures into one single figure (Figure 3) for easy comparison.

Reviewer 2

General Comments This study analyzes model simulations of two solar radiation management (SRM) schemes with regard to changes in mean and extreme temperature and precipitation and compares them to the RCP4.5 scenario simulations of current climate (i.e. year 2010). Results are based on 3 different global climate models and a comparison of 10year time slices to illustrate the effect of the two different SRM schemes compared to future climate change as represented by RCP4.5 scenario for year 2060 and the termination of SRM. The introduction and motivation for this paper are reasonable and well elaborated. However, the discussion of results as well as the conclusion is highly questionable regarding the methods used and illustrated in the paper. If more carefully assessed, some of the conclusions could potentially be very relevant (e.g., extremes are more effected by SRM termination than means, and the warming of the temperature's lower tail is not sufficiently offset by SRM), but the current paper is unable to robustly support these conclusions.

We thank the reviewer for her or his thorough review of our paper and hope that we were able to address the concerns satisfactorily.

Specific comments

The data basis for the analysis of extremes is insufficient to obtain robust results or support the strong conclusions drawn. 10 years are too short for this kind of analysis, even if simulations are in equilibrium and some models have more than one ensemble member. Information on the latter is not provided in the text, apart from a half sentence in section 3.1. Also I would argue that 3 models are not representative to conclude robust results just from at least two agreeing on the same sign of change(although this kind of measure is used in many GeoMIP papers).

We have changed the analysis to now assess longer time periods. 30 years of daily data (10950 days) are now analysed; except for the termination-effect analysis, in which case we investigated 20 years (7300 days). Results and conclusions are changed including more years in corresponding sections.

Information on the number of ensembles included in the analysis is now provided in Section 2, Page6, Line 161-166. We also substantially revised the analysis of the statistical significance and performed statistical tests to assess the robustness of the results and is discussed as a new Section 3.1.

The authors claim that T90 is representative for summer and T10 for winter season, which would reduce the effective sample size even further.

Since we didn't perform a seasonal analysis in the first version of the manuscript, this was just an additional statement to the annual mean analysis. In order to address this comment comprehensively, and following also the suggestion from reviewer 1, we have now performed the analysis also on a seasonal basis.

The signal-to-noise ratio in RCP4.5 is low, even in year 2060, to clearly distinguish the SRM effect from internal natural climate variability. The authors did not take this into account (e.g., could at least show variability between ensemble members of a single model). And again 10 years are fairly short to assess the contribution of natural variability.

Variability between ensemble members are shown in supplementary material. As mentioned also above, we have now redone the analysis to investigate 30 year-periods, and performed thorough statistical significance tests.

The discussion of results (and respective conclusions) are merely based on a qualitatively assessment of maps and numbers in various tables. The authors did not make effort to apply any statistical test to support the robustness or significance of their results. For instance, whether the change in the extremes really follows the mean cannot simply concluded by looking at global maps and numbers provided in the tables. (There are many papers on this issue!)

The reviewer is right, and the revision in light of this statement helped sharpen the discussion. Statistical tests are now applied to assess the robustness of the results. In addition to the geographical distribution of annual and seasonal variations and also following suggestion from reviewer1, we have also created tables summerising the results with refined regional values of (tropic, mid-latitudes and high latitudes) are shown in Table 2 and 3

There is no information on the baseline of the extremes indices indicated in the paper (i.e. values for RCP4.5 in 2010) to put the changes indicated in the tables into perspective to the overall magnitude of the respective index in the reference climate.

Values of the baseline of the extremes are now provided as supplementary material.

Results for consecutive dry days (CDD) are highly uncertain. Given the large disagreement between models, and the insufficient data basis, I don't think that such small changes as for instance 0.68 or 1.88 days/yr should even be discussed. See also my comment above. If the definition in Table 1 (i.e. of consecutive dry days per time period (=10years???)) is correct then one would expect very large numbers in some regions (e.g., Sahel zone) and then a difference of less than 1 or even 10 would be insignificant. Furthermore, the global mean of change in CDD is pretty much meaningless given the very heterogeneous distribution of positive and negative changes as illustrated in figure 3.

The unit for the consecutive dry days is days/year. This unit is now clarified where the CDD metric is introduced in Section 2.1, Page6, Line 194-195. For the revised work, as stated earlier, we have included more years and the robustness of the results are shown using statistical significance.

What about a model comparison with observation? Maybe the differences found between SRM schemes and RCP4.5 are just as big as the model bias compared to the observations.

We acknowledge that this study is a pure modelling study, and that ideally we should also investigate observations. However, a detailed model evaluation or observations-based study is beyond what we could possible do in this study.

Why are there no figures illustrating the termination effect if it is so important to draw conclusions from this analysis? The termination effect is not carefully assessed in this paper; see also my minor comments below.

Figures for the termination effect are now included in Figure 11 and 12 in the revised manuscript. We also respond to the minor comments below.

I highly criticize the conclusion drawn on the implications for the assessment of social costs of SRM (i.e.entire last paragraph of paper). This is pure speculation, not well-grounded in any literature reference,methodological approach or quantitative assessment. "There is no substantial indication for costly side effects" -How do you know? Given the substantial changes in spatial patterns of the extremes indices considered and associated uncertainties, I would not make such a strong statement without reconsidering the data basis and complementing the simple analysis of climate model data with some socio-economic data assessment and modeling approach.

We have now changed the conclusion in the revised manuscript. New conclusion is included in Section 4, Page 19, Line 623-631. Ne conclusion reads as "Overall, we conclude that the climate-change driven increases in the upper extremes of temperature and precipitation are simulated to be rather well mitigated by the two SRM climate engineering methods. However, we also find that the potential to mitigate effects of climate change by means of SRM differs around the globe and seasonally. Not very well dampened are in particular the increase in the mean temperatures is in the Arctic, and especially the increase in the lower temperature percentile in the Arctic winter. At the same time, it is not easily possible to locally engineer the climate by SRM methods, as the analysis of the SALT scenario shows. These findings indicate additional conflicts of interest between regions of the world if it should come to discussions about an eventual implementation of SRM.

More specific points:

Section 2.1, p. 32399, line 25: Aren't the extremes based on daily time series of minimum (2-m) and maximum temperature? It is mentioned further down (i.e. p.32400, line 5) but completely out of context as TX and TN are not further used in the paper.

Daily mean temperature and precipitation are used to discuss the percentile based extreme indices. However for Summer days and Frost days, TX and TN are used. Definition of the extremes are now defined more properly in Section 2.1, Page 6, Line 186-189.

Section 3.1, p.32401, line 12: "...models that simulated more than one..."??? Shouldn't that be "...models with more than one ensemble member..."?

Sentence removed since we have changed the analysis.

p.32402, line 6: “...SRM schemes are simulated to substantially narrow...”??? That sentence does not make any sense. Please rephrase.

Rephrased the sentence in Section 3.2, Page 10, Line 312, into ”Overall, both SRM schemes tend to substantially narrow the temperature distribution in the Arctic.”

p.32402, lines 13-16: This whole paragraph is unclear and does not make any sense as written. Please rephrase.

Removed the paragraph based on the new analysis.

Section 3.3.1, p.32406, lines 7-9: “The termination of the SRM ... compared to G3-SSCE method.” Makes no sense as G3-SSCE is a SRM method, isn’t it?

It is a typing error, changed in the sentence in Section 3.6, Page16, Line 517.

p. 32406, lines 12-14: “The models simulate drying ...”. The models don’t even agree on the sign of change in fig. 3, I doubt they do for the termination period. See also my major point on CDD above. “North of Africa”??? Is that Europe?

Sentence has been rephrased in Section 3.6, Page 16, Line 524. Fig 12 shows the change in precipitation for termination and the results are statistically significant.

Section 4, p. 32406, line 26: “...mean global warming caused by the RCP4.5 scenario...” Global warming is not caused by RCP4.5!

Sentence has be rephrased in Section 4, Page 16, Line 544.

p. 32407, line 3: “10-year temporal distribution” of what?

Sentence has be rephrased in Section 4, Page 16-17, Line 548-549.

p. 32407, lines 4-6: “In the simulations investigated, ...” This sentence makes no sense as written. Please clarify and rephrase.

Sentence has be rephrased in Section 4, Page 17, Line 552.

Referee 3

Firstly, we would like to thank the reviewer for her or his work in helping us improve our paper.

General Comment 0:
I am going to abbreviate Climate Engineering by Stratospheric Aerosols as "SACE" just so I can contrast it with SSCE.

Thank you for pointing out that the abbreviations used are bit confusing, so we changed the abbreviations for the experiments as SULF for Climate Engineering by Stratospheric Aerosols and SALT for Climate Engineering by Marrine Cloud Brightening, similar to Niemeier et al 2013. We also changed the denotation of 90th percentile of Precipitation as P90 and of temperature as T90/T10 for more clarity.

General comment 1: I think the way these simulations are usually analyzed can lead to misleading conclusions and I would like these authors to help to straighten out this problem. Alterskjær et al (2013) indicate that the "The (G3) SSCE experiment is de- signed so as to cancel the globally averaged radiative forcing relative to 2020 asso- ciated with the RCP4.5 scenario [Moss et al., 2010]". Berdahl et al say "G3 injects C12357 sulfate aerosols beginning in 2020 to balance the anthropogenic forcing and attempt to keep the net forcing constant (at 2020 levels) at the top of the atmosphere [Kravitz et al., 2011a]".

This study says (page 32396/L7-8) that G3 is designed "to balance the anthropogenic forcing and to keep the global temperature nearly constant (Kravitz et al., 2011).", and then says (page 32396/l10) that for G3-SSCE that "marine cloud brightness is altered, rather than stratospheric aerosols, are used to compensate the anthropogenic forc- ing." In all of these studies there is a comparison of the 2010-2019 decade with a later decade (2060-2069 or 2080-2089), with the implication that the scenarios were designed to maintain either temperature or precipitation at 2010 levels. Because of the climate forcing is not zero at 2010 or 2020, and there is a "climate commitment" (e.g. the planet will continue to warm even if forcing were to be fixed at those levels and forcing is not fixed) it is by no means clear that the CE would satisfy these as- sumptions. For the G3 scenarios, the appropriate forcing must be guessed ahead of time, and so one should anticipate that the planet will continue to warm after 2020 for both geoengineering scenarios, and as seen for example in figure 1 of Berdahl et al (2014), and in table 1 and figure 1 of Alterskjaer, the forcing is not fixed in time, does not balance the greenhouse gas forcing, and the climate continues to change over the integration period. Therefor the G3 behavior should be distinguished from the G1 sim- ulation, where there was a strict requirement that the TOA net flux be balanced for an extended period to 0.1 W/m2.

Because of this, one should not expect that there will be a reasonably complete com- pensation between the albedo modification and GHG forcing, and the study should reduce expectations about how strongly it should be achieved (e.g. on page 32394/l17 you say "are not completely alleviating the changes"). If the statistics do not stay con- stant it is as much an artifact of the experimental design as it is the choice of CE method.

Thank you for explaining the scenarios clearly. We have changed the manuscript text accordingly, reducing the high expectations associated with geoengineering. Please find the new text added in Section 3.2 Page 9, Line 293, text changed in Abstract, Page 1, Line 13-16, Section 3.4, Page 14, Line 444 and a text is removed in Section 4, Page 17, Line 560-562.

General Comment 2: I am getting tired of reading geoengineering papers that deliver the same messages over and over again. It is not that I don't believe the messages, I just don't think much detail is required when they are a repeat of previous studies. So I would encourage the authors to reiterate common conclusions with previous studies very briefly and very clearly, and then focus on aspects of these simulations where something new is to be learned from the simulations. I feel that the authors provide so much detail that one cannot see the forest for the trees. Here is my summary of the study:

Things that are the same as previous studies:

1. The planet cools with either geoengineering method (SSCE or SACE), and pre- cipitation decreases.
2. Warm/Moist precipitation events are mitigated by either CE method. Cold/Dry events are not.
3. The compensation is more effective in the tropics than in polar regions
4. There is a slight increase in width of the PDF of extremes events (as described in Curry et al) with geoengineering compared to the reference state, but, like Curry et al the resulting PDF is found to be much much closer to the "2010s world" than the RCP scenario with no geoengineering.
5. The signatures of weaker cooling over land with SSCE than SACE are the same as those found in the idealized studies by Bala and Caldeira (which I think should be cited)
6. The planet warms, and precipitation increases following termination. I feel like saying "duh", but if there is something new learned in this study I would be de- lighted to learn it.

I think these points should appear and be stated much more succinctly. Things that seem new or I have questions about:

Thank you for summarising previous studies. With your input, we revised our manuscript, and these points are now provided in Section4, Page 18, line 595-601. Also we have added concluding remarks to each Section, to make results easier to identify. Please note the new text in Section 3.2, Page 11, Line 335-342, Section 3.6, Page 16, Line 533-536 and in Section 4, Page 18, Line 589-593 602 -614

1. How different is the TOA forcing between G3-SSCE and G3? In the shortwave and in the longwave? Over land and over ocean?

TOA forcing between G3-SSCE and G3 is have discussed in Niemeier et al 2013. We have included a short discussion, and reference to their paper,in Section 3, Page 7, Line 221-230 and as a supplementary information, Figure of

TOA seperated for Land and Ocean for SALT and SULF is also provided.

2. Temperature statistics (differences with the 2010s world) seem relatively insensi tive to whether one uses SACE or SSCE.

Tempearture changes are relatively insensitive to the method indeed when treated globally, however regionally the influence is larger (refer Table 2 and 3).

3. I would like the authors to discuss how different the extreme events are from those predicted by the idealized studies (i.e. G1). Is G1 an adequate design strategy for understanding extreme events.

A direct one to one comparison with the G1 results is not easily possibly, since different indices are used to define extremes in Curry et al, 2014 paper. However, we have included a discussion in the introduction Section 1,Page 4, Line 101-109 as well as a discussion in the Summary and conclusion, Section 4, Page 18 Line 595.

4. It seems from table 2 that the precipitation over land is better treated by SACE than SSCE, but precip over ocean is better treated by SSCE.

Discussed in Abstract, Page 1 2, Line 19-21 26-30, Section 3.2, Page 11, Line 366-369 Page 12 , Line 376-381 and also in Section 4, Page 17, Line 566-571.

5. Are the differences between the two CE methods large enough that you feel they are robust and not artifacts of the scenario details?

To address the robustness of the results, we have included more years of data (30yrs) and statistical significance is also computed.

I quite liked the summary.
Thank you.

Page 32398/l20: “The G3 and experiment”. There is a word missing after “and”.

Change included.

Page 32417: “Realtive” should be “relative”.

Changed to minus.

Manuscript prepared for Atmos. Chem. Phys.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 30 May 2015

Climate extremes in multi-model simulations of stratospheric aerosol and marine cloud brightening climate engineering

Aswathy V N[1], Olivier Boucher[2], Martin Quaas[3], Ulrike Niemeier[4], Helene Muri[5], Johannes Mülmenstädt[1], and Johannes Quaas[1]
[1]Institute for Meteorology, Universität Leipzig, Germany
[2]Laboratoire de Météorologie Dynamique / IPSL / CNRS, Université Pierre et Marie Curie, Paris, France
[3]Department of Economics, Christian-Albrechts-Universität zu Kiel, Germany
[4]Atmosphäre im Erdsystem, Max-Planck-Institut für Meteorologie, Hamburg, Germany
[5]Department of Geosciences, University of Oslo, Norway
Correspondence to: J. Quaas (johannes.quaas@uni-leipzig.de)

Abstract. Simulations from a multi-model ensemble for the RCP4.5 climate change scenario for the 21st Century, and for two solar radiation management (SRM) schemes (stratospheric sulfate injection (G3), SULF), and marine cloud brightening ***by sea salt emission*** ~~(G3-SSCE),~~ SALT) have been analyzed in terms of changes in the mean and extremes for surface air temperature and precipitation. The climate engineered and termination periods are investigated. During the climate engineering period, both schemes, as intended, offset temperature increases by about 60% globally, but are more effective in the low latitudes and exhibit some residual warming in the Arctic (especially in the case of SALT that is only applied in the low latitudes). In both climate engineering scenarios, extreme temperature changes are similar to the mean temperature changes over much of the globe. ~~The exception is in northern Hemisphere mid and high latitudes~~ ***Exceptions are mid and high latitudes in the northern Hemisphere***, where high temperatures (90th percentile of the distribution) of climate engineering compared to RCP4.5 control period rise less than the mean, and cold temperatures (10th percentile), much more than the mean. ***This aspect of the SRMs is also reflected in simulated reduction of the frequency of occurrence of frost days for either scheme. However, the frequency of occurrence of summer days, is increasing less in the SALT experiment than the SULF experiment, especially over the tropics.***

~~When defining temperature extremes by fixed thresholds, namely number of frost days and summer days, it is found that both climate engineering experiments are not completely alleviating the changes relative to RCP 4.5.~~ ***Precipitation extremes in the two SRM scenarios act differently - the SULF experiment more effectively mitigates extreme precipitation increases over land compared to the***

SALT experiment. A reduction in dry spell occurrence over land is observed in SALT experiment. ~~is mostly similar to that for RCP4.5.~~ The SULF experiment has a slight increase in the length of dry spells. A strong termination effect is found for the two climate engineering schemes, with large temperature increases especially in the Arctic. ~~Mean temperatures rise faster than the extremes, especially over oceans, with the exception of the Tropics. Conversely precipitation extremes rise much more than the mean, even more so over the ocean, and especially in the Tropics.~~ . ***Globally, SULF is more effective in reducing extreme temperature increases over land than SALT. Extreme precipitation increases over land is also more reduced by SULF than SALT experiment. However, globally SALT decreases the frequency of dry spell length and reduces the occurrence of hot days compared to SULF.***

1 Introduction

Observed and projected global warming due to continuously increasing greenhouse gas emissions has promoted research focusing on the mitigation of greenhouse gas emissions as well as adaptation to climate change, and lately on alternative methods to counterbalance global warming. Climate engineering or geoengineering has been proposed as a means to counteract global warming in the case mitigation efforts prove insufficient or climate change becomes catastrophic (***e.g.,*** Crutzen, 2006; Schmidt et al., 2012). There are many proposed methods of climate engineering, which can be classified into two major categories, namely Solar radiation management (SRM) and Carbon dioxide removal (CDR). Solar radiation management aims to reduce solar radiation absorbed by the Earth system by increasing its albedo.

Several SRM techniques are being discussed, among them stratospheric sulfate aerosol injection has been suggested to be most feasible and least expensive (Lenton and Vaughan, 2009; Robock et al., 2009). SRM by marine cloud brightening is another technique, first proposed by Latham (1990). A number of single model studies have addressed both SRM techniques (Latham, 2002; Robock et al., 2008; Jones et al., 2009, 2010; Niemeier et al., 2013). Different experiment designs, however, hinder direct model-to-model comparisons (Kravitz et al., 2011). To answer the questions raised in independent studies, a suite of standardized climate modelling experiments has been performed within a coordinated framework, known as the Geoengineering Model Intercomparison Project (GeoMIP, Kravitz et al., 2013). GeoMIP consists of four solar climate engineering experiments namely G1, G2, G3 and G4, in which the G3 and G4 experiments investigate the effects of stratospheric sulfate aerosol injections. The GeoMIP G3 experiment is analysed in our study. Similarly, a first multi-model approach with common experimental setup to study sea salt climate engineering (SSCE), i.e. marine cloud brightening, has been performed within the “Implications and risks of engineering solar radiation to limit climate change” (IMPLICC) project (Alterskjaer et al., 2013).

The objective of this paper is to examine multi-model simulation results in terms of changes in mean and extreme temperature and precipitation as a consequence of reducing incoming solar radiation at the surface by these two different SRM techniques.

Kharin et al. (2007) found that the changes in temperature extremes can be expected to generally follow changes in mean temperatures in many parts of the world. However, especially over mid and high latitudes, temperature extremes may show larger relative changes, and over land, models show an increase in temperature variability in a warming climate (Kharin and Zwiers, 2005). According to the recent assessment report of the Intergovernmental Panel on Climate Change (IPCC), there will be more hot and fewer cold temperature extremes as well as a likely increase in precipitation extremes in a warmer world (Collins et al., 2013).

In this study, we compare the impact of stratospheric sulfate injection and sea salt climate engineering on changes in means and extremes of climate parameters. For statospheric sulphate injection, we use the GeoMIP G3 experiment, in which stratospheric aerosols are added gradually to a background ***following the representative concentration pathway 4.5 scenario*** (RCP4.5), to balance the anthropogenic forcing and to keep the global ***mean surface*** temperature nearly constant (Kravitz et al., 2011). The IMPLICC G3-SSCE is based on the GeoMIP G3 experiment, but sea salt emissions by which marine cloud brightness is altered, rather than stratospheric aerosols, are used to compensate the anthropogenic forcing. ~~Since this experiment is based on the GEOMIP G3 experiment, as described in Alterskjaer et al. (2013), we denote the experiment G3-SSCE.~~ ***Following Niemeier et al. (2013), we denote the G3 experiment (stratospheric sulfur injection) as SULF and G3-SSCE (marine cloud brightening by sea salt emission) as SALT.***

The SULF experiment exerts its forcing globally, whilst the SALT scheme is employed only over tropical oceans between 30°S and 30°N.

The climatic properties of the SULF and SALT experiments have been presented in previous studies. These focused mainly on the temporal and spatial distributions of climate engineering effects on the mean climate (Schmidt et al., 2012; Alterskjaer et al., 2013; Kravitz et al., 2013; Muri et al., 2015). Schmidt et al. (2012) studied the responses of four Earth system models to climate engineering in the G1 scenario. In this scenario, the radiative forcing from quadrupling of CO_2 is balanced by reducing the solar constant. Alterskjaer et al. (2013) investigated the simulation of SALT. Their results showed that a sufficiently strong application of SALT led to the compensation of the global annual mean warming by RCP4.5 in all models. The models showed a suppression of evaporation and reduced precipitation over low-latitude oceans and vice-versa over low-latitude land regions. Kravitz et al. (2013) summarized the current knowledge as gained from the GeoMIP simulations and remaining research gaps. They found that none of the participating models could maintain both global-mean temperature and precipitation to pre-industrial levels from a high CO_2 scenario, in agreement with theoretical considerations.

Presently, very few studies address the impact of climate engineering on extreme events and hardly any research has yet focused on the more realistic scenarios. Recent studies by Tilmes et al. (2013) and Curry et al. (2014) examined climate extremes in the multi-model climate engineering experiment (G1). The study by Tilmes et al. (2013) mainly focuses on the hydrological impact of the forcing as applied in the G1 experiment. As part of their study, they also analyze the upper percentile shifts in the annual and seasonal precipitation from monthly averaged model output in both G1 and abrupt $4\times CO_2$ experiments relative to the pre-industrial control state. In the Tropics, the G1 experiment tends to reduce heavy precipitation intensity compared to the control simulation. Their results showed a weakening of hydrological cycle under the G1 experiment.

Curry et al. (2014) investigated the temperature and precipitation extremes in the G1 scenario. They were found to be smaller than in the abrupt 4xCO_2 scenario, but significantly different from pre-industrial conditions. A probability density function analysis of standardised monthly surface temperature exhibited an extension of the high-end tail over land and of the low tail over ocean, while the precipitation distribution was shown to shift to drier conditions. The strong heating of northern high latitudes as simulated under $4\times CO_2$ is largely offset by the G1 scenario. However significant warming was found to remain, especially for daily minimum temperature compared to daily maximum temperature for the given time period. Changes in temperature extremes were found to be more effectively reduced compared to precipitation extremes.

The climate extreme indices used in this study are defined in Table 1 (see Methods described in Section 2). Details of the experiments considered in the study, models used and methods are described in Section 2. In Section 3, we discuss the geographical distribution of the climate extremes under the two climate engineering scenarios. Annual and seasonal variations of the extremes and the effect of termination on the extremes are discussed in the corresponding subsections of Section 3 ~~as well as the main differences in climate extremes under the two methods during climate engineering and after its termination.~~ In Section 4, we discuss the implication of our results and present the conclusions.

2 Data and Methodology

Results from three Earth system models (ESM) were available for the analysis. The models are the Max-Planck-Institute's ESM (MPI-ESM) (Giorgetta et al., 2013), the Norwegian Climate Centre ESM (NorESM) (Bentsen et al., 2013) and the Institute Pierre-Simon Laplace 5th-generation Coupled Model (IPSL-CM5) (Dufresne et al., 2013). The atmospheric component of the MPI-ESM lower resolution (MPI-ESM-LR), ECHAM6, runs at a resolution of T63 (triangular truncation at wave number 63, corresponding to approximately $1.9^\circ\times1.9^\circ$) with 47 vertical levels. The Norwegian Earth System Model1 - medium resolution (NorESM1-M) atmospheric model CAM4-OSLO has a resolution of $1.9^\circ\times2.5^\circ$ with 26 vertical levels, whilst LMDz, the atmosphere in the IPSL

Earth System Model for the 5^{th} IPCC report -low resolution (ISPL-CM5A-LR), runs at a resolution of 1.9°×3.75° with 39 vertical levels. The advantage of using models of such different components and resolutions is that the results from the different models are expected to span a large part of the uncertainty range of the results (Kravitz et al., 2013).

~~The G3 and experiment are compared to the RCP4.5 experiment for the period 2020 to 2070.~~ The aim of the climate engineering experiments is to balance the excess radiative forcing to remain at 2020 levels implied by the anthropogenic climate change in the Representative Concentration Pathway 4.5 (RCP4.5) post year 2020. [1] ***The experiments SALT and SULF follow the experiment design as given in Kravitz et al. (2011). For SALT only NorESM included sea salt emissions. The other two models prescribed the aerosols as calculated from NorESM (Alterskjaer et al., 2013).*** ~~In the SULF simulation, this is achieved by prescribing aerosol optical depth (AOD) and effective radius as given in Niemeier et al. (2013) in the stratosphere from an equatorial injection of SO_2 (Kravitz et al., 2011)~~. ***In the SULF simulation, the aerosol effects on radiation is included in the models via their optical properties (Niemeier et al., 2013). This is achieved by prescribing aerosol optical depth (AOD) and effective radius, which were calculated in previous simulations with an aerosol microphysical model ECHAM5-HAM (Niemeier et al., 2011); (Niemeier and Timmreck, 2015). This approach allows an impact of the aerosol heating on the dynamic of the ESM, while the feedback process of the dynamic on the areosols was only included in the previous simulations with ECHAM5-HAM . For both experiments, these are*** ~~This is~~ done increasingly in time, ***i.e., for 50 years from 2020 to 2070*** in order to reflect enough solar radiation to balance the increasing anthropogenic greenhouse effect. An additional 20 year extension ***of the simulation*** until 2090 is performed to explore the effect of abrupt ceasing of the SRM, which is referred to as the "termination effect" (Jones et al., 2013).

In the NorESM SULF experiment, an implementation inaccuracy leads to an overly large radiative effect in the terrestrial spectrum, by up to 0.5 to 1 Wm^{-2} in the last decade of the geoengineering. The consequence of a too high LW absorption by the aerosols in the stratosphere is moderately strong radiative warming in the stratosphere. This means a bit more SO_2 was needed in order to achieve the desired effect in NorESM1-M SULF.

In the ~~G3-SSCE~~**SALT** experiment, the globally averaged radiative forcing in RCP4.5 relative to the year 2020 is balanced via marine cloud brightening (MCB) by increasing injections of sea salt into the tropical marine atmospheric boundary layer (Alterskjaer et al., 2013). The seeding region chosen for the experiment extends between 30°N and 30°S over oceans. ~~Suitable~~ Seeding regions were chosen based on an earlier study by Alterskjær et al. (2012). For a detailed description of the SALT results and experiment design the reader is referred to Alterskjaer et al. (2013); Muri et al. (2015).

---

[1] RCP4.5 is a scenario that stabilizes radiative forcing at 4.5 W m^{-2} in the year 2100 (Taylor et al., 2012).

The MPI-ESM performed three realizations for both the SULF and SALT experiments. The NorESM1-M performed two realizations for both experiment, while IPSL-CM5A has one realisation for each experiment. Based on the time period chosen for analysis, ***firstly*** we compute the model statistics for each ensemble member for the models where more than one are available, and then consider the multi-model average. The multi-model mean results are given with an equal weight for all three models (i.e. first taking the ensemble-average for the models where more than one ensemble member was available). Prior to all calculations, all the three models ensembles are re-gridded to a common resolution, choosing the lowest of the model resolutions of $1.9^\circ \times 3.75^\circ$ (IPSL-CM5A-LR resolution).

2.1 Climate extreme analysis

In this study, climate extremes are defined by the lower and upper percentiles of the temporal distribution at each grid-point, as well as a set of indices defined by the Expert Team of Climate Change Detection and Indices (ETCCDI, Sillmann et al., 2013).

The daily-average model output is analysed for 30 year periods, except when analysing the termination effect, in which case a 20 year period is assessed. For the annual mean analysis, the data from which the extremes are drawn covers 10950 days and for termination it is 7300 days at each model grid point.

~~so the sample from which the percentiles are drawn covers approximately 3650~~.

Climate extremes ***are defined by the 90th and 10th percentile of*** the time-series of near surface air temperature (T90 and T10 respectively) and ***90th percentile of*** surface precipitation flux (P90) at individual model grid-points. ~~The 10th (P10) and 90th (P90) percentiles are analysed.~~ We also investigate higher percentiles (eg 99th), but this only as a ***global***-land- or ocean average ~~where statistics are large enough~~ (as shown in Table 2). ~~Since we do not de-seasonalize the data, the 10th percentile may be regarded as a measure of the winter- and 90th percentile that of the summer season.~~

The additional climate extreme indices used in this study are the frequencies of occurrence of Summer days (SU), Frost days(FD) and the maximum count of Consecutive Dry Days (CDD) in the period. These are computed from daily maximum temperature, daily minimum temperature and precipitation, respectively. Data for daily maximum (TX) and daily minimum (TN) temperature are directly provided from the models. The frost days index (FD) represents the number of days when TN<0°C and summer days (SU) define the number of days when TX>25°C for the given time period (usually three decade in our analysis). The consecutive dry days index (CDD) provides the largest number of consecutive days when daily precipitation is less than 1 mm day^{-1} in the analysed time period. ***In the Table 4 and Figures 4, 5 and 6 the units for CDD, FD and SU are converted to days/year.***

~~The climate extreme values are calculated separately for the RCP4.5 and the two SRM experiments.~~

To assess the influence of climate engineering on a changing climate, for every climate extreme index analysis, the last ~~ten years~~ ***three decades*** of climate engineering ***(2040 to 2069)*** are compared with ~~ten-year~~ the ***three-decades*** average at the beginning of the RCP4.5 scenario simulation ***(2006 to 2035, denoted as control period, CTL)*** ~~, i.e. the pre-SRM period).~~ The same analysis is conducted for the corresponding RCP4.5 scenario for the same ~~two ten-year~~ ***time*** periods. ***In addition to the annual mean changes, we also investigate extreme events for different seasons namely, December-January-February (DJF) and June-July-August (JJA), presented in Section 3.5.***

To determine the effect of abrupt ceasing of climate engineering on extremes, the upper ***and lower percentiles*** of ~~climate indices of~~ both temperature and precipitation for the ~~second~~ ***two*** decades after termination, i.e. years ~~2080 to 2089 (referred as 2080s)~~ ***2070 to 2089 (referred to as 2070s)*** are compared to the last ~~ten years~~ ***two decades*** of climate engineering ~~i.e.,2060s~~ ***(i.e., 2050 to 2069, represented as 2050s)***. A similar analysis is carried out for RCP4.5 as well, to investigate the changes during the same ~~two decades~~ ***time periods***.

Both climate engineering techniques are compared with the RCP4.5 (2040 to 2069) period, and the values are given in Table 6.

3 Results and Discussion

For reference, Tables 2, 3 and 5 show the changes in globally averaged values of mean and extreme (percentile based method) values of temperature and precipitation and Table 4 shows the globally averaged mean values of the other extreme event indices (Section 3.3 and 3.4). ***As a supplementary information, ensemble separated values for each model and for all scenarios are also provided, with the ensemble members showing relatively small variations between them.***

~~The tropical region in the table is defined as the area between 30°N and 30°S.~~

The main aim of the climate engineering experiment is to keep the globally averaged top-of-atmosphere radiative forcing at the RCP4.5 2020 level, hence it does not fully constrain the regional climate characteristics (Curry et al., 2014). ***Niemeier et al. (2013) computed the top of the atmosphere (TOA) flux changes in short wave (SW) and long wave (LW) for the last decade of climate engineering minus the RCP4.5 (2015-2024) for the MPI-ESM. They found that the top of the atmosphere short wave change*** ~~***in net radiative flux***~~ ***for the SALT in the MPI-ESM was smaller than the one for SULF over both ocean and land (Figure included in Supplementary material). However for SALT experiment, TOA SW fluxes are slightly larger over ocean relative to land. The difference of the solar radiation flux between land and ocean in SALT reflects the more local nature of this SRM, since SALT is applied only over tropical oceans. Long wave (LW) fluxes of both the SRMs are mostly similar, although SULF experiment is slightly larger than SALT experiment, except for all sky conditions over land.***

3.1 Statistical significance

To determine the roboutness of the results, we compute statistical significance test for the mean and extreme changes. Statistical significance of the change in mean temperature is computed using a two-sided Student t-test. For the mean change in precipitation we use Kolmogorov–Smirnov test, since the test is non parametric and make no assumptions about the probability distributions of the variable used (Conover, 1980).

The distribution of T90, T10, P90, SU, FD, and CDD is not sampled by the climate models (each ensemble member only provides a single value). To estimate the distribution function of these variables, we use sampling with replacement ("bootstrapping", e.g. Efron and Tibshirani, 1998). In the case of T90 and T10, the distribution of daily-mean temperature is sampled. In the case of P90, the distribution of daily accumulated rainfall is sampled. In the case of CDD, contiguous days with below-threshold precipitation (< 1 mm day^{-1}) are indexed, and the set of indices is sampled; this procedure preserves the temporal autocorrelation of the precipitation distribution. In the case of summer (winter) days, a binomial distribution with probability n/N is sampled, where n is the number of summer (winter) days and N is the total number of days in the model run. In all cases, 1000 samples of size N are used. The distribution is calculated independently at each grid point.

Once the bootstrapped probability distribution function for each model run i has been determined, the perturbed distribution $f_i(x)$ is compared to the reference distribution $g_i(x)$. The aim is to test the null hypothesis that $f_i(x)$ and $g_i(x)$ have been drawn from the same distribution. We calculate the overlap of the two distributions, denoted as

$$P(f_i > g_i) = \int_{-\infty}^{\infty} dx\, f_i(x) \int_{x}^{\infty} dx'\, g_i(x'). \tag{1}$$

The two-sided p-value for the null hypothesis is then

$$p_i = \min\left\{P(f_i > g_i), 1 - P(f_i > g_i)\right\}. \tag{2}$$

The p-value is calculated independently at each grid point.

To estimate the combined statistical significance in the multi-model ensemble, the p-values for each ensemble member are combined according to Fisher's method (Fisher, 1925). This method assumes that the same hypothesis test is carried out on k independent data sets (in our case, the different model runs), and yields the test statistic

$$X = -2\sum_{i=1}^{k} \ln(p_i) \tag{3}$$

with p_i calculated according to (2). Under the null hypothesis, this test statistic follows a χ^2 distribution with $2k$ degrees of freedom. The multi-model combined p-value is calculated from the χ^2

distribution function with $2k$ ***degrees of freedom*** $p_{\chi^2}(x;2k)$ ***as follows:***

$$p = \int\limits_{X}^{\infty} p_{\chi^2}(x;2k)\,dx \tag{4}$$

Geographical patterns of the changes in climate that remain despite climate engineering are examined in the following section and ***the regions where the changes are statistically significant at 95% are represented by hatches.***

3.2 Percentile based climate extreme analysis

Geographical distributions of ***change in*** mean, 90th percentile (T90) and 10th percentile (T10) of near surface temperature ***2040 to 2069 with respect to the reference RCP4.5 control period (CTL, 2006 to 2035)*** are shown in Figure 1 for RCP4.5 (left column), ~~G3-SSCE~~ ***SALT*** (middle column) and ~~G3~~ ***SULF*** (right column). ~~In these maps and those which follow, regions where the three models (ensemble average for the models that simulated more than one ensemble member) agree on the sign of the change are represented by hatches.~~

~~In the RCP4.5 scenario~~ ***For the mean and extremes simulated for the RCP4.5 scenario***, temperatures are ***warmer*** almost everywhere ~~warmer~~ in the ~~2060s than in the 2010s~~ ***2040 - 2069 period than in the control*** ~~***period***~~ (Fig. 1), with more warming over land than over ocean ***(Collins et al., 2013)***. ~~In both SRM scenarios, the three models agree on the positive sign of the temperature change (warming) over the extratropics and higher latitudes of the northern hemisphere, and also over much of the Southern hemisphere.~~ ***In both SRM scenarios, for the mean change in temperature, a residual, statistically significant warming is simulated over most regions globally for mean, upper and lower extremes of the temperature distribution. The warming compared to CTL in mean temperatures is larger than 0.5 K over the high latitudes (60°N-90°N) of the northern hemisphere. In the SALT experiment, the strong residual warming is extended over the continents to the mid-latitudes. Geographical distributions of the upper percentile (T90) of the two SRM techniques exhibit different warming patterns. The SALT experiment, being implemented in the marine tropical oceans, exhibits more uniform warming of 0.5-1 K over northern hemisphere mid- to high latitudes (30°N-80°N), emphasising more on the local influence of this experiment. Over most of the tropical oceans, change in temperature in the SALT experiment is close to or even less than zero with respect to CTL.***

In SALT, the pattern for the upper percentile temperature (T90) values are similar to those for the mean values in the northern hemisphere. ***The SULF experiment rather well mitigates the warming of the upper precentile, down to 0.5 K or less in most areas. This residual warming is still significant. For both the SRM methods, for the upper percentile, there is no warming north of 85° N. In contrast, most of the warming at the Arctic region occurs at the lower tail of the temperature distribution.***

~~In G3-SSCE, the low latitudes show a very slight warming. The warming is also not completely offset in the extratropics, especially in the northern hemisphere, which are simulated to warm by about 1 K. In G3-SSCE, the pattern for the P90 values are similar to those for the mean values, with not much offset of temperature (P90), indicating no shift in the tail of the temperature distribution. In some regions, such as the Arctic, the P90 values change somewhat less than the mean values. In G3, the result is very similar to the G3-SSCE experiment: the P90 changes largely follow the mean changes, except for a better mitigation of the warming in the Arctic for the upper percentile compared to the mean.~~

At the lower end of the temperature distribution, the 10th percentile increases in both SRM experiments broadly show ***in the tropics*** a distribution of small, positive changes very similar to the mean temperature change patterns. ~~, except for the~~ ***For the northern hemisphere high latitudes and continental regions in the northern mid-latitudes as well as sea-ice regions in the Southern hemisphere mid-latitudes,*** ~~***In these, rather than a smaller increase as for the upper percentile,***~~ ***A much stronger increase in the lower percentile of the temperature distribution (T10) is simulated.*** ~~This is consistently simulated by all three models~~. Overall, both SRM schemes ~~are simulated~~ ***tend*** to substantially narrow the temperature distribution in the Arctic. This is very likely due to the fact that both climate engineering schemes are solar radiation management approaches, by which only during Arctic day climate change can be mitigated (as seen in the upper percentile), while ~~in~~ ***during*** polar night, almost no local mitigation is achieved by construction. Warming in the lower tail of the temperature distribution may have important effects in the Arctic. ***This aspect of the SRM is more detailed in the Section 3.5.***

Table 2 lists global and regional mean, model-ensemble-mean values of changes in temperature of 2040 to 2069 minus the reference RCP4.5 control period (2006 to 2035). Difference values for global (all points, land only and ocean only), Tropics (30°N-30°S), mid-latitudes (30°-60° in both hemispheres) and high-latitudes (60°-90° in both hemispheres) are provided. For the SALT experiment, the models simulate a comparatively effective mitigation for the Tropics and mid-latitudes, and generally over oceans, with warmings of 0.17 to 0.26 K in the mean and an even more effective mitigation of the upper extremes. However, over northern hemisphere mid and high latitudes, the SALT experiment leaves a residual warming of 0.57 to 1.01 K, up to double the value simulated by the SULF experiment over the the same regions. As discussed earlier for the distributions, irrespective of the SRM technique simulated, warming at the lower tail of the temperature distribution (given by the lower percentile (T10)) at northern hemisphere high latitudes are much higher than the upper percentiles.

~~The conclusion that over much of the globe, the extremes change as the mean for the geoengineering schemes, except for the Arctic region, is clearly visible in the analysis of the geographical patterns, even if in the global mean (Table 2) slight differences are calculated.~~

In terms of both the mean and the extremes, the models simulate that the SALT experiment mitigates the warming better in the tropics and most of the Southern hemisphere, while it simulates a stronger residual warming, compared to the SULF experiment, in the northern hemisphere mid-latitudes, which may further affect the temperature gradient and circulation from tropics to mid-latitudes (Niemeier et al., 2013). Regarding the lower percentile (T10) warming, irrespective of the techniques, both the SRM tend not to mitigate warming in the Arctic well, and neither in some parts of the Southern ocean region. To get more insight into the warming patterns retained during SRM we also investigate the seasonal changes in Section 3.5

Changes in mean and the upper percentile (P90) precipitation are shown in Figure 2. As documented in earlier studies ***(e.g., Govindasamy and Caldeira, 2000)***, the RCP4.5 scenario shows an ***overall*** increase in precipitation ~~between the 2060s and 2010s~~ ***in the 2040-2069 period compared to the 2006-2035 period***, especially in the equatorial region between 5°N and 5°S. ~~The location of this increase, however, is not robust among the three models.~~ Changes in upper percentile (P90) precipitation in the RCP4.5 scenario are stronger than changes in mean precipitation.

~~Both SRM schemes show small area averaged changes, with decreases in mean precipitation in the subtropics and Tropics and slight increases elsewhere (Figure. 2and Table 2).~~ ***Mean changes in precipitation for the SRM are shown in Figure 2b) and c) and the changes in upper percentile (P90) in Figure 2e) and f).*** The ~~G3-SSCE~~***SALT experiment*** differs from ~~G3~~ the ***SULF*** experiment in the aspect that the precipitation is influenced by the emission of sea salt impacting cloud droplet number concentrations and subsequently precipitation formation in the clouds via the autoconversion process.

For both the mean and extreme precipitation, the ~~G3-SSCE~~***SALT*** experiment shows a rather strong positive anomaly over South-East Asia, ***as well as central Africa***. The Indian subcontinent and surrounding ***land*** regions are found to experience enhanced precipitation rates under the ~~G3-SSCE~~ ***SALT*** experiment. However, in the Amazon rainforest area, ~~G3-SSCE~~ the ***SALT*** experiment produces a negative anomaly in precipitation, in accordance with the simulation of Jones et al. (2009) on marine cloud brightening. In contrast to land regions, most of the tropical marine regions, including the ITCZ, Pacific, Atlantic and Indian Oceans show a negative anomaly for ~~G3-SSCE~~ ***the SALT experiment***. As discussed by Alterskjaer et al. (2013); Niemeier et al. (2013), in addition to the influence on autoconversion, these changes can be attributed to ~~a reduced vertical motion over the tropical Pacific under in the climate engineering experiment~~ ***large scale dynamics of increasing vertical motion in ITCZ and Walker circulations. This leads to an increase in the convective precipitation over land, compensating for the decrease in precipitation over the oceans. Thus over oceans, the SALT experiment is effective in reducing the extreme precipitation increases compared to the CTL period, which are stronger than the RCP4.5 2040s change relative to CTL.***

~~The mean value for the change in precipitation extremes (90th percentile) over global continents is positive at +0.08 mm day^{-1}, while over oceans it is slightly negative at –0.01 mm day^{-1}, yielding a global-mean change of +0.02 mm day^{-1} (refer Table 2)~~

The geographical distributions of the changes in precipitation ***of*** mean and upper percentile ***(P90)*** for the stratospheric climate engineering, ***SULF are shown in the right column of Fig. 2.*** ~~(G3; right column of Figure 2) are in general similar to the findings for the G3-SSCE simulations, albeit slightly smaller in magnitude.~~ ***In contrast to the SALT experiment, the SULF experiment effectively alleviates the precipitation extreme increases over land in the Tropics as well as northern hemisphere mid-latitudes compared to the CTL period, even shows decrease in extreme precipitation in these areas for P90 precipitation and a highly mitigated value for P99. When averaging globally, these features are prominent with SULF experiment resulting in more positive anomaly in precipitation over ocean and vice versa over most of land regions. Hence the changes in precipitation are almost opposite to SALT experiment, as pointed out in Niemeier et al. (2013) and the paper attributes the changes to the change in Walker circulation.***

Mean changes of precipitation for the 2040 to 2069 period with respect to the CLT period are given in Table 3. On global average, mean precipitation and 90th percentile are simulated to be well mitigated by both schemes, while the 99th percentile is still mitigated in its increase. Over land, the residual increase in the upper percentile (P99) precipitation simulated for the SULF scenario is 0.172 mm/day. For the SALT experiment, 0.359 mm/day increases are simulated, which is 50% less than the RCP4.5 scenario. Over ocean, the SULF experiment shows the same changes as RCP4.5, however less in magnitude. In the SALT experiment, the mean and 90th percentile precipitation is simulated to even decrease, while the 99th percentile is well mitigated in its increase.

~~and over oceans in particular, however, the G3 experiment produces more positive precipitation changes both for mean and extremes than the G3-SSCE experiment (Tables 2 and 3).~~

In Figure 3 the precipitation changes as simulated by the individual models are shown. In the SULF scenario, the tendency of all models to simulate moister equatorial tropics (ITCZ) and dryer sub-tropics is even more evident than for the ensemble mean. ~~In P90, both MPI-ESM and NorESM1-M are consistent in simulating a wetter Eurasian continent and wetter South Pacific Convergence Zone.~~ The signals are similar between mean and upper percentile, but stronger for the upper percentile. In the SALT, all models widely agree on reduced extreme precipitation over tropical marine regions and moister continents and ~~these factors are~~***this feature is more*** prominent in SALT compared to SULF experiment.

3.3 Changes in dry spells

Dry spells are measured as the number of consecutive dry days (CDD, Table 1). These are defined as the largest number of***consecutive*** days in the analysed period in which precipitation is less than

1 mm day^{-1}. In Figure 4, changes in ~~ten-year-average~~ CDD***, in units of days per year,*** for RCP4.5, SALT and SULF are shown for the 2040-2069 in comparison to the RCP4.5 2006-2035 control period. ~~CDD changes show little agreement among the three models, this is because it is derived from the more uncertain daily precipitation.~~

~~As expected, the desert regions show the extreme values. RCP4.5 has fewer CDD in the 2040s with a global mean reduction of 1.8 days year^{-1} over land, while over the tropical regions there is an increase in CDD of 1.44 days year^{-1}.~~

In the ~~G3-SSCE~~ ***SALT*** experiment, shorter dry periods are simulated, especially over the land regions. ***This could be because in SALT the precipitation has been shifted onto land.*** Australia, South Africa and most of Asia show a decrease by approximately 2-5 days year^{-1}. Over the Arabian peninsula, the decrease in CDD is up to 10 days year^{-1}. There are few regions where CDD increases in the ~~G3-SSCE~~ ***SALT*** experiment, mostly over parts of North Africa including Libya and Algeria. Overall the effect of SALT is most pronounced over global continents with a reduction of 0.29 days year^{-1}. Hence in ***global average values also*** ~~G3-SSCE~~, the overall increase in mean and extreme precipitation (discussed earlier) over continent and decrease over oceans is reflected in the CDD values as well.

Similar to the result for the SALT experiment, in general CDD for SULF also seems to decrease where there is increase in precipitation intensity and vice versa. Global mean values of CDD for land only and ocean only points also support this, with more CDD over land and less over ocean with values 0.41 days year^{-1} and 0.05 days year^{-1} respectively.

~~Similar to G3-SSCE, the G3 experiment results in increases in CDD, over the Tropics as well as global oceans and, reductions in CDD over large parts of extratropical continents, but more strong in magnitude than the G3-SSCE.~~

3.4 Changes in frequency of occurrence of cold days and hot days

The frequency of occurrence of cold days is quantified here as the number of frost days, defined as days per year when the minimum temperature (TN)***is*** less than 0°C. In RCP4.5, FD is reduced in the mid- to high latitudes especially of the northern hemisphere by up to one month per year, and widespread by 5 and more days per year over all extra-tropical continental areas of the northern hemisphere (Figure 5), ***with a global mean value of -3.03 days year^{-1} (Table 4)***.

Globally there ~~is less decrease in the~~***are fewer*** frost days under both SRM scenarios compared to CTL period with mean changes of -1.70 days year^{-1} and -1.34 days year^{-1} for SALT and SULF respectively (Table 4).

RCP4.5 scenario shows very few regions of increase in frost days. In comparison to RCP4.5, the SRM scenarios maintain more frost days over NH land. However, a strong reduction in the frequency of occurrence of FD is simulated for both SULF and SALT, with patterns very similar to the simulated increase in the RCP4.5 scenario. It may be concluded that the warming especially at the lower end of the temperature distributions, ***which is*** not offset by the SRM scenarios ***(Section 3.2)*** is sufficiently strong. Hence, it reduces the frequency at which the freezing threshold is reached and subsequently FD are reduced. ~~is largely not alleviated.~~ For all regions, the SULF experiment is simulated to be more effective in mitigating the decrease in frost days, possibly because the forcing is applied globally, and is more effective towards higher latitudes than SALT.

The frequency of occurrence of hot days can be quantified as the number of Summer days (SU), defined as the the total number of days per year in which TX is greater than 25°C. Figure 6 shows the yearly change in SU for the 2040 - 2069 period vs the CTL period. As expected, RCP4.5 shows an increase in SU. This is most pronounced in the sub-tropics with increases by up to more than one month per year, but is widespread over low- to mid-latitude continents (LIU Yunyun, LI Weijing, ZUO Jinqing and Zeng-Zhen, 2014). In the Tropics the maximum increase of 86 days year^{-1} corresponds to an entire season more of SU, and the average increase is as much as 11 days year^{-1} (Table 4). This strong increase over the Tropics is well reduced by the ~~G3-SSCE~~ ***SALT*** scenario, however, the still substantial increase of 10-20 days year^{-1} over North America and Eurasia is only slightly offset. In contrast, the extra-tropical changes in SU are effectively reduced by the globally-applied ~~G3~~ ***SULF*** scheme, where, in turn, still substantial increases in SU over the tropics (up to 30 days year^{-1}) are simulated. ***Looking at the global mean values and also ocean and tropics seperately, it is clear that the increases in the occurrence of summer days are more effectively reduced in the SALT experiment, which is not unsurprising considering this is the region of the forcing.***

3.5 Seasonal changes in Extremes

Temperature and precipitation extreme events depend a lot on the seasonal variations. Hence studying the annual changes is not enough to explain the extreme event analysis. So we also analyse the change in extreme events based on two different seasons; namely DJF and JJA. This analysis is done for the percentile based method i.e, upper percentile (90th percentile) and lower percentile (10th percentile).

Zonal mean change in mean temperature, upper percentile (T90) and lower percentile (T10) for annual, DJF and JJA season is shown in Figure 7. During DJF season, there is a noticeable warming over the northern hemisphere high latitudes existing for the upper percentile (T90) for both SRM methods. This signal was completely absent in the annual change analysis Section 3.2. The SRM techniques are ineffective during the winter season over the high latitudes. Therefore, even with SRM implementation warming in the northern hemisphere polar regions still persists. This result shows one of the major caveats of the SRM techniques. Change in upper percentile (T90) for JJA is similar to the annual change in temperature. Lower percentile (T10) analysis for DJF seasonal temperature also exhibits profound warming over northern hemisphere, higher in magnitude and spatial extent than the upper percentile (T90) warming. Warming pattern in

lower percentile is mostly similar to the annual change analysis. Warming in the lower tail of the temperature distribution has implications to permafrost and ice melting and sea level rise. These are some of the major issues of anthropogenic climate change that can inherently not be addressed by SRM techniques.

However, for JJA season lower percentile (T10) temperature there is much less warming over the northern hemisphere high latitudes, indicating the effectiveness of SRM during summer season. Even though there is less warming in the Arctic, there is still residual warming of 0.5 to 1K over the northern hemisphere mid latitudes in SALT experiment. Since JJA corresponds to winter in southern hemisphere, there is a net warming in the lower percentile (T10) in the southern hemisphere.

In conclusion, irrespective of both the SRM techniques, there is net warming at the lower tail of the temperature distribution at high latitudes during winter season. Extend of warming is more in the SALT experiment compared to the SULF experiment. Annual changes in the upper percentile (T90) is essentially that of the JJA season and lower percentile (T10) is that of DJF season.

Precipitation changes are highly dependent on seasons and Figure 8 shows the zonal mean change in precipitation for annual, DJF and JJA season. Since precipitation pattern is different over land and ocean, zonal mean curves for land only (top row) and ocean only points (bottom row) are shown separately in Figure 8. For JJA season, which corresponds to the monsoon season over northern hemisphere, SALT leads to increase in extreme precipitation compared to the CTL scenario. DJF seasonal precipitation mostly behave similar to the annual mean. In general for both the seasons, similar to annual mean precipitation over land is better treated in SULF experiment and ocean in SALT experiment.

3.6 Termination effect

The termination effect of the SULF and SALT experiments are investigated for both temperature and precipitation and shown in Figure 9 and 4. ***We only consider the annual changes in this section and*** the values are summarized in Table 5.

As expected, the termination of SRM leads to a rapid net global warming. When following the ***mean temperature of*** RCP4.5 scenario ~~(2080s vs 2060s)~~ ***in the 2070 - 2089 vs the 2050 - 2069 period,*** a gradual warming is simulated which is stronger for the average temperatures in the northern polar and mid-latitude regions than the global average of +0.30 K. T90 temperatures rise at a slower rate than the average ones ~~except in the Tropics where the warming rates are slightly larger for the upper percentiles~~.

The termination of the SRM lead to strong warming of ~~both~~ average and extreme temperatures for both schemes, with slightly larger values for the ~~G3~~ ***SULF*** simulations. ***For both the methods,*** changes are stronger over land. ***For both the SRMs,*** mean values rise the most in the northern polar regions, while T90 values increase more at mid- and low latitudes over land, with only moderate warming in the polar regions. The global mean values of the temperature changes for the ~~G3-SSCE~~ ***SALT*** scenario for mean, T90 and T99 are +0.59 K, +0.59 K and +0.65 K, respectively. In the ~~G3~~ ***SULF*** scenario, simulated patterns are similar to ~~G3-SSCE~~ ***SALT***, but stronger. The termination of the ~~SRM~~ ***SULF*** leads to stronger changes in extreme temperatures also in the mid- and polar regions, compared to the ~~G3-SSCE~~ ***SALT*** method. The global mean change for temperature extremes over land for SULF is +0.84 K. ***In lower percentiles (T10) due to termination, temperature rises much faster than the mean and upper percentile (T90) in both the SRM schemes. Particularly strong warming is simulated over the northern high latitudes as well as some regions of the southern ocean.***

~~A~~ Similar analysis is carried out for precipitation as well. Termination of ~~G3-SSCE~~ ***SALT*** leads to strong increases of precipitation over most regions. However, the models simulate ***reduced precipitation*** over some subtropical land regions, namely northern Africa, Europe and some regions of Indian subcontinent due to the termination effect. The global mean change of precipitation extremes over land is +0.461 mm day^{-1} (P99), half the magnitude over ocean. Tropics experience a large increase in precipitation extremes (P99) with a net value of +1.001 mm day^{-1}. Under ~~G3~~ ***SULF*** termination, ~~similar results are found, albeit with less drying anomalies in the subtropical areas.~~ ***there is large increase in precipitation over most of the land, mainly, the south east Asia, south of Africa as well as the Amazon region. Overall the precipitation over land regions are increased by +0.561 mm day***$^{-1}$

In conclusion, the termination effect of SULF on temperature is stronger than for the SALT experiment. In the SALT experiment, the termination results in larger precipitation increases over ocean than land. Hence, in general the termination of the SRMs results in a reversal of the patterns simulated to occur during the climate engineering time.

4 Summary and conclusions

In this study, the results of simulations with three different Earth system models within the SRM climate engineering model intercomparison studies of IMPLICC and GeoMIP have been analyzed with respect to surface air temperature and precipitation and their corresponding extreme indices. Two solar radiation management methods were implemented in these simulations, namely the injection of stratospheric aerosols (SULF) and marine cloud brightening by sea salt injections (SALT). Both solar radiation management climate engineering methods are effective at counteracting the mean global warming ~~caused by the RCP4.5 scenarios~~. In the marine cloud brightening experiment, SALT, however, where SRM is implemented only in the Tropics, extra-tropics and high latitudes warm up during the climate engineered time.

The focus of this study was on the changes in extreme temperatures, defined here as the upper percentile (90th) and lower percentile (10th) ***of the 30-year temporal distribution of near surface***

temperature and precipitation at each grid-point. We also define the temperature and precipitation extremes based on the fixed threshold; namely dry-spell (consecutive dry days), frost-day and summer-day indices.

In the simulations investigated, ~~over much of the globe,~~ upper percentile (T90) temperature show small positive changes over tropics except northern hemisphere mid and high latitudes. ~~similar to the mean temperatures despite a small global-mean difference in changes. Exception is~~ In northern hemisphere high and mid latitudes, ~~where the~~ warm temperatures (T90) rise less than the mean, but the cold temperatures (T10) much stronger than the mean. This is consistent with the expectation, since SRM is effective only during polar day.

Defining temperature extremes by fixed thresholds, namely frost days as those where the minimum temperature is colder than the freezing point, and summer days as those where the maximum temperature is warmer than 25 °C, it is found that ~~neither SRM scheme effectively alleviates the decrease in the RCP4.5 frost days simulated for the mid- and high latitude continental regions, nor the scheme offset the increase in summer days.~~ ***the spatial patterns for the two SRM techniques differ.*** SULF better reduces the increase in the extra-tropics while SALT better reduces the increase in the sub-tropics. ***Globally, SALT is better in reducing the increase in the summer days compared to SULF. However Frost days are better mitigated in SULF experiment.***

The change in precipitation pattern mostly contrast each other in both the SRM techniques compared to the reference CTL period (2006 to 2035). ~~In terms of precipitation, it was found that the increases in mean and also P90 precipitation are in general well offset by both schemes.~~ ***In the tropical marine regions, the SALT scheme leads to an overall reduction in precipitation compared to CTL period. Extreme precipitation increases over land is more effectively reduced by SULF than SALT experiment.*** The geographical patterns of the P90 precipitation change show large variability which averages out when considering large regions.

~~In SULF, consistent among the models is a tendency in the low latitudes to have precipitation more concentrated in the equatorial tropics, indicating an intensification of the ITCZ, and to produce more precipitation over Eurasia. As for dry spells, the results are different for the two SRM climate engineering scenarios: The reduction in dry spell duration in G3-SSCE is mostly similar to what is produced for RCP4.5. G3, in turn, shows slight increases in the length of dry spells.~~

Extremes in temperature and precipitation vary with the season. We thus analyzed the percentile extremes separately for the boreal (Dec-Jan-Feb) and austral (Jun-Jul-Aug) winter seasons. Changes in the upper percentile (P90) for the annual distribution represent the changes of the summer seasons (JJA for northern hemisphere and DJF for southern hemisphere), and lower percentile (P10) is that of winter seasons (DJF for northern hemisphere and JJA for southern hemisphere). Results indicate that for both the SRM techniques there is net warming at the lower tail of the temperature distribution at high latitudes in the boreal and austral winter season.

Strong temperature increases are simulated after the ceasing of SRM climate engineering. SULF termination results in a rapid warming of entire globe, stronger over land in both tropical and extra-tropical regions than over oceans, and less strong over the Arctic for the 20-year time-frame analysed. SALT termination effect is more confined to the Tropics. Also precipitation responds strongly to the termination of SRM climate engineering measures with strong increases over land regions. ***In conclusion, termination effect of SULF on temperature is stronger than for the SALT experiment. SALT experiment termination result in more precipitation increases over ocean than land. Hence, in general termination of the SRMs result in the complete reversal of the patterns observed during the climate engineering time.*** Extreme values, both for temperature and precipitation, show stronger increases than the mean values for the termination effect.

Our results support some of the previous findings regarding the effectiveness of SRM over the lower latitudes compared to the high latitudes especially in winter seasons (Curry et al., 2014). Our results also reaffirm the fact that the regulation of global mean temperature does not necessarily control the regional climate (Ban-Weiss and Caldeira, 2010; Irvine et al., 2010). The SALT experiment result in a large increase in precipitation over land, which reinforces the result from an idealized scenario by Bala et al. (2011). Moist events over land is better mitigated in SULF than in SALT (Niemeier et al., 2013).

Our results show that SALT is more localised and more effective over the tropical regions. Most of the tropical marine regions show small, changes in extreme temperature compared to the CTL period. We found that the SULF experiment is effective in mitigating increase in extreme precipitation over land while SALT over ocean. In terms of the extremes based on threshold values, namely changes in the occurrence of frost days, summer days and length of consecutive dry days both the SRMs somewhat alleviates the effect of warming. But globally, the SALT experiment tend to reduce consecutive dry days and also reduce increase in summer days than the SULF experiment. Globally over land in temperature, termination due to SULF is more in magnitude than corresponding RCP4.5 and SALT scenarios. Warming over the lower tail of temperature distribution due to termination is much higher in magnitude compared to mean and higher temperature. By termination, besides an increase in precipitation over most of the globe, we also found a decrease in precipitation in SALT experiment over Indian subcontinent, North Africa as well as Europe.

~~This study has implications for the assessment of social costs of SRM climate engineering. Typically the extreme values of temperature or precipitation give rise to economic damages. Since changes in temperature and precipitation extremes in general are more or less equally affected by the two SRM climate engineering methods as the mean values, there is no substantial indication for costly side effects in this regard, except for the Arctic region. Other non-mitigated costs of greenhouse gas emissions, e.g. sea level rise, ocean acidification, have not been assessed here. Furthermore,~~

~~some effects are subject to considerable uncertainty, such as the exact geographical pattern of P90 precipitation change.~~

Overall, we conclude that the climate-change driven increases in the upper extremes of temperature and precipitation are simulated to be rather well mitigated by the two SRM climate engineering methods. However, we also find that the potential to mitigate effects of climate change by means of SRM differs around the globe and seasonally. Not very well dampened are in particular the increase in the mean temperatures is in the Arctic, and especially the increase in the lower temperature percentile in the Arctic winter. At the same time, it is not easily possible to locally engineer the climate by SRM methods, as the analysis of the SALT scenario shows. These findings indicate additional conflicts of interest between regions of the world if it should come to discussions about an eventual implementation of SRM.

Acknowledgements. The IMPLICC EU project (grant no FP7-ENV-2008-1-226567) and its researchers (Ulrike Niemeier and Hauke Schmidt, Max Planck Institute for Meteorology, Hamburg; Michael Schulz, met.no Oslo; Kari Alterskjær and Jón Egill Kristjánsson, University of Oslo); and as well as the GeoMIP project are acknowledged for conducting the simulations and providing the model output. This study was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) in the Priority Programme (Schwerpunktprogramm 1689 "Climate Engineering - Risks, Challenges, Opportunities?"), project "Learning about cloud brightening under risk and uncertainty: Whether, when and how to do field experiments", GZ QU-311/10-1 and GZ QU 357/3-1. The IPSL-CM5A climate simulations were performed with the HPC resources of [CCRT/TGCC/CINES/IDRIS] under the allocation 2012-t2012012201 made by GENCI (Grand Equipement National de Calcul Intensif), CEA (Commissariat à l'Energie Atomique et aux Energies Alternatives), and CNRS (Centre National de la Recherche Scientifique). Ulrike Niemeier is funded by SPP 1689 project Ceibral. MPI-ESM simulations were preformed on the computers of Deutsches Klimarechenzentrum (DKRZ). Helene Muri is funded by the Norwegian Research Council Project EXPECT (grant 229760/E10) and NorESM1-M computing resources was provided by NOTUR ***(NN9182K, NN2345K)*** and NORSTORE ***(NS9033K, NS2345K)***. ***T*****he authors are grateful to Simone Tilmes and two anonymous reviewers for their constructive and helpful comments.**

References

Alterskjær, K., Kristjánsson, J. E., and Seland, Ø.: Sensitivity to deliberate sea salt seeding of marine clouds – observations and model simulations, Atmospheric Chemistry and Physics, 12, 2795–2807, doi:10.5194/acp-12-2795-2012, http://www.atmos-chem-phys.net/12/2795/2012, 2012.

Alterskjaer, K., Kristjánsson, J. E., Boucher, O., Muri, H., Niemeier, U., Schmidt, H., Schulz, M., and Timmreck, C.: Sea-salt injections into the low-latitude marine boundary layer: The transient response in three Earth system models, Journal of Geophysical Research: Atmospheres, 118, 12,195–12,206, doi:10.1002/2013JD020432, http://doi.wiley.com/10.1002/2013JD020432, 2013.

Bala, G., Caldeira, K., Nemani, R., Cao, L., Ban-Weiss, G., and Shin, H. J.: Albedo enhancement of marine clouds to counteract global warming: Impacts on the hydrological cycle, Climate Dynamics, 37, 915–931, doi:10.1007/s00382-010-0868-1, 2011.

Ban-Weiss, G. a. and Caldeira, K.: Geoengineering as an optimization problem, Environmental Research Letters, 5, 034 009, doi:10.1088/1748-9326/5/3/034009, 2010.

Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, a., Seland, Ø., Drange, H., Roelandt, C., Seierstad, I. a., Hoose, C., and Others: The Norwegian Earth System Model, {N}or{ESM1-M–P}art 1: Description and basic evaluation of the physical climate, Geosci. Model Dev., 6, 687–720, doi:10.5194/gmd-6-687-2013, 2013.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change[Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2013.

Conover, W. J.: Practical nonparametric statistics, John Wiley and Sons, New York, 2 edn., 1980.

Crutzen, P. J.: Albedo Enhancement by Stratospheric Sulfur Injections: A Contribution to Resolve a Policy Dilemma?, Climatic Change, 77, 211–220, doi:10.1007/s10584-006-9101-y, http://link.springer.com/10.1007/s10584-006-9101-y, 2006.

Curry, C. L., Sillmann, J., Bronaugh, D., Alterskjaer, K., Cole, J. N. S., Ji, D., Kravitz, B., Kristjánsson, J. E., Moore, J. C., Muri, H., Niemeier, U., Robock, A., Tilmes, S., and Yang, S.: A multi-model examination of climate extremes in an idealized geoengineering experiment, Journal of Geophysical Research: Atmospheres, pp. n/a–n/a, doi:10.1002/2013JD020648, http://doi.wiley.com/10.1002/2013JD020648, 2014.

Dufresne, J.-L., Foujols, M.-a., Denvil, S., Caubel, a., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, a., Cugnet, D., Noblet, N., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, a., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, a., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, Climate Dynamics, 40, 2123–2165, doi:10.1007/s00382-012-1636-1, http://link.springer.com/10.1007/s00382-012-1636-1, 2013.

Efron, B. and Tibshirani, R. J.: An introduction to the bootstrap, Chapman and Hall/CRC, Boca Raton, 1998.

Fisher, R.: Statistical methods for research workers, Oliver and Boyd, Edinburgh, 1925.

Giorgetta, M. a., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, Journal of Advances in Modeling Earth Systems, 5, 572–597, doi:10.1002/jame.20038, http://doi.wiley.com/10.1002/jame.20038, 2013.

Govindasamy, B. and Caldeira, K.: Geoengineering Earth's radiation balance to mitigate CO_2 induced climate change, Geophysical Research Letters, 27, 2141–2144, 2000.

Irvine, P. J., Ridgwell, A., and Lunt, D. J.: Assessing the regional disparities in geoengineering impacts, Geophysical Research Letters, 37, 1–6, doi:10.1029/2010GL044447, 2010.

Jones, A., Haywood, J., and Boucher, O.: Climate impacts of geoengineering marine stratocumulus clouds, Journal of Geophysical Research, 114, D10 106, doi:10.1029/2008JD011450, http://doi.wiley.com/10.1029/2008JD011450, 2009.

Jones, a., Haywood, J., Boucher, O., Kravitz, B., and Robock, a.: Geoengineering by stratospheric SO_2 injection: results from the Met Office HadGEM2 climate model and comparison with the Goddard Institute for Space Studies ModelE, Atmospheric Chemistry and Physics, 10, 5999–6006, doi:10.5194/acp-10-5999-2010, http://www.atmos-chem-phys.net/10/5999/2010, 2010.

Jones, A., Haywood, J. M., Alterskjaer, K., Boucher, O., Cole, J. N. S., Curry, C. L., Irvine, P. J., Ji, D., Kravitz, B., Egill Kristjánsson, J., Moore, J. C., Niemeier, U., Robock, A., Schmidt, H., Singh, B., Tilmes, S., Watanabe, S., and Yoon, J.-H.: The impact of abrupt suspension of solar radiation management (termination effect) in experiment G2 of the Geoengineering Model Intercomparison Project (GeoMIP), Journal of Geophysical Research: Atmospheres, 118, 9743–9752, doi:10.1002/jgrd.50762, http://doi.wiley.com/10.1002/jgrd.50762, 2013.

Kharin, V. V. and Zwiers, F. W.: Estimating Extremes in Transient Climate Change Simulations, Journal of Climate, 18, 1156–1173, 2005.

Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C.: Changes in Temperature and Precipitation Extremes in the IPCC Ensemble of Global Coupled Model Simulations, Journal of Climate, 20, 1419–1444, doi:10.1175/JCLI4066.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI4066.1, 2007.

Kravitz, B., Robock, A., Boucher, O., Schmidt, H., Taylor, K. E., Stenchikov, G., and Schulz, M.: The Geoengineering Model Intercomparison Project (GeoMIP), Atmospheric Science Letters, 12, 162–167, doi:10.1002/asl.316, http://doi.wiley.com/10.1002/asl.316, 2011.

Kravitz, B., Robock, A., Forster, P. M., Haywood, J. M., Lawrence, M. G., and Schmidt, H.: An overview of the Geoengineering Model Intercomparison Project (GeoMIP), Journal of Geophysical Research: Atmospheres, 118, 13,103–13,107, doi:10.1002/2013JD020569, http://doi.wiley.com/10.1002/2013JD020569, 2013.

Latham, J.: Marine Cloud Brightening, Nature, 347, 339–340, 1990.

Latham, J.: Amelioration of global warming by controlled enhancement of the albedo and longevity of low-level maritime clouds, Atmospheric Science Letters, 3, 52–58, doi:10.1006/asle.2002.0048, http://doi.wiley.com/10.1006/asle.2002.0048, 2002.

Lenton, T. M. and Vaughan, N. E.: The radiative forcing potential of different climate geoengineering options, Atmospheric Chemistry and Physics Discussions, 9, 2559–2608, doi:10.5194/acpd-9-2559-2009, http://www.atmos-chem-phys-discuss.net/9/2559/2009/, 2009.

LIU Yunyun, LI Weijing, ZUO Jinqing and Zeng-Zhen, H.: Simulation and Projection of the Western Pacific Subtropical High in CMIP5 Models, JOURNAL OF METEOROLOGICAL RESEARCH, pp. 327–340, doi:10.1007/s13351-014-3151-2.1., 2014.

Muri, H., Niemeier, U., and Kristjánsson, J. E.: Tropical rainforest response to marine sky brightening climate engineering, Geophysical Research Letters, 42, 2951–2960, doi:10.1002/2015GL063363, http://dx.doi.org/10.1002/2015GL063363, 2015.

Niemeier, U. and Timmreck, C.: What is the limit of stratospheric sulfur climate engineering?, Atmospheric Chemistry and Physics Discussions, 15, 10 939–10 969, doi:10.5194/acpd-15-10939-2015, http://www.atmos-chem-phys-discuss.net/15/10939/2015/, 2015.

Niemeier, U., Schmidt, H., and Timmreck, C.: The dependency of geoengineered sulfate aerosol on the emission strategy, Atmospheric Science Letters, 12, 189–194, doi:10.1002/asl.304, http://doi.wiley.com/10.1002/asl.304, 2011.

Niemeier, U., Schmidt, H., Alterskjaer, K., and Kristjánsson, J. E.: Solar irradiance reduction via climate engineering: Impact of different techniques on the energy balance and the hydrological cycle, Journal of Geophysical Research: Atmospheres, 118, 11,905–11,917, doi:10.1002/2013JD020445, http://doi.wiley.com/10.1002/2013JD020445, 2013.

Robock, A., Oman, L., and Stenchikov, G. L.: Regional climate responses to geoengineering with tropical and Arctic SO 2 injections, Journal of Geophysical Research, 113, D16 101, doi:10.1029/2008JD010050, http://doi.wiley.com/10.1029/2008JD010050, 2008.

Robock, A., Marquardt, A., Kravitz, B., and Stenchikov, G.: Benefits, risks, and costs of stratospheric geoengineering, Geophysical Research Letters, 36, L19 703, doi:10.1029/2009GL039209, http://doi.wiley.com/10.1029/2009GL039209, 2009.

Schmidt, H., Alterskjær, K., Bou Karam, D., Boucher, O., Jones, a., Kristjánsson, J. E., Niemeier, U., Schulz, M., Aaheim, a., Benduhn, F., Lawrence, M., and Timmreck, C.: Solar irradiance reduction to counteract radiative forcing from a quadrupling of CO₂: climate responses simulated by four earth system models, Earth System Dynamics, 3, 63–78, doi:10.5194/esd-3-63-2012, http://www.earth-syst-dynam.net/3/63/2012, 2012.

Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, Journal of Geophysical Research: Atmospheres, 118, 1716–1733, doi:10.1002/jgrd.50203, http://doi.wiley.com/10.1002/jgrd.50203, 2013.

Taylor, K. E., Stouffer, R. J., and Meehl, G. a.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-11-00094.1, 2012.

Tilmes, S., Fasullo, J., Lamarque, J.-F., Marsh, D. R., Mills, M., Alterskjaer, K., Muri, H., Kristjánsson, J. E., Boucher, O., Schulz, M., Cole, J. N. S., Curry, C. L., Jones, A., Haywood, J., Irvine, P. J., Ji, D., Moore, J. C., Karam, D. B., Kravitz, B., Rasch, P. J., Singh, B., Yoon, J.-H., Niemeier, U., Schmidt, H., Robock, A., Yang, S., and Watanabe, S.: The hydrological impact of geoengineering in the Geoengineering Model Intercomparison Project (GeoMIP), Journal of Geophysical Research: Atmospheres, 118, 11,036–11,058, doi:10.1002/jgrd.50868, http://doi.wiley.com/10.1002/jgrd.50868, 2013.

Table 1. Climate extreme indices

Index	Description	Index definition	Units
T90, T99/ P90, P99	$90^{th}/99^{th}$ percentile	$90^{th}/99^{th}$ percentiles of the temporal distribution for given time period from temperature and precipitation	mm day^{-1}/ °C
T10/T1	$10^{th}/1^{st}$ percentile	$10^{th}/1^{st}$ percentiles of the temporal distribution for given time period from temperature	°C
CDD	Consecutive dry days index	Number of consecutive days where precipitation rate $<$1 mm day^{-1} in given time period	days year^{-1}
FD	Frost days index	Number of days per time period when TN $<$ 0°C	days year^{-1}
SU	Summer days index	Number of days per time period when TX $>$ 25°C	days year^{-1}

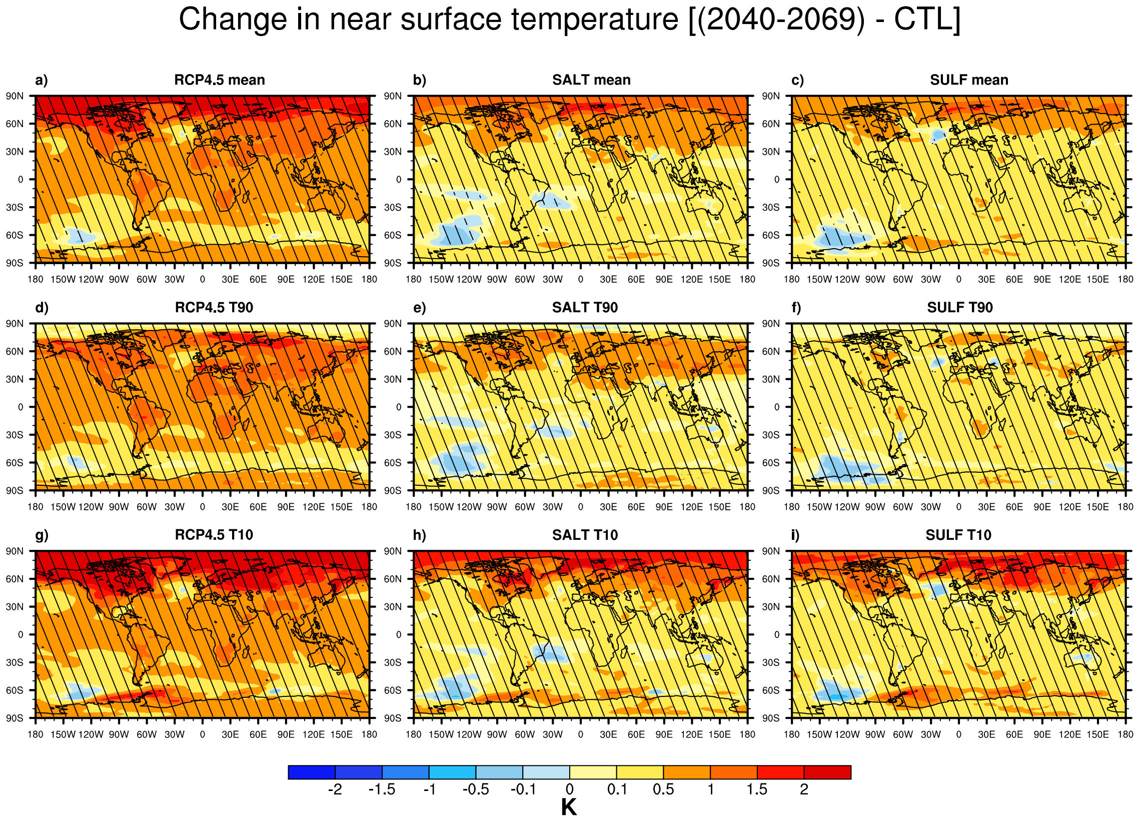

Figure 1. Multi-model mean change in near surface temperature (K) for RCP 4.5 (left column), SALT (middle) and SULF (right column) for 2040-2069 minus the RCP4.5 control period (CTL) (2006-2035). Panels a) to c) denote changes in mean values, d) to f) same as a) to c) but for the 90th percentile and g) to i) same as a) and c) but for the 10th percentile of the temporal distribution at each model grid point. Hatches denote regions where the changes are 95% statistically significant.

Table 2. Change in temperature 2040 to 2069 minus the RCP4.5 control period (2006-2035)

		Change in Temperature (K)							
		Global			Tropics	NH Mid-lat	NH High-lat	SH Mid-lat	SH High-lat
		All points	Land	Ocean	(30°N-30°S)	(30°N-60°N)	(60°N-90°N)	(30°S-60°S)	(60°S-90°S)
RCP4.5	Mean	0.77	1.05	0.65	0.73	0.96	1.76	0.44	0.45
	T90	0.74	1.07	0.61	0.75	1.02	1.03	0.46	0.34
	T99	0.76	1.07	0.63	0.77	1.02	1.03	0.51	0.34
	T10	0.85	1.18	0.71	0.70	1.12	2.58	0.45	0.66
	T1	0.92	1.26	0.78	0.70	1.37	2.68	0.53	0.72
SALT	Mean	0.31	0.47	0.24	0.17	0.57	1.01	0.19	0.26
	T90	0.28	0.48	0.19	0.14	0.65	0.63	0.18	0.21
	T99	0.25	0.44	0.16	0.10	0.58	0.64	0.20	0.18
	T10	0.38	0.54	0.31	0.20	0.63	1.43	0.22	0.41
	T1	0.45	0.65	0.37	0.25	0.81	1.52	0.27	0.49
SULF	Mean	0.30	0.39	0.26	0.25	0.35	0.80	0.23	0.27
	T90	0.26	0.35	0.22	0.25	0.33	0.33	0.22	0.13
	T99	0.23	0.29	0.20	0.22	0.28	0.25	0.24	0.09
	T10	0.38	0.51	0.32	0.24	0.50	1.35	0.27	0.46
	T1	0.44	0.58	0.39	0.27	0.64	1.50	0.37	0.44

Table 3. Change in precipitation 2040 to 2069 with respect to the reference RCP4.5 2006-2035 period.

		Change in Precipitation (mm/day)							
		Global			Tropics	NH Mid-lat	NH High-lat	SH Mid-lat	SH High-lat
		All points	Land	Ocean	(30°N-30°S)	(30°N-60°N)	(60°N-90°N)	(30°S-60°S)	(60°S-90°S)
	Mean	0.045	0.039	0.047	0.051	0.044	0.076	0.021	0.038
RCP4.5	P90	0.119	0.122	0.119	0.132	0.133	0.188	0.055	0.106
	P99	0.774	0.666	0.819	0.976	0.677	0.613	0.537	0.297
	Mean	-0.001	0.029	-0.013	-0.011	0.012	0.042	-0.003	0.009
SALT	P90	-0.004	0.096	-0.046	-0.041	0.046	0.113	-0.003	0.032
	P99	0.121	0.359	0.021	0.114	0.194	0.377	-0.008	0.083
	Mean	-0.001	-0.006	0.001	-0.008	0.004	0.029	-0.004	0.022
SULF	P90	0.008	-0.004	0.014	-0.006	0.015	0.075	0.004	0.058
	P99	0.182	0.172	0.186	0.194	0.204	0.192	0.149	0.106

Table 4. Change in CDD, FD and SU for the 2040-2069 period with respect to the CTL period.

	CDD (days/yr)				FD (days/yr)				SU (days/yr)			
	Global	Land	Ocean	Tropical	Global	Land	Ocean	Tropical	Global	Land	Ocean	Tropical
RCP4.5	0.15	0.31	0.08	0.48	-3.03	-4.91	-2.24	-0.26	11.51	9.68	12.28	19.13
SALT	-0.04	-0.29	0.07	-0.05	-1.69	-2.52	-1.34	-0.14	3.41	4.35	3.01	4.84
SULF	0.16	0.41	0.05	0.47	-1.34	-1.72	-1.18	-0.06	4.35	3.61	4.67	7.41

Table 5. Change in temperature and precipitation for the 2070-2089 period with respect to the 2050-2069 period.

		Temperature (in K)				Precipitation (in mm day^{-1})			
		Global	Land	Ocean	Tropical	Global	Land	Ocean	Tropical
RCP4.5	Mean	0.30	0.39	0.26	0.26	0.021	0.153	0.023	0.021
	T90/P90	0.29	0.37	0.25	0.30	0.069	0.542	0.075	0.081
	T99/P99	0.30	0.38	0.27	0.31	0.415	0.194	0.508	0.601
	T10	0.34	0.48	0.29	0.22	–	–	–	–
	T1	0.41	0.62	0.32	0.20	–	–	–	–
SALT	Mean	0.59	0.75	0.53	0.64	0.054	0.021	0.067	0.071
	T90/P90	0.59	0.73	0.53	0.73	0.152	0.070	0.187	0.207
	T99/P99	0.64	0.81	0.58	0.81	0.771	0.461	0.902	1.001
	T10	0.61	0.80	0.53	0.56	–	–	–	–
	T1	0.62	0.80	0.54	0.50	–	–	–	–
SULF	Mean	0.62	0.84	0.52	0.61	0.054	0.056	0.054	0.067
	T90/P90	0.65	0.93	0.53	0.65	0.135	0.167	0.121	0.157
	T99/P99	0.70	1.02	0.57	0.72	0.678	0.561	0.727	0.850
	T10	0.63	0.83	0.55	0.58	–	–	–	–
	T1	0.65	0.83	0.57	0.57	–	–	–	–

Table 6. Change in temperature and precipitation of SALT and SULF of the 2040-2069 period minus the corresponding period in the RCP4.5.

		Temperature(in K)				Precipitation (in mm day^{-1})			
		Global	Land	Ocean	Tropical	Global	Land	Ocean	Tropical
SALT - RCP4.5	Mean	-0.46	-0.58	-0.41	-0.55	-0.045	-0.009	-0.061	-0.062
	T90/P90	-0.46	-0.57	-0.41	-0.61	-0.123	-0.025	-0.165	-0.173
	T99/P99	-0.51	-0.63	-0.46	-0.66	-0.653	-0.307	-0.798	-0.862
	T10	-0.47	-0.63	-0.40	-0.49	–	–	–	–
	T1	-0.47	-0.61	-0.41	-0.45	–	–	–	–
SULF - RCP4.5	Mean	-0.47	-0.66	-0.39	-0.48	-0.046	-0.045	-0.046	-0.059
	T90/P90	-0.48	-0.71	-0.39	-0.51	-0.111	-0.125	-0.105	-0.138
	T99/P99	-0.53	-0.77	-0.43	-0.55	-0.592	-0.494	-0.633	-0.782
	T10	-0.47	-0.66	-0.39	-0.45	–	–	–	–
	T1	-0.48	-0.68	-0.39	-0.44	–	–	–	–

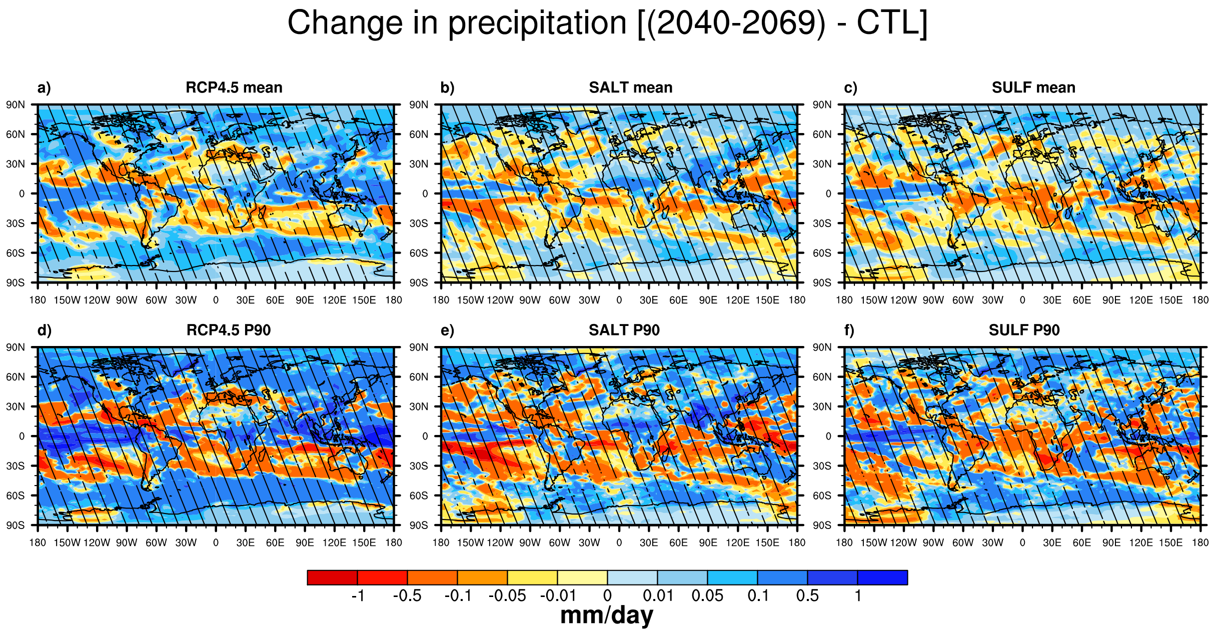

Figure 2. Multi-model mean change in precipitation (mm day^{-1}) for RCP 4.5 (left column), SALT (middle) and SULF (right column) for the 2040-2069 period minus the RCP4.5 2006 - 2035 control period (CTL). Panels a) to c) denote changes in mean values, d) to f) same as a) to c) but for the 90th percentile of the temporal distribution at each model grid point. Hatches denote regions where the changes are 95% statistically significant.

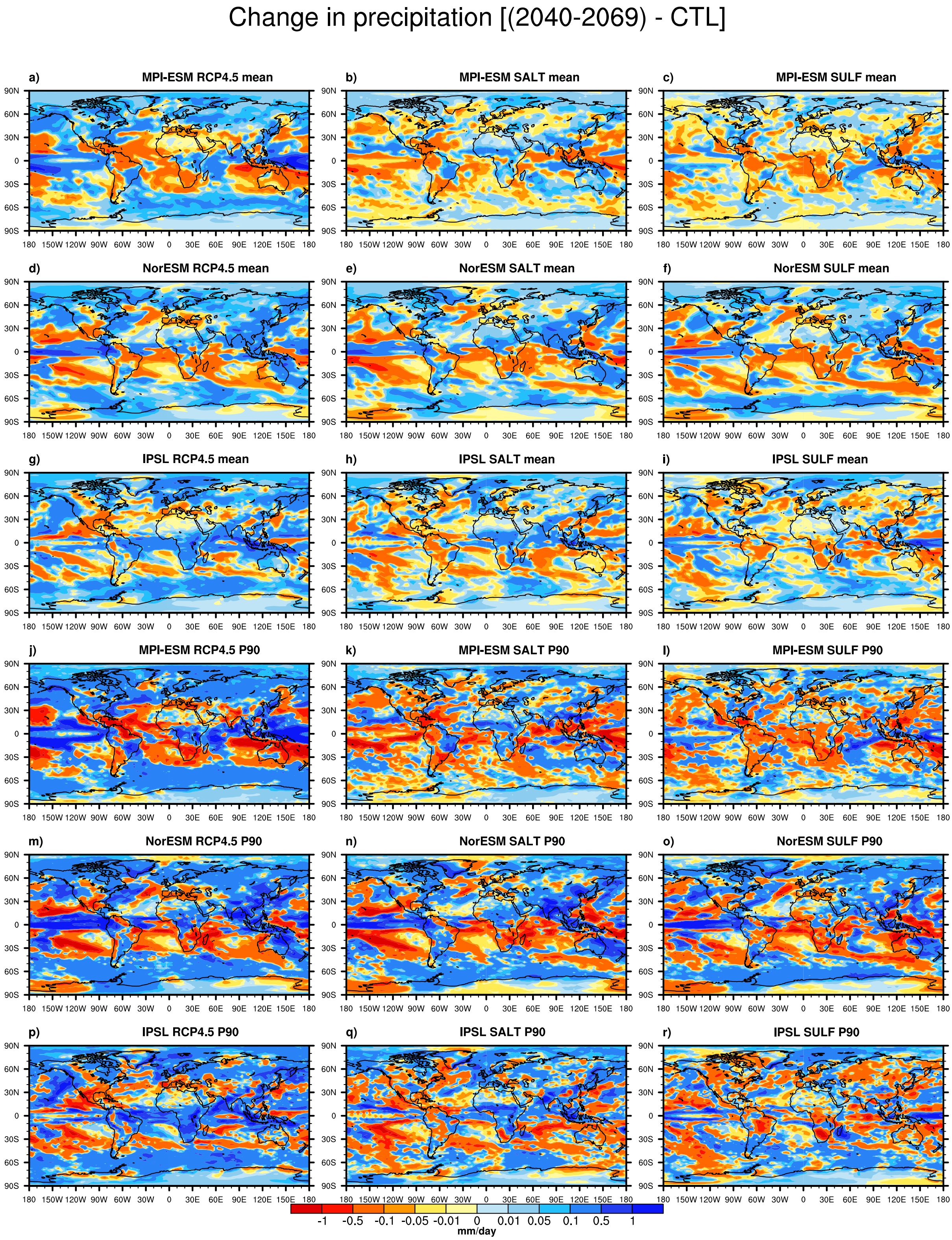

Figure 3. Change in precipitation (mm day^{-1}) for three scenarios RCP4.5, SALT and SULF and three models MPI-ESM, NorESM, IPSL for mean (first three rows) and P90 (last three rows) for the 2040-2069 period minus the RCP4.5 2006-2035 control period (CTL).

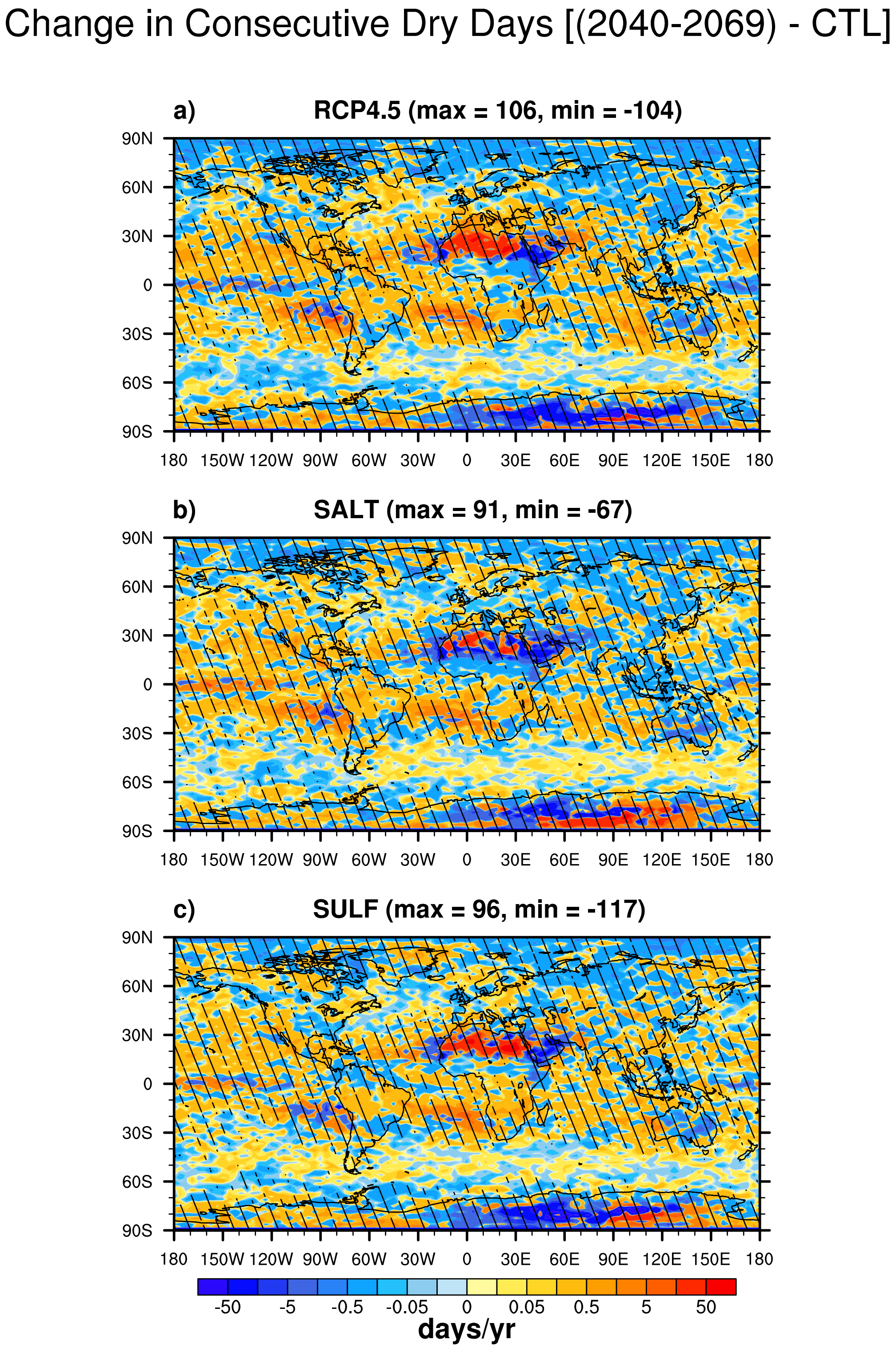

Figure 4. Multimodel mean of change in consecutive dry days RCP4.5 (top panel), SALT (middle) and SULF (bottom panel) for the 2040-2069 period minus the RCP4.5 2006-2035 control period (CTL) period. Hatches denote regions where the changes are 95% statistically significant.

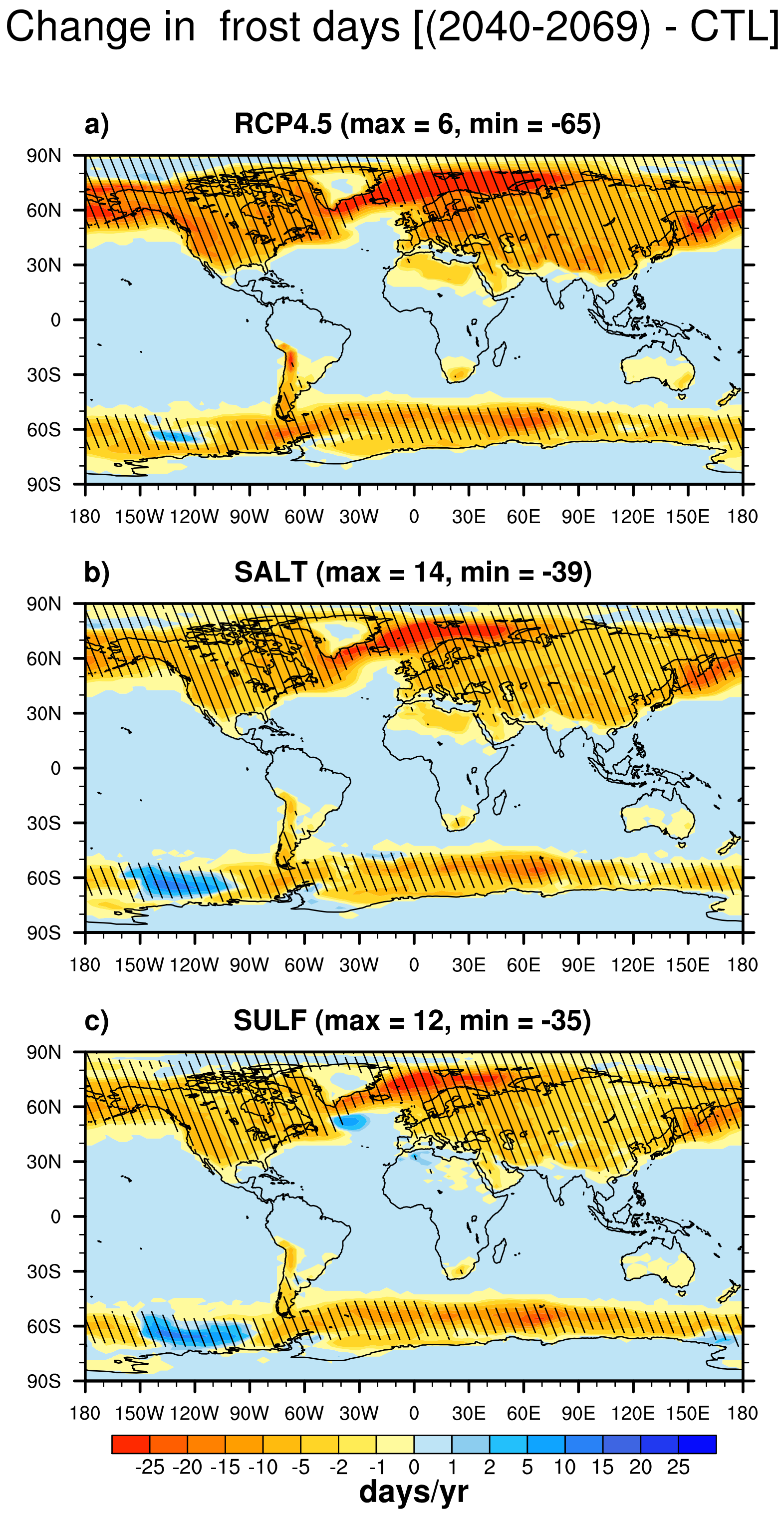

Figure 5. As Fig. 4, but for the mean change in frost days.

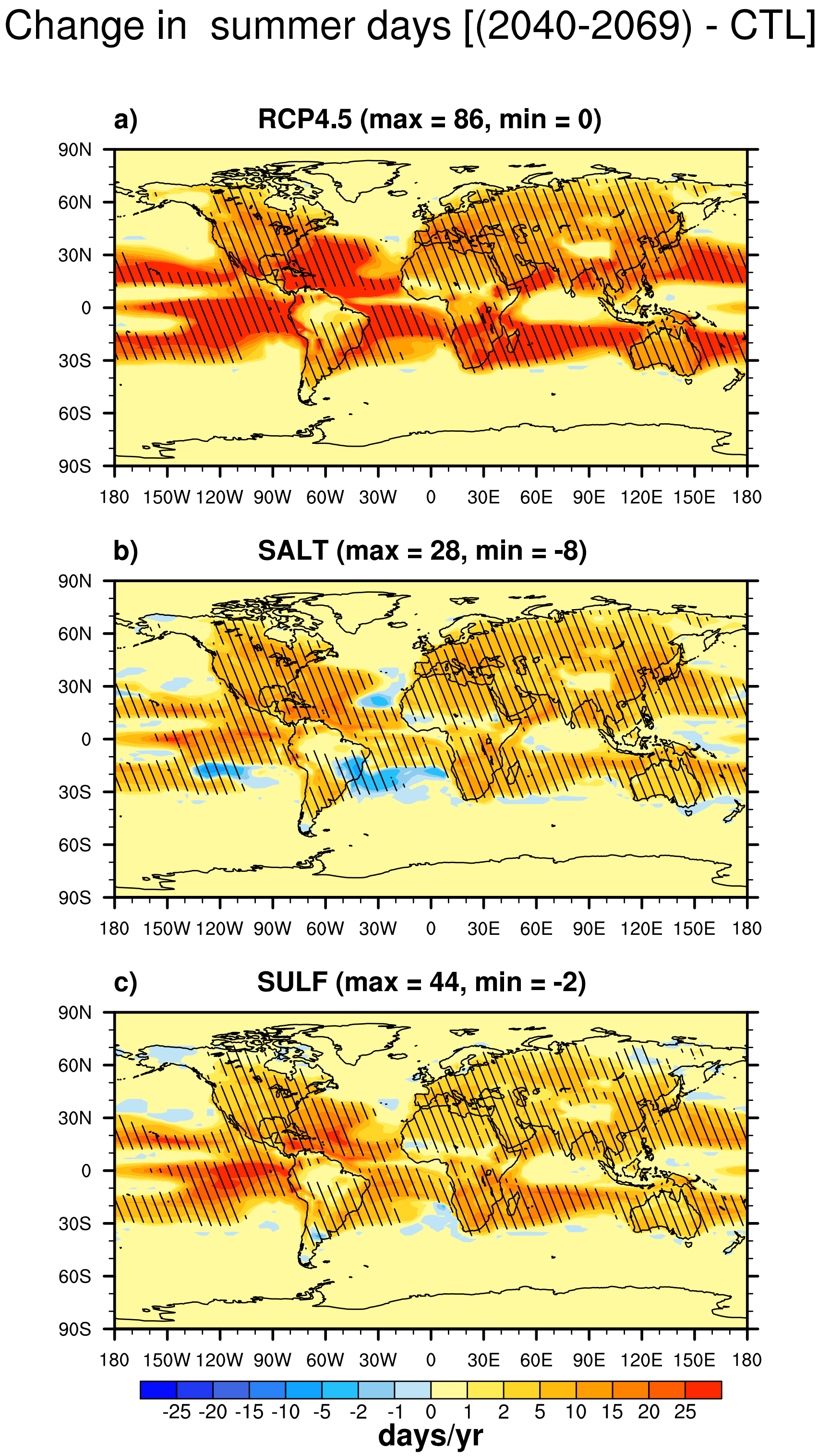

Figure 6. As Fig. 4, but for the mean change in summer days.

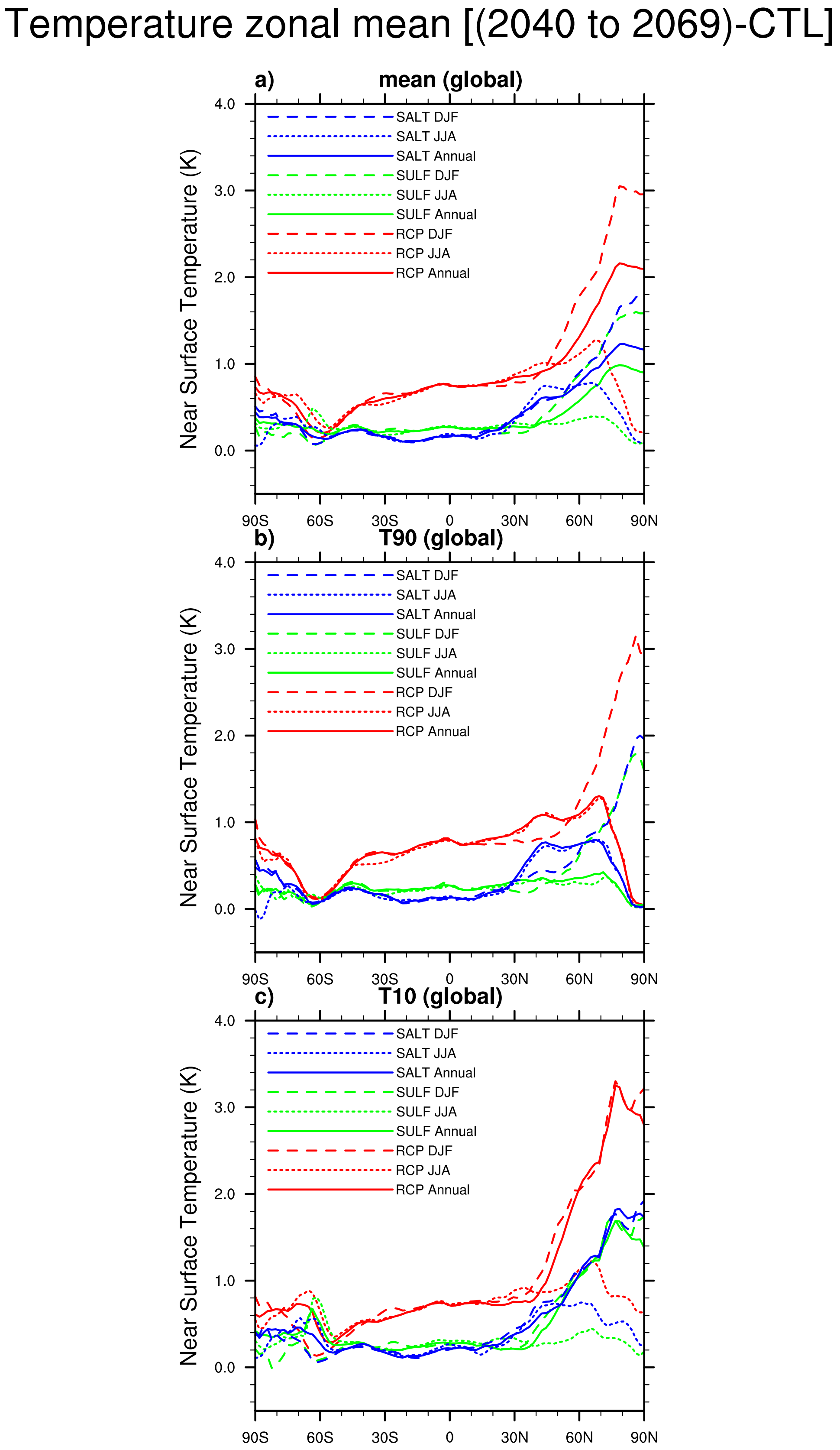

Figure 7. Multi-model zonal mean change in temperature (K) of RCP 4.5 (Red), SALT (Blue) and SULF (Green) for the 2040-2069 period minus the RCP4.5 2006-2035 control period (CTL) for annual mean, DJF and JJA season. The top panel shows changes in mean values, middle panel for the 90th percentile and bottom panel for the 10th percentile.

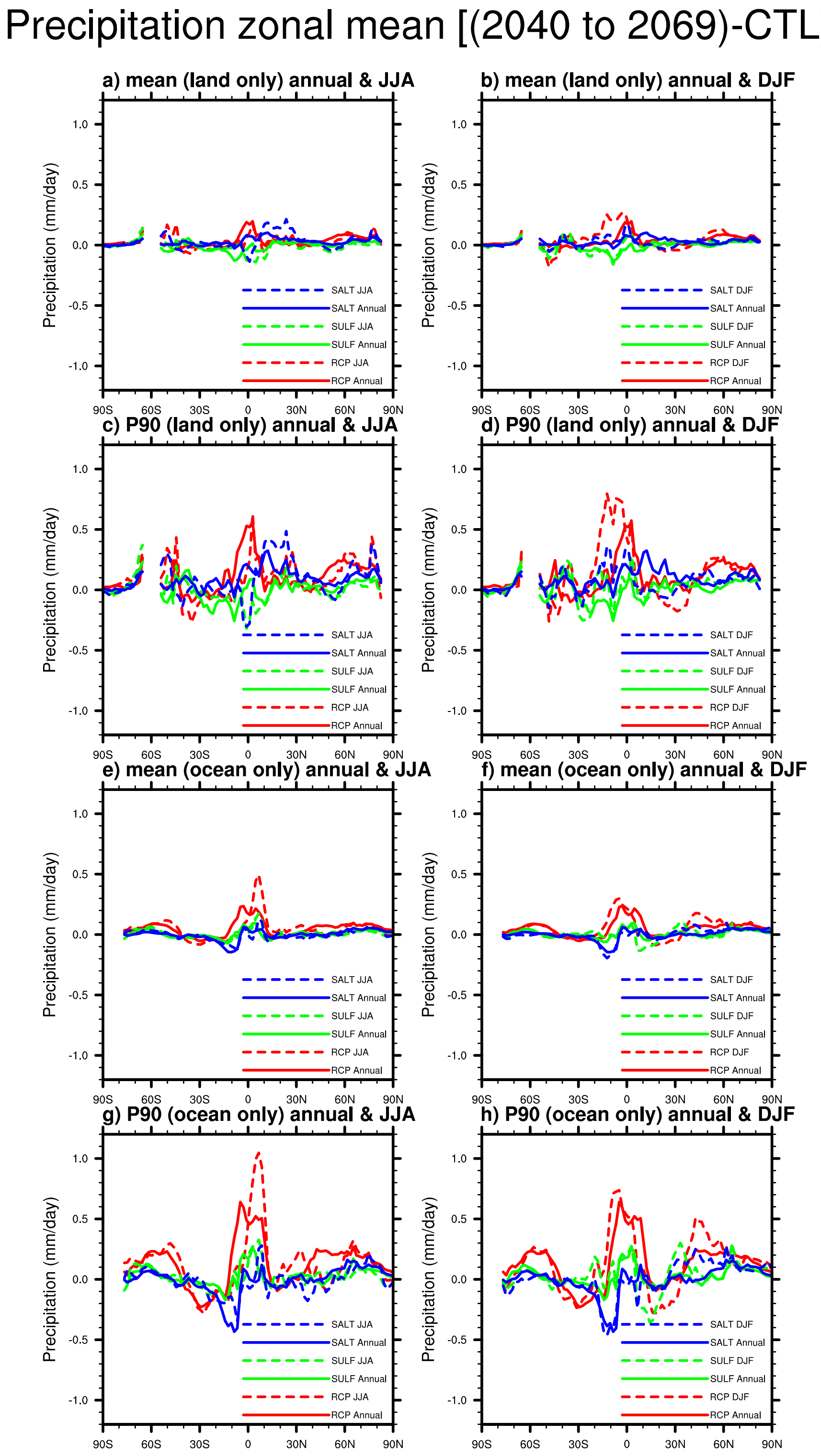

Figure 8. Multi-model zonal mean change in precipitation (mm day^{-1}) of RCP 4.5 (Red), SALT (Blue) and SULF (Green) for the 2040-2069 period minus the RCP4.5 2006-2035 control period (CTL) for annual mean, DJF and JJA season. Left column for JJA season and right column for DJF season. First two rows for mean and P90 of land only and the bottom two rows for mean and P90 of ocean only points respectively.

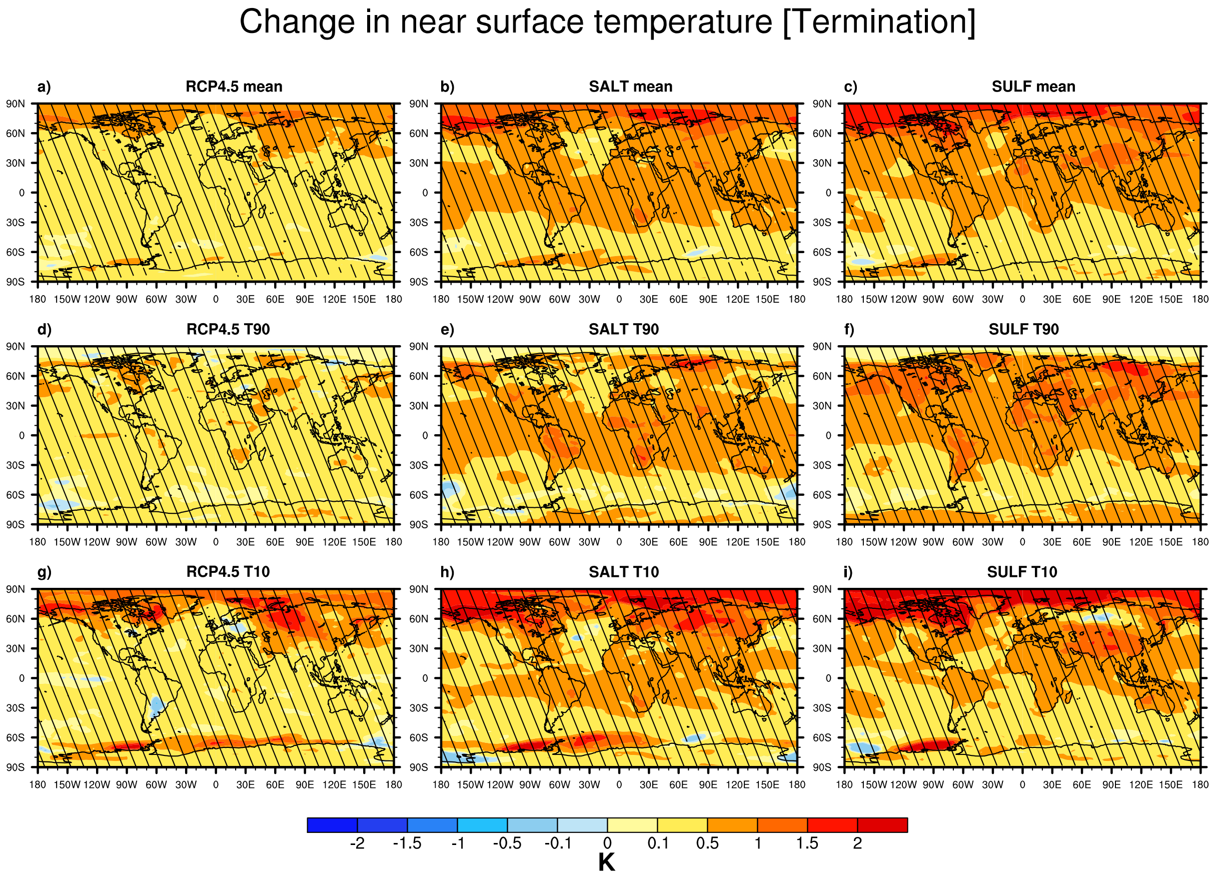

Figure 9. Multi-model mean change in temperature (K) during climate engineering termination period for RCP 4.5 (left panel), SALT (middle) and SULF (right panel). Panels a) to c) denote changes in mean values, d) to f) same as a) to c) but for the 90th percentile and g) to i) same as a) and c) but for the 10th percentile of the temporal distribution at each model grid point. Hatches denote regions where the changes are 95% statistically significant.

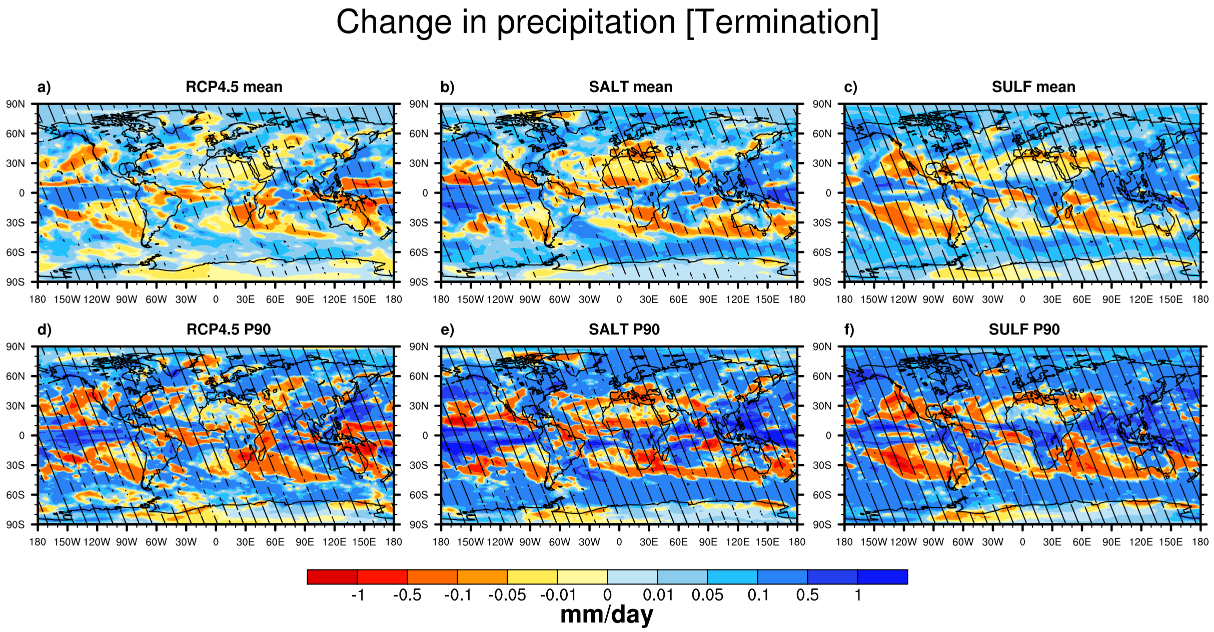

Figure 10. Multi-model mean change in precipitation (mm day^{-1}) during climate engineering termination period for RCP 4.5 (left panel), SALT (middle) and SULF (right panel). Panels a) to c) denote changes in mean values, d) to f) same as a) to c) but for the 90th percentile of the temporal distribution at each model grid point. Hatches denote regions where the changes are 95% statistically significant.