

Response to reviewers

We thank the reviewers for their comments and efforts towards improving this manuscript. Below, the reviewers' comments are given in blue italic font, and our response follows in black font. Text in small font are excerpts from the new manuscript.

Reviewer #2

Generally I thought the style of writing, and the quality of the figures is fine and would support publication in a modified form. The main issues I have with the paper in its current form are (1) it is unclear what has been done in parts of the manuscript; (2) as presented the results may be misinterpreted unless more care / accuracy is taken in the wording

We thank the reviewer for raising this general concern, and for the opportunity to improve our manuscript by making both methods and interpretation of results more accurate and transparent. We follow the reviewer's suggestions below to make these improvements.

The G4CDNC experiment is used. I am not familiar with this experiment myself, and had to do some searching in the text to find that it is a 50% increase in CDNC of low clouds. This should be brought to the fore, so that it is clear.

To clarify at an earlier stage that it is the G4cdnc experiment we focus on in this paper (and also what this experiment involves), we have made changes to the text that we hope improves it. Text marked in red below is added to the existing text:

In the abstract: "The nine contributing models prescribe a 50% increase in the cloud droplet number concentration (CDNC) of low clouds over the global oceans **in an experiment called G4cdnc**, with the purpose of counteracting the radiative forcing due to anthropogenic greenhouse gases under the RCP4.5 scenario."

At the end of page 2: "To attempt to alleviate this problem, the Geoengineering Model Intercomparison Project (GeoMIP) initiated a series of experiments where the models were to simulate MCB in a particular manner, with different degrees of complexity in the simulation design (Kravitz et al., 2013). The G1ocean-albedo experiment prescribes an increase in ocean albedo in the models at a rate intended to offset increasing global temperatures in response to a quadrupling of atmospheric CO₂ concentrations. The G4cdnc experiment **investigated in this work** simulates a **50 %** increase in the CDNC of low marine clouds as described in more detail in Section 2. Finally, the G4sea-salt experiment involves an increase in sea salt emissions over tropical oceans at a rate intended to produce a radiative forcing of -2.0 Wm^{-2} under the RCP4.5 scenario. The three experiments have an increasing degree of complexity or realistic representation.»

It would be useful to know how much of an overestimate applying the CDNC increase to all low clouds over the ocean may be to those that could realistically be subjected to a geoengineering scheme. What area of the earth does this correspond to in each of the different models? This is an important parameter to know: generating the requisite spray for MCB is no trivial task in terms of energy and scale-of-operation requirements.

GeoMIP G4cdnc is an idealized experiment, which is not meant to realistically represent what could be done with marine cloud brightening. They were designed to test the robustness of models in simulating geographically heterogeneous radiative flux changes and to see their effects on climate. Increasing CDNC by 50% over all oceans is clearly an exaggeration of what could credibly be done. A more realistic GeoMIP marine cloud brightening experiment, G4sea-salt, is analyzed by Ahlm et al. (2017, ACP). The results identify the areas in the participating models where the sea salt seeding is the most effective at brightening the clouds. This includes the decks of persistent low level clouds off the west coasts of the major continents in the sub-tropics. A brief discussion around this is now

included at the very end of Section 5. As for calculating the ocean area that could realistically (energy-, scale-of-operation- and economical-wise) be subjected to MCB, compared to the area with low cloud amounts for the different models, we do not believe that we have the right data to quantify this in a way that is defensible.

In the introduction there is reference to the finding of Alterskjaer et al., that seeding may lead to warming under certain conditions. Connolly et al. (2014, Phil Trans) discuss a similar finding in some detail. In short, the finding is that the Abdul Razzak et al parametrization does not work outside of the region they were originally tested and, in some ranges of parameter space, increasing the aerosol number concentration can lead to a spurious reduction in droplet number concentration. I note from table 1 that some of the models used in this study employ the same parameterisation, including NorESM. While CCN activation schemes are not a focus of this particular paper, some of the reasons for the effects talked about in the introduction may be due to CCN activation schemes, so it is an important point for other people working in this area.

Thank you for pointing us towards the Connolly et al paper. This is an interesting point, and we agree it deserves mentioning. We have now added a sentence in the Introduction, referring to the Connolly et al. paper and pointing to the importance of the models' activation schemes. We have also added a small paragraph on this topic (second last paragraph) in Section 5. That being said, in this experiment, models did not increase their aerosol number concentrations, but rather increased their cloud droplet number concentrations directly, essentially bypassing this issue.

In the introduction: However, comparison between studies has been difficult partly due to different experimental design. Also, in studies based on one or only a few models, results will be very dependent on the particular models' parameterizations. For instance, Connolly et al. (2014) suggested that the unintended warming found in Alterskjær and Kristjánsson (2013) for seeding of some particle sizes (as mentioned above) was likely to be an artifact of the cloud parameterization scheme used in the model.

In Section 5: As seen in Table 1, there are large differences between models as to how the aerosol-cloud processes are parameterized, and this presumably has an impact on the MCB climate response. For instance, Morrison et al. (2009) found that the complexity of the cloud schemes had large impacts on stratiform precipitation; Penner et al. (2006) compared several models and concluded that the method of parameterization of CDNC can have a large impact on the calculation of the first indirect effect. CSIRO-Mk3L-1-2 and MPI-ESM-LR are the only models that have prescribed CDNC levels in their simulations, and these models stand out as having ERFs well above any of the other models. While an evaluation of which liquid cloud parameterizations are more appropriate is beyond the scope of this paper, this might be an indication that prescribing CDNC levels can lead to exaggerated responses to marine cloud brightening.

I was not familiar with the Gregory regression method (section 3) until I did a search on the Internet. I think the original source should be cited. Would it be more complete to show / present the correlation coefficients associated with the analysis, to allow the reader to assess its suitability?

In Section 2.2, we already state that "To estimate the effective radiative forcing of increasing CDNC by 50% in the different models, we use the method of Gregory et al. (2004), whereby the top of atmosphere radiative flux imbalance is regressed against the global annual average surface air temperature change compared to the RCP4.5 simulations."

To remind the reader of this explanation, we now refer back to this section when the ERF results are presented. We have also added an extra sentence after the above extract from Section 2.2, to further detail how this Gregory regression was performed;

To put the most weight on the initial changes, we use annually averaged values for years 2020-2030, but decadal mean values for years 2031-2070.

Regarding correlation coefficients between change in temperature and TOA radiative flux imbalance, we have now added a column in Table S1 showing the coefficients for each model as well as

indications of statistical significance. This table is now referred to in the results, in the section where the ERF are presented. Unsurprisingly, the models with the weakest ERF have the weakest correlations.

	Correlation between surf. temp. change and TOA rad. flux imbalance	Correlation between surf. temp. change and baseline low cloud fraction	Correlation between surf. temp. change and baseline LWP for oceans
	Global mean correlation	Spatial correlation	Spatial correlation
BNU-ESM	-0.60	-0.59	-0.49
CanESM2	-0.83	-0.09	+0.43
CSIRO-Mk3L-1-2	-0.92	-0.18	+0.44
GISS-E2-R	-0.26*	+0.08	+0.04
HadGEM2-ES	-0.69	+0.06	+0.50
IPSL-CM5A-LR	-0.35*	+0.03	+0.27
MIROC-ESM	-0.56	-0.49	+0.47
MPI-ESM-LR	-0.85	+0.16	+0.20
NorESM1-M	-0.29*	-0.51	-0.32
Median	-0.55	-0.13	+0.45

Table S1: For the global mean correlations (leftmost column), globally and annually averaged temperature changes over the years 2020-2060 are correlated against corresponding changes in TOA radiative flux imbalance. For the spatial correlations (two rightmost columns), global matrices of annual mean changes (average of years 2020-2060) are correlated against baseline (20 first year of RCP4.5) values. All correlations except those marked by an asterisk are statistically significant at the 95% level. The lowest row shows correlation between model-median quantities (not the median of the individual model values above).

Fig 4a was not referred to directly in the text.

Figure 4a was referred to in the first paragraph of Section 4 (previously 3): “The model median geographical pattern of this radiative perturbation is shown in Fig. 4a.”

Although the models in the paper are important tools for climate assessment, the major issue with this kind of analysis is that the processes under investigation are not resolved at high enough grid resolution. There should be some caveat in the paper so that readers who are unfamiliar with the details are not mis-lead.

We absolutely agree that in spite of all the advantages of using Earth system models in this type of study, this is a clear disadvantage. We have added the following paragraph to the end of Section 4, covering the reviewer’s comments on both model resolution and overestimation by seeding clouds above all oceans:

A caveat of the present study is that the spatial resolution of global climate models is too coarse to resolve important processes such as convection or precipitation formation. These processes are instead parametrized, which may lead to unknown artifacts in the responses, dependent on the specific formulations used (see e.g. Clark et al. (2009) who compared precipitation forecast skills for convection-allowing and convective-parametrized ensembles). It is also important to point out that the G4cdnc experiment was not designed to give a realistic representation of the magnitude of cooling and other climate responses to MCB, but rather to test the robustness of models in simulating geographically heterogeneous radiative flux changes and to see their effects on climate. Increasing CDNC by 50% over all oceans is clearly an exaggeration of what could credibly be done. A more realistic GeoMIP marine cloud brightening experiment, G4sea-salt, is analyzed by Ahlm et al. (2017). Their results identify the areas in the participating models where the sea salt seeding is the most effective at brightening the clouds. This includes the decks of persistent low level clouds off the west coasts of the major continents in the sub-tropics. Alterskjær et al. (2012) investigated the susceptibility of

marine clouds to MCB based on satellite as well as model data, and reached a similar conclusion; large regions between 30°S and 30°N, and especially clean regions in the Pacific and Indian oceans, along with regions in the western Atlantic, are susceptible.

The statement: “Our results suggest that liquid cloud parameterizations ought to be of appropriate complexity in order to attempt to model marine cloud brightening and the climate response.” I did not see how this conclusion was arrived at. How do you know that liquid cloud parameterisations are complex enough?

The reviewer is of course right – we have not performed any evaluation of different parameterizations. We have now modified our language from “our results suggest” to “we hypothesize”;

We hypothesize that liquid cloud parameterizations ought to be of appropriate complexity in order to attempt to model marine cloud brightening and the climate response.