Response to Referee 2

We thank the referee for their useful comments

This brief communication shows results on the likelihood of an ice-free Arctic under the low-warming IPCC targets, using experiments with solar radiation management with the HadGEM2-ES model. It shows that limiting the global temperature rise to around 1.5◦C reduces the likelihood of an ice-free Arctic to 0.2%, compared to 43% for warming around 2◦C. This study is interesting and shows similar results to previous work using other models, in particular with the CESM in Sandersen et al. (2017) and the CMIP5 RCPs in Screen and Williamson (2017). This shows that those results are robust across different models. So while the results are not new, they are worthwhile to be published, in particular in light of the upcoming IPCC report on the low warming targets. However, I find the submitted manuscript not publishable in its current form. Even for the category "Brief communication" I find the results section too short – it is shorter than the methods and about the same length as the conclusions and introduction individually. This paper really seems to be pushing the "least publishable unit" into the unacceptable category, and I cannot support that. There is lots of additional analysis that could easily be done with these simulations, even expanding directly on the figures shown. For example Figure 2 is not discussed much and that could be expanded to add new insights not previously shown to this discussion (the spatial aspect of the ice edge under these targets). That said, I have a long list (as long as the paper...) of things that need to be addressed, detailed below. Furthermore, the article has many typos, fuzzy figures, and imprecise and confusing statements and captions. The figures also need to be revised. Overall, I see merit in publishing these results as brief communication, but not in the present form, and therefore recommend major revisions.

- Results section expanded

- Figure quality and information content improved

- More motivation provided

Detailed comments:

Page 1, Line 8-9: Why would we want to "reduce the internal variability" if we want to produce a probability density function of an ice-free state? Is this a typo, and it should be "deduce"?

Clarified with 'improve the signal to noise associated with the internal variability'

Page 1, line 18: Seems to be missing a noun at the end of the sentence? Should be "one such system"

This has been corrected

Page 1, Line 25: Should be "uptake", not "up-take"

This has been corrected

Page 1, Line 28: Should be “increased”, not “increase”

This has been corrected

Page 1, Line 3: Why “then” here, doesn’t really make sense, as it does not relate to the previous sentence (“Sea ice then hits its smallest extent sometime in September…”)

This has been corrected

Page 2, Line 13: two .. at the end of the section

This has been corrected

Page 2, Line 13: I think the introduction needs to reference the of previous work on this topic, and not make it sound like this hasn’t been done yet (Screen and Willamson, 2016, Sanderson et al. 2017)

This has been corrected

Page 2, Line 15: Degree symbol is incorrect, should be ◦ not â˚U¸e . This is correct for◦ C, but not for degree related to N/S.

This has been corrected

Page 2, Line 30 to Line 2 Page 3: Is there just one simulation for each RCP? Sounds like it, since just one year is given for reaching these target temperatures.

The appropriate sulphate loading was calculated just once for each target temperature from the mean of the relevant 4-member RCP scenario ensemble. This was initially mentioned P3 line 1+2 but has been moved up and expanded for clarification. Thus, it is indeed the case that ensemble mean is used to determine the dates for the temperature thresholds.

But then it says “For CMIP5 a historical + scenario initial condition ensemble of 4 HadGEM2-ES members was completed.”. So does the ensemble mean of these four cross the threshold at that time? That needs to be made clearer. The figure 1 only shows one line for each RCP. But then says that that is the ensemble mean of 4 RCP members. But then each member of the 16 member ensemble for reduced temperatures is shown. Why is that done? It makes no sense to me. The 4 ensemble members for each RCP should be shown as well, so it is clear whether the new ensemble is outside the internal variability of the initial RCP. That would help to substantiate the statement on page 3, line 5 that (The resulting ensemble spread in global mean temperature is larger than that for the initial 4-member ensemble). Not that I doubt that, but it is always nice to see this, and to see how much larger the spread is. Adding the ensemble members for the RCPs would also explain why some of the red lines start well above or below the black line.

As per the reviewer’s suggestion, the figures have been updated to show all 4 RCP scenario ensemble members (rather than the mean as before)

Page 2, Line 30 to Line 2 Page 3: What is the reference period for the 1.5C and 2C and 2.5C warming? That’s not mentioned and various papers use different periods.

The reference for pre-industrial is the mean of the parallel pre-industrial control simulation for the period 2005-2100. Pre-industrial refers to 1860 forcing and this has now been added to the ensemble description summaries on page 3.

Page 3, line 3-6: This seems like quite a large perturbation, mixing the ocean initial conditions with an atmosphere that it isn't equilibrated with it. Has this approach been tested before? I have only heard of others using much smaller initial condition perturbations.

The imposition of a random atmosphere on an ocean state has been demonstrated by Griffies, S. & Bryan, K. Climate Dynamics (1997) 13: 459. https://doi.org/10.1007/s003820050177. The atmosphere equilibrates in days and the ocean loses memory of its initial state within 20 years.

This is now mentioned, and cited, in the text.

Page 3, line 7-10: Why were these target temperatures of 1.3 and 1.7C chosen, when the IPCC target is 1.5C?

The idea was to investigate the region around the 1.5C target to sample for thresholds in the temperature sensitivity. Internal variability was expected to bracket the 1.5 target.

Page 3, Results section: This is a VERY short result section, shorter than the methods and about the same length as the conclusions. That seems to me to be very much below the "least publishable unit" standard, and while interesting, I can't support publishing research in such minimal increments. There is lots of additional analysis that could easily be done with these simulations related to Arctic sea ice to make this acceptable, or even expanding the current description of the results. I know this is a "Short communication", but even for that I find the results too brief.

The results section has been expanded to include further discussion of the updated Figure 2.

Page 3, line 29: Extra "In" at start of sentence

Corrected

Page 3, line 30: Why is the "quantitative result described here" only listed a paragraph later? That makes it hard to compare it with the other studies. Which should be discussed in the introduction already, in my opinion.

Page 4, line 4: Should be "to", not "do"

Corrected

Page 4, Line 1-7: Didn't Screen and Williamson (2017) look at the probability of just one ice-free Arctic, while here and in Sandersen et al. (2017) the overall probability is assessed?

Screen and Williams use a regression line to produce a single value but then apply Bayesian statistics to obtain a probability.

Page 4, Conclusions in general I think these need to be more carefully written. I don't think the statement "since the Arctic sea ice in CMIP5 models is effectively in equilibrium with the instantaneous global temperature" is defensible, as that would mean that Arctic sea ice in a given

year is dependent on the global temperature. The cited work shows that this is true in the long-term sense, but not for year-to-year variability. Furthermore, this doesn't apply to all properties of Arctic sea ice, so this needs to be more precisely formulated (September, extent).

Agreed. This comment and associated references has been removed

Generally: Please use a consistent number of significant digits, and not 1.5C and 2.5C but then 2C.

Consistency corrected throughout

References: Please check these for inconstant formatting – not my job as a referee to fix those

Done

Figure 1: As mentioned earlier, this figure should either show the ensemble mean or the individual members, not a mix of both. Unless there is a good reason for that, but none is articulated. Furthermore, the figure is fuzzy, so it doesn't seem to be saved in a vector format. Need to be switched out to be acceptable.

Figure 2: It is not clear to me what this figure really shows. First of all, this figure does not show the "spatial pattern of the Arctic sea ice extent", but the "spatial pattern of the sea ice edge". Which is only the 15% contour, and hence is very noisy. Furthermore, it is unclear to me what both the red and black lines are. It says early on that it shows the "mean of years 2080-2099", for certain temperature thresholds in certain simulations. But then it goes on to say it shows "the mean (years 2006-2025) of the four" RCPs. I also don't understand "starting at +1.5◦C above preindustrial in RCP2.6 (left), +2◦C in RCP4.5 (centre) and +2.5◦C in RCP4.5 (right)." This clearly needs to be rewritten to make sense to a reader who hasn't made the figure. It says the red lines are for the 16 individual ensemble members, but those are not the RCPs. Going back to how the ensembles are defined, I think that's what the red lines show, but then these are not for temperatures 1.5C, 2.0C and 2.5C, but for 1.5C, 1.3C and 1.7C? Panels also should have labels (a) and (b) and (c) for ease of reading/referencing. Furthermore, based on what I see, the black line in each panel is different. I assume that is because it is for the different temperature thresholds in the two RCPs mentioned above, and probably for 2006-2025, as they look like present-day. But why use different baselines for each, and what is the third one, since there are only two RCPs? That all makes no sense to me. It would make more sense to use the present-day period for all of them, potentially both for the model and observations, to show how realistic the model is. And then show the ice edge for the 2080-2099 for each of the three ensembles. But since their temperatures aren't that different, and are lumped together in the next figure, why they should be shown is unclear. To show the difference between the ones that end at 1.5 and 1.3 degree C versus ensemble 3 that ends at 1.7 degree C?

Figure 3: This figure is also fuzzy and out of focus. It needs to include a higher resolution figure to be considered for publication. Furthermore, it is unclear what exactly it shows. The caption needs to state what the size of the boxes is (+/- 0.1 degree C around the mentioned thresholds, according to my reading of the graph), as it does not show the probability right at the quoted temperature thresholds. Otherwise, this figure is the most interesting one. I do wonder how it compares to the transient RCPs with that model though, and hence how much it adds to just using those. Can those be added here? They all cross the same temperature range, so could be included in the PDF. Currently it seems to only include the 3 ensembles with 16 members each (3*16=48). That leads to

many fewer members that are in the 2 degree warming box than in the 1.5 degree warming box, and hence could be influencing the probability distributions shown in panel 2. Adding the RCPs in here would help to rectify this. Panels also should have labels (a) and b.

The figures have all been updated in line with all the above suggestions