Dear referee,

Thank you very much for the valuable comments on our manuscript. We have considered all the comments carefully, and will revise the manuscript according to the comments and suggestions by you and other referees, as well as community reviewers. Below is a point-by-point response your comments.

With warm regards,

Ou Wang, Ju Liang, Yuchen Gu, Jim M. Haywood*, Ying Chen, Chenwei Fang, Qin’geng Wang*

**General comment:**
This paper examines the impact of changes in different measures of precipitation over China under a high emissions scenario and with a small ensemble of climate intervention simulations at the end of the century. One climate model (UKESM1) is used for comparison with two different realisations of stratospheric aerosol injection - G6solar, using a constant solar dimming, and G6sulfur, with gradually increasing injections of sulfur into the stratosphere. The paper is well organised and generally clearly and well written. My main criticism is that the discussion does not evaluate the results with respect to other research that was carried out from other experiments such as GLENS. Suggestions for other relevant literature is at the end together with references made in the comments below.

**Major comments:**

Consider reducing the content of Section 2.2. A lot of this is repetition from the literature that is cited and doesn’t necessarily need to be included in this article.

**Response:**

Thank you very much for your suggestion. We have realized many repetitions from the literature are not necessary. The section 2.2 will be revised as:

“In this study, data from G6sulfur and G6solar experiments are used from the sixth phase of GeoMIP from the U.K. Earth System Model UKESM1 (Sellar et al., 2019). UKESM1 is a fully coupled Earth system model with an atmospheric resolution of
1.25° latitude by 1.875° longitude (Storkey et al., 2018; Walters et al., 2019; Mulcahy et al., 2018, Sellar et al., 2019), and contributes to both CMIP6 and GeoMIP6 (Jones et al., 2020). The Scenario MIP high GHG forcing scenario SSP5-8.5 (O’neill et al., 2016) is used as the baseline scenario of both G6solar and G6sulfur experiments (Kravitz et al., 2015). UKESM1 simulates SO₂ injection in the stratosphere along the Greenwich meridian at an altitude of 18-20 km between 10° N and 10° S in G6sulfur over the period 2020-2100 (Kravitz et al., 2021; Haywood et al., 2022). A parallel experiment to G6sulfur, the G6solar experiment, reduces ScenarioMIP Tier 1 high forcing scenario to the medium forcing scenario by reducing solar irradiance. Notably, it is anticipated that G6solar will exhibit reduced inter-model disparities in the spatial distribution of forcing when compared to G6sulfur owing to model differences in representing the complexities of the sulfur cycle within global models. Therefore, G6solar is proposed as a parallel experiment to G6sulfur for the purpose of comparing the impacts of solar reduction with those of stratospheric aerosols (Kravitz et al., 2015).”

To avoid confusion, consider changing any reference to “present-day” to the control period and sticking with this consistently, rather than switching between ‘historical’, ‘baseline’ and ‘present-day’. The WMO has adopted 1990-2020 as the “current climate period”, which suggests that the period used in this article is historical.

Response:
Thank you. We have changed all references to "present-day" to "control period" throughout the manuscript.

Is there a benefit in the data validation against APHRODITE? The article does not include an assessment of extreme precipitation metrics with respect to observations, nor do the results or conclusions refer back to the observations. I think you could drop this, and instead refer to other assessments of the validity of extreme precipitation in climate models such as (Sillmann et al., 2013; Donat et al., 2020; Tebaldi et al., 2021)

Response:
Thank you for this important comment. Indeed, this study has benefits in the data validation against APHRODITE. First of all, the observations were used to validate the direct results of the model (i.e., amount of precipitation). We think this is the most fundamental for our study. In this regard, further description and analysis will be added in the revised manuscript. In addition, a scatter plot will be added as a new panel in
Figure 2 (as shown below), and the bias as a percent instead of an absolute value between the observations and the models (ensemble mean and the three model members) will be indicated, as suggested by another referee (RC3). Please refer to the response to the RC3 for more detail. More comparison on the extreme precipitation metrics between the observations and the model results will be performed and discussed in the revised manuscript.

![Figure 2(d) Scatter plots between the observations and model results at different level of precipitation during the control period (CP).](image)

The observations were classified into several level (intervals): P10 (the smallest 10%), P10-50, P50-90, P90-95, and P95 (the largest 5%).

In addition, according to your suggestion, some assessments in other relevant studies (e.g., Sillmann et al., 2013; Donat et al., 2020; Tebaldi et al., 2021) on the validity of extreme precipitation in climate models will also be mentioned in our revised manuscript. We appreciate it very much for providing the valuable references.

Given that this research uses one climate model, and three ensemble members for each scenario, it isn’t really appropriate to state that G6sulfur/G6solar abates or ameliorates...
climate change as depicted by SSP5-8.5. There aren’t sufficient model members to remove model uncertainty, and without observations we do not know which model realisation (if any) adequately simulates the effects of SAI. All you can state with confidence is that using this methodology and data, the SAI experiments produce results consistent with a lower emissions target. I would prefer to see all of the statements on improvements with respect to climate change removed, or at least reduce the emphasis of the statements.

Response:

We agree with you, and thank you very much for the suggestion that makes our paper more rigorous and scientific. Really, considering many possible uncertainties in the climate models, as well as in the research scenarios, we cannot assert that G6sulfur/G6solar abates or ameliorates climate change as depicted by SSP5-8.5. Accordingly, relevant statements in the manuscript will be revised to reflect the limitations of our methodology and the uncertainties associated with the findings.

Minor comments:

L28 change impacts to efficacy

Response:

L28: "impacts" has been changed to "efficacy".

L30 rephrase this sentence as noted above.

Response:

L30: The sentence “While the results from both G6sulfur and G6solar show encouraging abatement of many of the impacts on detrimental extreme events that are evident in SSP5-8.5 there are some exceptions.” has been change to “In all, the results from both G6sulfur and G6solar are encouraging, showing a reduction in the efficacy of detrimental extreme events, consistent with the lower emissions target of SSP2-4.5.”

L31 remove trends

Response:

L31: "trends" has been removed as suggested.
L42 is the higher risk of flooding associated with increased extreme precipitation? Remove the time periods studied, this is implicit.

Response:

Yes, the higher risk of flooding is indeed associated with increased extreme precipitation. This is indicated in the results by Ying et al. (2014), where the flood risk is understood as an extreme climate index. Indeed, the time periods are implicit, and will be removed.

L44-60 is this level of detail on historic events warranted? You do not examine the changes in jets or other sources of extremes.

Response:

The details on historic events are mostly quoted from published papers or media reports. To be honest, we could not warrant their reliability since we have not conducted further investigation in this regard. Now, according to the comments by another referee (RC2), we have come to realize that the detailed description on the historic events is not necessary as it is not directly relevant to the study here. Therefore, we the part (L44-60) will be deleted or shortened in our revised manuscript.

L44 Why was summer of 2020 anomalous - it seems in keeping with the other extreme events you reported.

Response:

The using of the word “anomalous” was not appropriate. Besides 2020, flooding events also frequently happened in other years. We will correct it later. In addition, as mentioned above, the detailed description on the historic events will be shortened in our revised manuscript.

L61 change appears to has, and cite relevant literature such as (Donat et al., 2016; Pendergrass and Knutti, 2018), which also discuss changes in the hydrological cycle.

Response:

The sentence has been changed as: “On a global scale, climate change has been influencing hydroclimatic conditions (Donat et al., 2016; Pendergrass and Knutti, 2018).”

L64 Should causes go before faster?
Response:

Sorry for the wording mistake. The sentence should be “Climate change causes faster evaporation, and higher atmospheric temperature induce more moisture-laden air in the storm tracks.”

L66 Update this to the more nuanced and recent research that shows extreme precipitation generally goes up everywhere (Pendergrass and Knutti, 2018)

Response:

We will update that based on relevant studies (e.g., Pendergrass and Knutti, 2018).

L79 SAI does not mitigate anthropogenic climate warming, it may mitigate some of the impacts.

Response:

You are right. The sentence will be revised as: “To some extent, SAI partially counteract climate warming by injecting reflective particles, or their gaseous precursors, into the stratosphere (Zarnetske et al., 2021).”

L79-89 This paragraph needs rephrasing to explain that SAI is premised on reproducing the effects associated with volcanic eruptions. However, you do not need to list the different volcanic eruptions themselves - just point to a large body of literature that supports these effects.

Response:

Thank you for this comment which has helped make our manuscript more concise and less verbose. The paragraph (L79-89) will be changed as: “SAI is premised on replicating the effects associated with volcanic eruptions, wherein reflective particles or their gaseous precursors are injected into the stratosphere. These resultant aerosols reflect and scatter solar radiation back into space, leading to a cooling effect that counterbalances the warming caused by increased concentrations of greenhouse gases (e.g. Bluth et al., 1992; Self et al., 1996; Robock, 2000; Soden et al., 2002; Haywood et al., 2014; Schmidt et al., 2018). In addition to reducing the temperature, SAI also influences tropospheric and stratospheric ozone, terrestrial ecosystem, terrestrial carbon, and hydrological cycle by changing the physical climate system and atmospheric chemistry. Numerous studies support these effects associated with volcanic eruptions and their simulation through SAI techniques (e.g. Liang and Haywood., 2023; Jones et
al., 2018; Jones et al., 2020; Cao, 2018; Plazzotta et al., 2019; Lee et al., 2021; Visioni et al., 2022; Imai et al., 2020; Mclandress et al., 2011).”

L99 include other recent research that explored changes in temperature and precipitation in other SAI experiments not just the GeoMIP archive (e.g. Tye et al., 2022; Simpson et al., 2019).
L100 What about (Tew et al., 2023)?
Response: 
Thank you for the suggestion. We have incorporated other recent research, including studies exploring changes in temperature and precipitation in various SAI experiments beyond the GeoMIP archive, such as those by Tye et al. (2022), Simpson et al. (2019), and Tew et al. (2023).

L105 remove maximise the signal-to-noise in the simulations as 
L114 remove according
L126 See comment above, but at the very least remove The GeopMIP G6sulfur simulations that reduce global mean temperatures from the SSP5-8.5 scenario to the SSP2-4.5 are described in detail elsewhere.
L143-145 remove this last sentence.
Response:
Your above suggestions are all accepted in revising the manuscript.

L154 I believe that the extreme indices were defined by the WCRP not IPCC.
Table 1: the authors should be Frich and Klein Tank. Also refer to (Sillmann et al., 2013; Zhang et al., 2011) for the correct definitions
Response:
Thank you for pointing out the mistake, and we have corrected it in the text.

How did you calculate the 95th percentile? Did you bootstrap the individual years to avoid data inhomogeneities (Zhang et al., 2005)
Response:
For each grid, the 95th percentile was calculated based on 30 years (1981-2010) of precipitation data. We calculated the 95th percentile directly without using
bootstrapping methods, as recommended for calculating temperature indices (https://www.climdex.org/learn/indices/#index-TX90p).

L161 This may not be relevant if you remove the Aphrodite data as suggested above. However, I am concerned about regridding the larger data to the smaller grid. No additional information is gained in this respect (just several grid boxes with the same values) and may show errors and biases that are not true. Instead it would be more robust to regrid the observations to match that of UKESM. See https://climatedataguide.ucar.edu/climate-tools/regridding-overview for more information.

Response:

Thank you for your valuable suggestion. The observations were regridded to match that of UKESM.

L165 remove instead of the more commonly used Student’s t-test. Wilcoxon Rank Sum Testse work as

Response:

The text was removed.

L167 change “with p-value <0.05 suggesting” to “with a 5% confidence level of”

Response:

The expression was changed.

L169-182 How did you establish the CDFs? Did you fit distributions, or are these empirical CDFs from the data? Did you examine the uncertainty in the CDFs, and were they fitted for each model member (correct) or the model mean (as the figures suggest)? I am also wary about CDFs for very small sample sizes - i.e. 30 values of the annual maximum rainfall.

Response:

Thank you for pointing out this problem that we didn't explain it clearly about how we established the CDFs. In fact, for establishing the CDFs, firstly, for an extreme precipitation index at each grid point, the yearly mean of the ensemble model members was calculated. Then, the annual extreme precipitation indices for each grid point was
obtained by averaging over yearly means during the 30 years. Finally, the cumulative probability distribution of the extreme precipitation index over all grid points was statistically analysed for each of the seven regions, as well as the whole China. Therefore, we have a large number of samples for calculating the CDFs, stead of 30 values.

We computed the empirical cumulative distribution functions (ECDFs) of our data using histograms. To achieve a smoother representation of the distribution, we applied a Gaussian smoothing technique. By doing so, we were able to obtain smoothed representations of the empirical distributions, which provided clearer insights into the underlying patterns of the data.

As far as the uncertainty in the CDFs is concerned, in the original manuscript, we just placed emphasis on the uncertainty in the direct results of the UKESM, which we think is the most fundamental for our study. The model results (amount of precipitation) were validated with the observations (APHRODITE), and only means of the ensemble models were considered. In our revised manuscript, examination on the uncertainty in CDFs will be added by comparing the model results with that from the APHRODITE (for the historical period) and comparing the results among different models (for the future).

**L185-194 put this into the previous section.**

**Response:**

Thank you for your suggestion. We'd like to clarify that since we have already employed the Wilcoxon test for significance testing, the field significance calculations mentioned here are redundant, and was not used in this study. Therefore, we have removed this portion (L185-194) in the revised manuscript. We are sorry for this confusion.

**L193 Include this statement in the figure caption instead of the text - and check which way you have represented significance, this is opposite from the figure.**

**Response:**

As said above, the description of the method is removed.

**L211 Comment on the increase in drought in west and Taiwan under G6sulfur.**

**Response:**
At present, we are not sure about the mechanism about the increase in drought in west and Taiwan under G6sulfur. We think this is possibly because of the effects of relatively complex terrain in the areas, in particular, high mountains, which may disturb the atmospheric circulation and block transport of the moisture air. We will add some explanations in the revised manuscript.

L215 Change this statement to something like projected changes are similar to those of SSP2-45, meaning that the SAI simulations are approximately successful.

Response:

The statement has been revised as “G6sulfur (Fig. 3c) shows projected changes similar to those of SSP2-4.5 (Fig. 3b), indicating that the SAI simulations are approximately successful.”

L246 There are no absolute values&le;100mm or no increases from the control period $>100mm$?

Response:

Thank you for pointing out this problem. We have clarified as “In the other three G6 models, the increase in RX5day is considerably smaller than that under SSP5-8.5, with none exceeding 100 mm compared to the control period (Fig.S2b-d).”

L250 remove (p-value $<0.05$) and every other instance - this has already been stated.

Removed.

L251 remove which is generally. Is there really only one research paper on increases in extreme precipitation in this region?

The words are removed. Some other research papers (e.g., Qin and Xie, 2016; Peng et al., 2018) are include.

L257 Remove this sentence.

Removed.

L265 should this be depicting?

Yes, corrected. L266 remove sentence “The comparisons confirm….”
L278 remove “This suggests that G6sulfur….” This is the results section, so discussions aren’t appropriate.

   Removed. We appreciate it very much for pointing out the grammar and wording problems.

L280 Stick to reporting the differences between the simulations in this section, then interpret (with appropriate caveats) in the discussion sections and remove the sentences on the implications or efficacy of SSPs vs G6.

Response:
Thank you for the thoughtful suggestion. We have removed the interpret in line280-4, and now the revised sentence is “In comparison to SSP2-4.5 (Fig.S3b), G6sulfur exhibits an increase in RX5day, primarily in the region between 100°E and 120°E. For 'G6sulfur-G6solar'(Fig.S3c), positive values of RX5day are more pronounced in certain areas between 100°E and 120°E, especially in the low latitude zone between 20°N and 30°N.”. In revising the manuscript, we will stick to reporting the differences between the simulations, and interpret (with appropriate caveats) in the discussion sections.

L291 Remove ameliorating

Response:
L291: "Ameliorating” has been removed.

Table 2: I am not sure that this adds to the interpretation of the results and could be removed. If retained, re-phrase as difference between G6 and SSP or something similar.

Response:
We think that Table 2 provides a useful summary of the results. If you look at Table 2, you get the general idea that SRM does ‘good’ things for the precipitation extremes, but ‘bad’ things for the droughts and dry days. That is a very useful take-home message. So, we would like to retain the table. Some descriptions and discussions about the results will be reconsidered and revised.

L311 Why is this interesting? Elaborate please.

Response:
We find this interesting because it highlights a unique pattern in the data. Despite observing mitigation effects in other regions, we notice that G6solar does not exhibit the same effect in the SC region. In other words, unlike other regions, G6solar does not mitigate the extreme high values of the RX1day index in the SC region. This suggests a nuanced relationship between the G6solar and their impact on RX1day in the SC region, warranting further investigation.

L325 It might be more meaningful to look at the relative changes (e.g. in percentage terms) rather than absolute values. With regard to the “arbitrary” regions, why are they somewhat arbitrary? Surely they relate to some geographical or political definition, the point to make is that they may not correspond with climatological regions. Note that smaller regions would just emphasise noise in the results.

Response:

We agree that relative changes such as in percentage terms would be more meaningful in some cases. However, in the case of our study, a great deal of the results are small values, and consequently, the relative changes (percentages) could be very large even for minor absolute changes. For this reason, we think the absolute changes can be more appropriate here. In the revised manuscript, we will add relative changes for some results or discussions where they are appropriate.

We agree that the conclusions would be different if based on different criteria for dividing the regions. Unfortunately, in this regard, there is no standard criteria for dividing the regions. In this study, the division of regions is a conventional way and has been widely adopted in many statistics reports and relevant studies (e.g., Luo et al., 2017; Fan et al., 2020; Yang and Shao, 2021; Liang et al. 2023). The sentence “It should be stressed here that the regions that are chosen for aggregation are somewhat arbitrary and the results could well change should smaller sub-regions be chosen for analysis.” is not expressed accurately, which is removed. Some discussion on the uncertainty from regions’ division will be added, including the scaling effect (smaller regions would just emphasise noise in the results).

L329/30 remove this sentence.

Response:

L329/30: The sentence has been removed.
Figure 6 as noted above the uncertainty across model members should be included in these curves. Please also check the colour scheme for colour blind appropriateness, and use the same x-axis for each variable for all regions (i.e. one x-axis for Rx1day, another for R95p). This also applies to Figure 9 and 12.

Response:

Yes, it would be more informative if the uncertainty across model members be included in the CDF curves in Figure 6. We do have the results of the three model members, and we tried to added results of each model along with the curves to indicate the uncertainty. However, since the curves are closely overlapped, we couldn’t find a way to make the figure clear. Actually, the lines or colour blocks could be blended and overlapped together, and make the figure difficult to distinguish. For this reason, we gave up the idea of directly including the results of model members in the figures. Instead, we will add some statistical metrics of the model members in our revised manuscript. In this way, we think, at least to some extent, the uncertainty across model members can be indicated.

The colour scheme has been checked for colour blind appropriateness. For example, the purple line has been changed to black.

For the suggestion on using the same x-axis for each variable for all regions, there is also a difficulty. Because the range of index changes varies big across different regions, when plotting them on a large-scale x-axis, the curves with small range (or values) could be very close or even overlapped with each other, and difficult to be distinguished. For this reason, x-axis is not the same for all the regions. This is also the case for Figure 9 and 12.

L350 Decreases in CWD do not necessarily equate to reductions in precipitation intensity. You can only make this interpretation if there is a reduction in the total number of wet days AND an increase or no change in the annual total.

Response:

Thank you for this thoughtful comment. After recalculating the total number of wet days, the sentence in lines 349-350 has been revised. “While CWD is projected to decrease under SSP5-8.5, annual total precipitation amounts are projected to increase (Fig. 3). Additionally, the total number of wet days is expected to decrease in the future. These findings suggest that daily extreme precipitation intensity may rise in southern areas of China in the future (Zhu et al., 2018).”
L356 R50mm is not one of the formal ETCCDI indices, it is user defined.

Response:
Thank you for your reminding. We have revised the sentence as follows:
“The R50mm index is derived from the Rnnmm index, as suggested by ETCCDI. The Rnnmm index represents the count of precipitation above a user-chosen threshold. In this case, the threshold is set to 50 mm, as recommended by the China Meteorological Administration (CMA).”

L371 remove effectively ameliorates the

Response:
“effectively ameliorates the” has been removed.

L373-375 This is discursive and needs more references to support it (and moving to the discussion section). There is likely a combination at play including changes in the location of the jet streams and ITCZ, as well as interactions with topography and changes in maritime temperature gradients.

Response:
Thank you for your suggestion. More references in this regard will be included in our revised manuscript, and move it to discussion section.

L380/1 Remove this sentence.

Response:
L380/1: The sentence has been removed.

L387 CWD=200 days is right at the far end of the tail, I don’t think it’s appropriate to make this statement without any error estimates or uncertainty information. Further, the duration estimates of CWD and CDD add up to longer than a year - this is particularly obvious in comparison with Figure 12. Please check.

Response:
Thank you for the comments. As mentioned above, for the uncertainty in the CDFs, examination on the uncertainty in CDFs will be performed and added by comparing the model results with that from the APHRODITE (for the historical period) and comparing...
the results among different models (for the future). Basing on that, relevant statements will be checked and the manuscript will be revised accordingly.

The phenomenon of the combined CDD and CWD exceeding 365 days in the same region arises because the high values of CWD and CDD may occur at different grid points, resulting in the possibility of the total exceeding 365 days.

L402 Should this refer to Figure 9? This statement would be more meaningful with uncertainty envelopes to clearly demonstrate whether there are or are not differences between each model.

Response:

Yes, the statement refers to Figure 9. Also, as mentioned previously, since the curves are closely overlapped, we couldn’t find an appropriate way to include the results of model members in the figures. Instead, we will add some statistical metrics of the model members in our revised manuscript. In this way, we think, at least to some extent, the uncertainty across model members can be indicated.

L413 Remove, and would have the ….

Removed.

L458 It is not noteworthy that these yield similar results - that’s the objective of the cooling. Remove this sentence.

The sentence has been removed.

L468 Remove last sentence of paragraph.

Response:

L468: The last sentence of the paragraph has been removed.

L475-481 There are other aspects related to drought risk - not least evaporation - that don’t show up in the dry day count, such a strong statement about changes in drought risk aren’t appropriate. See (Cheng et al., 2019; Dagon and Schrag, 2019, 2016) for more results related to climate intervention.

Response:

The sentence line 476-481 has changed to “This reflects a potential decrease in drought risk in northwest regions and an increase in extreme drought events in low-latitude southeast coastal areas in the future according to four G6 simulations. Changes in precipitation affect soil moisture, thereby influencing evapotranspiration (ET) and
ultimately precipitation patterns. Assessing whether changes in DD and CDD affect drought risk also requires consideration of variations in ET and soil moisture (Cheng et al., 2019; Dagon and Schrag, 2016). Furthermore, solar radiation management (SRM) increases drought risk compared to SSP5-8.5 and SSP2-4.5 scenarios in northern regions (NEC, NC, and NWC).”

G6solar and G6 Sulfur are different ways of simulating a possible climate, we have no way of knowing whether one or the other is more valid without observations, and so can’t be described as outperforming each other. See (Bednarz et al., 2022; Visioni et al., 2021) for more discussion on this.

Response:
We agree that we cannot tell which one is better between G6solar and G6sulfur. What we did in our study is comparing both the results of G6sulfur and G6solar in extreme precipitation events against the lower emission target (SSP2-4.5) or that in control period. We will add more discussion on this according to the references you suggested.

Section 4 I suggest condensing the bullet points to short concluding statements, which can then be followed by the explanations. Given that the focus of this article is on the climate intervention, it would also make more sense to emphasise those results rather than the future projections that have been published elsewhere. This is also the point to discuss how valid the results are with respect to other research - including experiments outside the GeoMIP project.

Response:
Thank you very much for the thoughtful and valuable suggestions. In revising the manuscript, we will consider to condense the bullet points, and add more explanations. In addition, more comparisons will be included and discussed with relevant results from published studies elsewhere. Thanks to so many thoughtful and valuable suggestions by you and other referees, we believe our manuscript will be improved greatly.

References


