Reviewer comments are in bold. Authors' responses are in blue.
Response to reviewer #1 (Peter Irvine)

This article provides a thorough analysis of the similarities and differences between the responses of the GeoMIP G6 sulfur and G6 solar experiments, and a discussion of the uncertainties involved. The article is very strong and would make a valuable contribution to the literature, and I think it is ready for publication after a few minor points are addressed.

We thank the reviewer for their very nice comments! We address all of his points below.

A very minor point but I would suggest reviewing the use of parentheses as there are an awful lot of them. I would suggest reserving parentheses for those points which are truly not relevant to the thrust of the sentence, and restoring some of the details held in many back into the sentences.

Following this and other comments by all the reviewers, we have tried to simplify some of the longer phrases.

The figures look great but one common issue is that they don't make use of greater than and less than arrows (triangles) on the colorbars. If all data falls within the plotted bounds, perhaps this could be stated or if not these colorbars should be changed.

We changed all the colorbar where the issue was correctly raised by the reviewer.

One area I think that the article could elaborate on is the differences in the response to prescribed stratospheric AOD as compared to fully simulated stratospheric aerosols, covering stratospheric heating, chemistry changes, etc.

We have included in Section 2 a discussion of the differences in response for the models with prescribed aerosols.

Specific comments – Note many of these are suggestions to clarify the text and should be taken on or disregarded as the authors' see fit.

L13 – should these sentences be linked by a colon? They seem distinct points to me.

We've separated the two phrases.

L15 – aerosol's?

We've changed the phrase to make sure it's clear.

L19 – should that be a full stop rather than a colon?

Done

L21 – Is there something that could be said for precipitation change?

We mentioned precipitation changes above, giving some numbers: given the complexity of the response, we don't feel like we could say anymore without going into too much detail.

L38 – missing space, 4x.

Fixed.

L60 – by the stratospheric circulation

Fixed.

L62 – missing close bracker.

Added.

L70 – perhaps add a short phrase linking this list of analysis to the goal of exploring these uncertainties.

Added.

L71 – experimental

Fixed.

L80 – distributions

Corrected.

L81 – put parentheses and cites after "stratospheric processes"

Done.

L93 – I think here or elsewhere it is worth reflecting on the view that RCP8.5 is not just a high emissions scenario but an implausibly high emissions scenario, or at least adding a few words of caution around this scenario. This might be raised in the conclusion or introduction instead.

We had included some comments about this in the conclusions, but we have included some more information on that as suggested.

L94 – I'd suggest parentheses have been overused in this document, here for example.

We've removed them here and elsewhere.

L95 – drop spatially, put the parenthetical statement between commas.

Done.

L105 – Perhaps flip the order of this sentence to make it easier to follow: "The teams updated the reduction in solar constant, and the prescribed aerosols … at different intervals, two did so every decade, …."

We prefer to keep the order of the phrase like this, since then we specify what CESM2 did in the next phrase and it makes more sense.

L112 – to within 0.2 C of SSP2-4.5 levels.

Added

L114 – there are a variety?

Changed.

L115 – produce a large spread for the two scenarios.

Changed.

Figure 1 – there's a lonely degree symbol, should it just be a K? I'd guess that the ensemble members have been plotted for each model but this isn't stated.

It should be °C (it's not anomalies). We added a note about the single realizations, thank you for noticing.

L126 – What about radiative forcing? Surely the response to CO2 forcing and solar forcing are related, even if the % change in insolation and the response per doubling of CO2 are not.

In a transient climate, the relation might not be obvious (especially considering that there are other evolving forcings, like tropospheric aerosols and other shot-lived GHGs). Some more in-depth analyses could be done with the G1 experiment (described in Kravitz et al., 2021), but it falls outside the scope of this work.

L130 – Figure 2f? And does it show this? Where is the model spread without normalization? Is this a generalizable result or a chance occurrence due to the make-up of ensemble?

We fixed the figure label. We have included in the revised version of the supplementary the non-normalized version of Fig. 2f. Due to the modest amount of models, we're not sure we can say if this is a generalizable result.

Figure 2d – R-squared is 0.0, is that right?

Yes.

L140 – Should that be the other way round, incoming minus outgoing, i.e. positive = more energy input to the earth system?

Thank you for catching that, fixed.

L142 – a much more consistent

Thank you. Fixed.

L144 – two forcings

Corrected.

L145 – rephrase? And drop "indeed"

Changed.

L146 – is necessary

Added

L146 – drop indeed

Done

L148 – more in general? More generally?

Changed

L161 – models'

Fixed.

L171 – new paragraph?

Done

L178 – can't be made.

Corrected.

L183 – Are there papers comparing the simulated Pinatubo response and observed Pinatubo response for these models? That would be a valuable point of comparison.

Sadly, there aren't, although some models will take part in other experiments (like VolMIP and ISA-MIP). These experiment aim to standardize the initial meteorological conditions before the eruption (QBO, winds, ENSO etc.) so that the differences in transport and microphysics can be analyzed. One of these experiments is being analyzed right now, for Pinatubo (https://meetingorganizer.copernicus.org/EGU21/EGU21-13387.html) but results are not final yet (and some models are still missing)

Figure 4 – the Pinatubo box is clunky and hard to read, is there a better way to present this information? Here and in other relevant figures, it's probably worth mentioning that the yellow and orange lines fully overlap. Panel c – the panel title is too long, how about: "G6 sulfur AOD compared with Pinatubo". In the caption: "the year with a global AOD CLOSEST to that OF Pinatubo"

Done! Thanks for the suggestions.

L194 – this sentence is hard to follow, consider revising.

We rewrote this sentence.

L200 – this last sentence is also a little muddled, consider revising.

Rewrote.

L21 – I think this is reasonable, isn't there a citation to back that up?

We're not sure what the reviewer is referring to, as the line is probably wrong?

L214 – this comparison is not fair, the forcing from Pinatubo has had only a year to act, whereas for G6sulfur it has had a few decades. This suggests that the simulated response may be weaker than the observed Pinatubo response.

Given the feedbacks from other reviewers, we have tried to explain further some of the limitations of comparing with Pinatubo.

Figure S1 – are there negative AOD values?

Considering they are derived by considering the AOD in G6 minus that from the background ssp585 experiment, it is possible that some of the initial years produce a local negative result.

Figure 5 – Is SSP2-4.5 a typo? Should that be SSP5-8.5?

Yes! Thank you for noticing.

246 – difference rather than differences. And should that be radiative perturbation from CO2.

Fixed.

L250 – stratospheric heating repeated?

Yes. We removed the repetition.

L251 – is this inability to restore the ocean state of the same character as the other things listed? It seems to be a consequence rather than a driver of differences in response. Though perhaps I've misunderstood the point.

It would depend on what's driving the particular feature of the oceanic circulation (of the AMOC, in particular). And even more important, what is the driver in the models (which could be different) compared to the real world.

L255 – Is Macmartin the right reference to introduce this analysis? Won't the main point of Macmartin still hold? I.e., that the uncertain degree of warming under RCP8.5, that is absent or reduced in solar geo scenarios, is itself associated with regional climate changes and hence drives model spread (and uncertainty). The Macmartin argument applies to the spread in RCP8.5 projections as compared to those of G6.

In a way it still holds, as we show in Fig. 6, but much less for G6Sulfur than for G6Solar. And, MacMartin et al. (2015) analyses used G1, which is solar dimming over 4xCO2, so it was a much simpler experiment in many ways.

L260 – I'd suggest not switching the order of comparison here, i.e., the spread is larger in G6 sulfur.

Fixed.

L261 – I think it is a mischaracterization to say that G6sulfur is unable to cool down the Norther high latitudes. It looks to me as if ~2/3 of the warming difference has been offset. This should be rephrased to make clear that it is at least partially effective.

We added the word "completely" to better clarify the point.

L263 – Is this definitely due to heating rather than some other factor?

It is one of the possible causes (and in CESM we have verified in other works that this is the case). But we've toned down the phrase.

L265 – although, if I recall correctly it's of smaller magnitude than would be expected from just the difference in radiative forcing due to countervailing circulation changes, right?

Yes. We have fixed the phrase.

Figure 6 – Why do the colobars not have "greater than" extensions? The lack of such extensions suggests that all points in the arctic in 6c, for example, see less than 1.1C of warming. Is that correct? Should note the change in colorbar range between a and c and e.

See response above. We changed all the colorbars.

Figure 7 – no colorbar label.

Added.

L305 – point to?

Corrected

L333 – Should this be split into 2 sentences?

Done!

L345 – tends.

Fixed.

L346 – will occur over the tropics, rather than observed.

Changed.

L358 – this last sentence is very long, consider splitting. Should that be "direct effects of CO2", as in the CO2 physiological effect or is this also CO2 radiative effects too? Changes in dynamical changes?

We have changed the phrase to make it clearer.

L360 – Could this effect be quantified, i.e., what's the land-mean precipitation error for those two projections?

We have included some discussion and a supplementary figure of the land-mean precipitation.

Table 1 – CMIP6 doi? Is that a typo for MPI models, i.e., should it read: "r1,r2,r3" and not "r1,r2,r"?

The DOI (sorry, we changed that to DOI from doi) is the one for the various datasets uploaded to CMIP6. We fixed the problem.

Table 2 – should "last decade (2081-2100)" be "last 2 decades (2081-2100)"?

Fixed.

Figure 10 – The side-panels don't line up with the maps and it's hard to read given the number of colourful lines, I'd suggest pulling them out to a separate figure as little is gained from combining them with the maps. I also would only show one experiment for absolute plots.

We have separated the two figures.

L370 – "reduction in global-mean precipitation" should be elaborated to make clear what that reduction is relative to, as there is also a reduction seen in SSP2-4.5 relative to SSP5-8.5.

Done!

L372 – This sentence has a strange structure, consider revising.

We've tried to clarify the phrase.

L385 – no need to specify air, although it could be specified in air temperature.

Removed.

L386 – this second half could be rephrased.

Done!

L415 – (and more) at wrong point.

Changed.

L422 – rephrase - not even the most ideal.

Done

L425 – restoring or maintaining the ITCZ location?

Changed.

Figure 11 – stratospheric aerosol intervention – first time used. "more particularly". Just a suggestion, but this might be easier to follow if the boxes were ordered from top to bottom: stratospheric heating and aerosols top (the inputs) à stratospheric chem, dynamics and aerosol chem / microphysics à Radiative forcing + surface climate (the outputs).

Thanks for the suggestions. We have made the changes in the revised figure.