We would like to thank Doug for the time taken to provide such a detailed and constructive review of our manuscript. We have addressed his comments by making the following changes.

Overall this is an important paper that makes a valuable contribution and should ultimately be published, though there are some things in the presentation that could be improved.
Thanks for the time taken to review the work and for your appreciation of its main findings.

A particular example (and this is redundant with my comments below) is the omission of much on the targets used in the simulation design for both SAI datasets considered, the fact that those targets are simply choices, and being clear that the reason that the uncertainties discussed herein manifest themselves as a big difference in the distribution across latitudes of injection rates (and corresponding aerosol distribution) rather than as a difference in the eventual interhemispheric temperature gradient T1 is a result of the choice of those targets. (Indeed this paper would likely have been easier to write if the simulations had been conducted with only a T0 target!) Doesn't really affect the paper, but it would be confusing as written to someone who didn't know the datasets well already.
- Agreed this fact was assumed in much of the original text and now has been made more explicit and the discussion of targets used has been expanded. While the original targets used by GLENS were described in lines 84-86 of the original draft, the discussion was brief and there was no mention of their subjectivity or the implications of a different set of targets being used for model dependence. We have therefore added discussion to the text acknowledging the subjectivity of the targets used and the consequences for the interpretation of uncertainty. We also mention some of the motivation behind the targets that were chosen as clearly there are consequences for dropping T1 and T2. We agree that had the simulations been conducted with only a T0 target, this paper would likely have been about the potentially broad range of climate responses rather than the potential range of injection distributions.

Also, while it again doesn't ultimately affect any conclusions from the paper, the eof-based metric for evaluating AMOC strength is more complicated to interpret than implied herein since a shift in the latitude of overturning circulation, for example, will manifest as a reduction in the magnitude of the principal component… without more work it's hard to disambiguate true reductions in the strength of the overturning (e.g., looking at maximum of streamfunction) from where that maximum occurs.
- Agreed. To address this issue we have further detailed the spatial structure of AMOC changes, expanded its discussion, and added a figure to the supplement. Doug is right to point out that it is both difficult to compare quantitatively the time series in Figure 5 and discern their relevance to the spatial structure of AMOC changes. We have therefore performed additional analysis and added a figure to the SI documenting the ensemble mean trends in EOF1 sealed by PC1. The discussion has also been expanded to point out that changes are PC1 are normalized and therefore difficult to compare across experiments.

Finally, it might be worth noting that the original simulation in Kravitz et al 2017 that pioneered the strategy used in both GLENS and ARISE yielded a different hemispheric asymmetry in the injection rates than GLENS, despite both being in CESM1(WACCM); the only difference being a change in the land model from CLM4 to CLM4.5. Neither the cloud adjustment to CO_2 nor the tropospheric aerosol changes can be responsible for that. I don't recall having looked at AMOC in that run, but if it's still around, that might be worth looking at… at least worth acknowledging. Might be something in fast response to CO_2 again, but for vegetation in the land model… that is, there may be other factors beyond the 3 identified here that are also relevant when looking at different models

Yes, this is a nice (and somewhat perplexing) result and is now mentioned in the introduction of model dependence.

There've been a few recent papers on scenarios for SRM, and NCAR is organizing a workshop on the topic soon; I think the scope of what's envisioned on SRM scenarios is so much broader than the scope in this paper that including that word in the title is potentially misleading… yes, scenario is relevant to what you're doing, but you certainly aren't remotely covering the scope of scenario dependence. The title is not wrong, just may not be what people think of when they read the words.

Title: we agree that some readers might have a different context for the scenario dependence cited and we have made changes to the title and throughout the text to be explicit in citing the "background scenario" to distinguish it from the SAI scenario. That said, context is provided in the abstract, which should avoid confusion, and we have further clarified that we're referring to the background scenario.

Abstract line 1, I don't think "model dependence" is itself an uncertainty, but rather demonstrates the presence of uncertainty. (Nitpicky, perhaps…)

- Abstract Line 1: We're not sure we understand the nature of this comment. Model dependence seems equivalent to model structural uncertainty (and thus is a component of overall uncertainty) unless certain models can be discounted outright? Is the suggestion that if we can rule out a model then model disagreement is no longer a component of uncertainty?

L12, is the AMOC behavior "associated" with the rapid adjustment? (If not, that's the wrong word). I gather from the later discussion that that is a hypothesis for why AMOC responds as it does, but not clear that that is proven.

- Line 12: Agreed. The word "associated" is removed as the reference is a bit unclear.

L14, shouldn't be over the timeframe of deployment, but of research prior to deployment. (And the statement is true regardless of whether the target of SAI is to stay below 1.5C or not, so should reword anyway)

- Line 14: Agreed. The statement in question is removed as it presumes an approach for dealing with the uncertainty identified that is subjective.

L15 (and elsewhere), the adjectives “significant” in front of flexibility, and “large” in front of uncertainties, don’t actually convey any information. What makes an uncertainty “large” or “small”? It would seem to me that the useful sense of that word ought to be whether some resolutions of the uncertainty would result in a choice to deploy and other resolutions wouldn’t. If any possible value of the uncertainty doesn’t change whether or not someone would choose to deploy, is it meaningful to call it “large”? And, given that that is hard to prove (and I don’t think proven here), not clear to me that there’s any basis for using these adjectives. Ditto L25, for example. (And one could level the same criticism at the IPCC report too, where qualifiers are better defined for climate change, but used arbitrarily for SRM.) Given that these words convey no actual information, but they do convey emotion, I would argue that such qualifiers don’t belong in a scientific journal article (though I’m aware that this is a generic problem with many papers throwing words around without thinking through what those words do or don’t mean).
- Line 15: Agreed, in some cases the meaning of “significantly” and “large” is vague or arbitrary and in those cases it has been removed.

L37… seems like some rewording is in order here, in that all 3 of the mechanisms that are identified here aren’t actually specific to SAI, so they’d lead to uncertainty in the response to solar reduction too. (And if there was no control over the interhemispheric gradient, then the uncertainty would lead to uncertainty in that gradient, rather than uncertainty in where to inject to compensate it – a point that should be made much more clearly somewhere in the paper.)
- Agreed that this is a useful point to make. We have included mention of the tradeoffs between model dependence in the SAI implementation and climate response as a function of target metrics chosen.
Note that a lot of the degree symbols didn’t wind up correct in the pdf.
- Sorry. This have been fixed.

L124-125, I agree with the plausibility statement, though strictly speaking this should come with some reference to support the assertion.
- Line 124-125: Reference added .

Section 3 might benefit from some subsections (on CO2 fast response, AMOC, and tropospheric aerosols)
We agree that this is a useful suggestions and these section headers have been added.

First paragraph of Section 3… Before this (probably somewhere in section 2) it would be critical to better describe the goals of the strategy used in GLENS and ARISE that is currently only briefly noted in passing lines 84-85, because otherwise a reader not intimately familiar with these datasets would not understand why there is a difference in the aerosol injection rates across latitudes. Really only takes an extra sentence to reiterate that the injection rates are adjusted to maintain not only global mean temperature but also interhemispheric and equator-to-pole, and that the algorithm determines the distribution of injection across the 4 latitudes that is needed to compensate all 3 metrics, with the balance between NH and SH injection based only on the desire to balance interhemispheric temperatures, and the balance between 15 and

30 degree injection based on the desire to balance equator to pole gradient. (One could further point out that as the radiative forcing from co2 is hemispherically symmetric, to first order one might expect a symmetric injection strategy to optimally compensate for the co2 forcing, even with a goal of managing interhemispheric balance.) This comment on the needed forcing seems essential context prior to the current first paragraph of section 3.
Revisions to address this issue have been made in response to earlier comments.

In that first paragraph, one could also point out that the very first simulation of the control strategy used in GLENS and ARISE, from Kravitz et al 2017, resulted in a nearly hemispherically symmetric injection profile. (The sole difference between that and GLENS being the switch from CLM4 to CLM4.5…) Or maybe this is worth explaining elsewhere… do you know why the change in the land model can also change things? Is that also a vegetation-based fast-response to CO2?
This is quite interesting and a bit of a mystery but our initial hypothesis would likely focus on the relative humidity of the advected flow from North America and its influence on salinity trends in the North Atlantic. That said these aspects are beyond what we can assess in the current work.

L170, if you're going to use "FSNT", then while that makes perfect sense to those of us who use CESM, might be better to state what it is an acronym for, for the rest of the folks.
- Line 170: Yes, this was also pointed out by Alan. Acronyms have been changed to be more intuitive.

L182-3, of critical relevance here, but not stated, is the role of the SAI strategy. If both simulations had been conducted without any intent to balance interhemispheric temperature gradient, then perhaps the correlations would be stronger – that is, it is the fact that the controller used is trying to deliberately compensate for model differences that matters; this will make the intermodel differences in temperature more similar while the intermodel differences in injection rates (and hence FSNT) will be less similar. (This is similar to my comment above in that you're glossing over the relevance of the injection strategy in interpreting the results, I don't think you can do that.)
Revisions to address this issue have been made in response to earlier comments.

L200-202, yes, but (worth pointing out somewhere) that to first order, hemispheric asymmetry in the slow response doesn't matter… that is, if there is hemispheric asymmetry in the *response* to a symmetric CO2 RF, one would expect to also have a counteracting hemispheric asymmetry in the response to a symmetric AOD… see next comment too.
Agreed, we've added text that the slow response is not important to the extent that changes in temperature are small and rather it is their inability to compensate the rapid adjustments that really is key.

L205-206 yes, but with a qualifier… that being that *some* of the uncertainty in how the climate responds to CO2 (climate feedbacks that determine the slow response) is actually reduced in the CO2+SAI case, because the same temperature-dependent climate feedbacks operate in response to both forcings (by definition; see e.g. MacMartin, Kravitz and Rasch, 2015 in GRL, which one of the reviewers claimed was too obvious to publish). g., since the purely radiative forcing from CO2 (prior to thinking about cloud adjustments and their resulting RF) is roughly understood (roughly uniform spatially and seasonally), then a hemispherically symmetric AOD would to first order compensate for that, regardless of how hemispherically asymmetric the climate system response to that forcing might be. So this seems a bit too simplistic – the critical observation from your analysis here is that the SAI needs to compensate for the radiative results from the fast-response to CO2 as well…

Agreed. We've added mention that it is specifically the rapid adjustments that are key. We think the importance of rapid CO2 responses (in both energy and water cycle realms!) remains underappreciated in the community and so the comment made by the reviewer of MacMartin et al. GRL is unfortunate to the extent that it limited this type of discussion.

L220, important to stress what you're comparing to. Of course, even in ARISE-SAI-1.5, AMOC is stronger than in SSP2-4.5, it's just weaker than the reference period… though looking at your figure, if I guess on the missing information not given there, this may be dependent on the metric one chooses to evaluate AMOC strength (at an absolute minimum that needs to be acknowledged). In discussing AMOC response and comparison between GLENS and ARISE, important to stress that in both cases the presence of SAI strengthens AMOC relative to no SAI, but in the GLENS case it is overcompensated (relative to change in global mean temperature) while in ARISE it is undercompensated (relative to that). I don't think this comes across well… comes across as a fundamental difference in sign which is simply not true – it's more a question of degree of compensation by SAI. (That is at least true for AMOC metric focused on strength alone; see e.g. Figure 3 in https://www.pnas.org/doi/10.1073/pnas.2202230119; that's the middle-atmosphere version rather than TSMLT, but the plots are nearly identical for ARISE). The metric considered here that risks confounding changes in pattern with changes in strength… if the conclusion that ARISE strengthens AMOC relative to SSP245 isn't true for the eof-based method you use here, once you include the relevant baseline case for comparison, then that calls into question how to interpret the eof-based metric.)

We have added SSP245 time series to Fig. 5, which was an inadvertent omission. It now shows that SAI, somewhat surprisingly, has a limited effect on ARISE versus SSP245, as opposed to in GLENS, where the effect is striking. Again this may point to the role of water cycle processes, rather than warming, in driving AMOC trends.

L261, the latitude of SAI injections only depends on it if one wants it to depend on it… of course, there are good reasons to want it to depend, but the current wording is too concise. (Again, one could simply fix the latitudes of injection, set the NH and SH injection rates to be the same, choose them to balance T0 only, and then instead of the uncertainty being in the injection rates, it would manifest as uncertainty in the resulting shift in T1 under SAI…)

Agreed. The discussion has been expanded to be explicit about this.

L264, wouldn't it be fairer to say two versions of the same climate model… the similarities between CESM1 and CESM2 are much more than between them and some other modeling center model.
We are hesitant to call these the same model given the substantial changes in clouds between the model versions. We now describe them as two models from the same climate modeling center.

L271, the sentence is about SAI but then switches to SCI. Everything in this paper would apply equally well to MCB, but this sentence as written shouldn't switch.
L271, the last bit of this sentence is wrong. SAI is *already* a "promising" risk-mitigation measure; "promising" generally suggests that you don't need to reduce all of the uncertainties
This sentence is now revised, also to address Alan's concerns.

Fig 2, units on panel f are correctly shown in the figure as "K" and wrong in the legend. (IMO you could get rid of the units on all the subpanels and just state in the caption) BUT, panel f should also be scaled by the amount of cooling offset, otherwise it will greatly overemphasize the residual in GLENS-SAI relative to residuals in ARISE. Ditto Figure 3f. (Actually, now I'm not sure how to interpret panel f at all… I was thinking that if you subtracted the respective reference time-period from each simulation, and then normalized by the amount of cooling, then that would tell you something about the pattern of the residual GHG+SAI in each case… but did you just take some time period and subtract the two, despite the different background emissions and different temperature targets? How is one supposed to interpret that?)
Good catch. Corrected. Panels f are not normalized by warming (as they are in the other panels) as temperatures are stabilized under SAI. It is true that they could be normalized by avoided warming but this doesn't differ much from what is shown. The interpretation of (f), and the point being made, is that the difference in the patterns of unmitigated warming (c) is actually quite similar to the residual pattern under SAI (f) and even the low top model versions (d).

Fig 5a, should also include the line for CESM2-WACCM-SSP2-4.5, for context, either in addition to or instead of CESM2-WACCM6-SSP585. Having said that, this figure confuses me in multiple ways. I presume the units are Sv? Is the eof calculated once, or is it a different calculation for each model? Each simulation? Each decade? If you keep the same eof pattern, how do you disentangle a change caused by a shift in the strength of the circulation from a shift in the latitude of the peak? I know, for example, because I've plotted it, that if you use the maximum of the streamfunction as your AMOC metric, then you will clearly show that ARISE-SAI *recovers* AMOC strength relative to SSP2-4.5 (but not back to 2030 levels), so if that does not hold for your choice of metric, then that point is pretty relevant to point out to the reader – that the conclusions on AMOC depend on what metric you happen to use to calculate it. (I'm guessing that's true as I'd expect SSP2-4.5 to not be worse than SSP5-8.5.)
- Yes, thanks - the figure benefits from this addition though the change is actually quite close to that of SSP585. The PCs are unitless and the units are retained in the EOFs. This does result in challenges in anything more than a qualitative comparison. Discussion has been added to the text and a new figure on the

patterns are in the appendix.

Fig 5 panels b-e, change relative to what? Relative to their respective reference periods, or relative to unmitigated at the same time period?
- Fig 5 b-e: Good catch! These are changes between 2020-2039 and 2050-2069 and this was omitted from the caption. Now added.

Panel 7f isn't particularly meaningful, given that it isn't scaled in any way… what's the message?
- Panel 7f: The message here is that, consistent with Fig 1, the injections at 15S are much greater in ARISE-SAI than GLENS-SAI, which are instead in the NH. It provides useful context for the hemispheric asymmetry also seen between SSP370 and RCP85 in Fig 7e.
***