Response to Review of "Global streamflow and flood response to stratospheric aerosol geoengineering" by Wei et al.

We first thank the referee for his/her insightful comments, which helped us clarify and greatly improve the paper. In the reply, the referee's comments are in *italics*, our response is in normal and changes to the text are shown in blue.

Anonymous Referee #1

The authors analyze the results of GeoMIP G4 simulations on future streamflow and flood risk. Terrestrial hydrology has been looked at in detail in only a few solar geoengineering papers; this study is novel in that it utilizes a river routing model to connect climate model output with streamflow and flood risk. The study could use some additional analysis to back up claims made in the discussion section, and there are some issues with wording and language. I recommend publication after these and other comments are addressed.

General Comments:
1. *I would encourage the authors to add connections to atmospheric chemistry and physics implied by their results, perhaps through precipitation and evaporation feedbacks.*

   Reply: Thanks for your constructive suggestion. We analyze the precipitation, evaporation and related runoff changes under G4 stratospheric aerosol injection geoengineering. The following new Figure 1 is about the precipitation, evaporation and runoff changes from G4 to rcp45. The previous Figure 1 in main text is labeled as Figure 2. The paper fits well under the scope of ACP now, as we add the following paragraphs in a new "Results" section as "3.1 Precipitation, evaporation and runoff changes":

   G4 stratospheric aerosol geoengineering lowers net radiation fluxes at the top of the atmosphere (TOA) by ~ 0.36 W m^{-2}, and reduces mean global temperature by ~ 0.5 K and overall slows down of the global hydrological cycle (Ji et al., 2018). Global precipitation decreases by 2.3±0.5 % per Kelvin in response to G4 stratospheric aerosol injection (Ji et al., 2018). Precipitation and evaporation rates are strongly influenced by incoming radiation and the water vapor content of the troposphere. Hence solar geoengineering produces changes in both atmospheric circulation and thermodynamics. Several studies have analyzed changes in large scale circulation under the G1 solar dimming experiment (e.g., Tilmes et al., 2009; Davis et al., 2016; Smyth et al., 2017; Guo et al., 2018), but the more subtle changes under G4 have not yet been analyzed in similar depth. Broadly speaking, increasing greenhouse gases tend to produce a stronger Hadley circulation and enhanced hydrological cycle, increasing precipitation in the tropics and lowering it in the subtropics (the wet gets wetter and dry gets drier response) (Chou et al., 2013). Geoengineering, under both G1 solar dimming, and G4 aerosol injection, counteracts this response, decreasing tropospheric temperatures, and maintaining a higher pole-equator meridional temperature

gradient than under greenhouse gas forcing alone (Guo et al., 2018), and tending to reverse the wet dry patterns under greenhouse gas forcing (Ji et al., 2018; Wang et al., 2018). Stratospheric aerosol injection geoengineering produces a more complex climate response than produced by simple solar dimming (e.g. G1), as the aerosol layer not only scatters shortwave radiation, but also absorbs near-infrared and longer-wavelength radiation (Lohmann and Feichter, 2005; Niemeier et al., 2013; Ferraro et al. 2014). The net result of these changes in the GeoMIP experiments is model-dependent (Wang et al., 2018; Ji et al., 2018).

Under G4 stratospheric aerosol geoengineering, the global annual precipitation over land (excluding Greenland and Antarctic) decreases 9.8 mm relative to the reference experiment rcp45 experiment. The tropical Africa and south Asia regions suffer large precipitation reduction with values up to 17.2 mm and 24.8 mm per year (Figure 1a), southeastern Northern America and Alaska also see large precipitation decreases. In contrast, precipitation increases significantly over southern Africa and eastern Brazil under G4. Previous studies based on Global Land-Atmosphere Climate Experiment–Coupled Model Intercomparison Project phase 5 (GLACE-CMIP5) suggest strong coupling between local soil moisture and precipitation over southern Africa and eastern Brazil, both of which are simulated to experience large precipitation reduction under global warming (Seneviratne et al., 2013), which is reversed under G4. Although the precipitation increase under G4 over the Mediterranean region is not statistically significant, May et al. (2017) note soil moisture and precipitation both decrease under global warming. Lower temperatures under G4 result in a reduction of 7.6 mm in mean global land (excluding Greenland and Antarctic) evaporation relative to rcp45.

Under G4 stratospheric aerosol geoengineering, there is large precipitation reduction over the Indian subcontinent and East Asia monsoon regions of 7.4% and 4.2% respectively. Under G1, these are related to a reduced latitudinal seasonal amplitude of the ITCZ (Schmidt et al. 2012; Smyth et al., 2017), and a reduction in the intensity of the Hadley circulation (Guo et al., 2018). Precipitation over other monsoon regions in G4 sees less significant changes. The monsoonal precipitation reduction in Indian and East Asian regions is consistent with a weakened summertime monsoonal circulation under G4 (Fig. S9 f).

Displacement of mid-latitude westerlies and changes to the North Atlantic Oscillation, especially during winter, will change regional precipitation variations under G4. Ferraro et al. (2015) and Muri et al. (2018) found that the tropical lower stratospheric sulfate aerosol injection leads to a thermal wind response that affects the stratospheric polar vortices. The polar vortices guide winter mid-latitude jets and cyclone paths across the mid-latitudes. These circulation changes result in more moist maritime air into the Mediterranean region which increases precipitation (Fig. S9 e, f). Under a warming climate, an earlier spring snowmelt over northeastern Europe and a later onset of the winter storm season would both alter flooding conditions (Blöschl et al., 2017). Both these will also be affected by G4 stratospheric aerosol geoengineering.

Increased evaporation forecast under rcp45 is suppressed under G4 geoengineering, due to reduced downward surface radiation (Kravitz et al. 2013; Yu et al., 2015). Evaporation decreases over a significantly ($p<0.05$) broader area than precipitation, especially in the Northern Hemisphere (Figure 1b). The change of precipitation minus evaporation (P-E) basically follows the change of precipitation and evaporation, but is of a smaller magnitude (Figure 1c), due to their simultaneous reductions. There are significant reductions of P-E over south Asia, tropical eastern Africa and the Amazon basin, and significant increases over Southern Africa and eastern Brazil. Increased P–E in northern Asia caused by global warming could be partly counteracted by solar geoengineering (Jones et al., 2018; Sonntag et al., 2018). The simulated precipitation and evaporation changes under the G4 stratospheric geoengineering implies potentially significant changes in the terrestrial hydrological cycle. P–E can be used as a simplified measure of runoff and water availability. Under the G4 experiment, P–E increases over Europe during summer time, implying more water availability and shortened return period of river discharge. Soil moisture also reflects local water mass balance, i.e. the difference between P-E and runoff. Soil moisture increases over Southern Africa, southwestern North America and several parts of South America, where P-E and runoff both increase. The regions with both significant reductions in P-E and runoff also show decreases soil moisture, such as tropical Africa, south Asia and most of middle Northern America.

The spatial pattern of runoff change from rcp45 to G4 resembles that of P-E (Figure 1c,1d) with a broader area of significant changes. The annual runoff decreases by 2.5 mm, similar to the change in P-E. There are large runoff decreases over tropical Africa, South Asia, southeastern Northern America, the Amazon basin and Alaska. Runoff slightly increases over Southern Africa, southwestern North America and several regions of South America. Variability in runoff and streamflow is greater than for precipitation and evaporation (Figure 1, 2), due to spatial heterogeneity in soil moisture and because streamflow spatially integrates runoff (Chiew and McMahon, 2002).

Precipitation, evaporation and runoff changes show that land areas dry slightly, especially around the equator, south Asia and at northern high-latitudes under G4. Increases in P-E are predicted on the western parts of Europe and North America, with their eastern sides becoming drier with decreasing P-E and runoff.

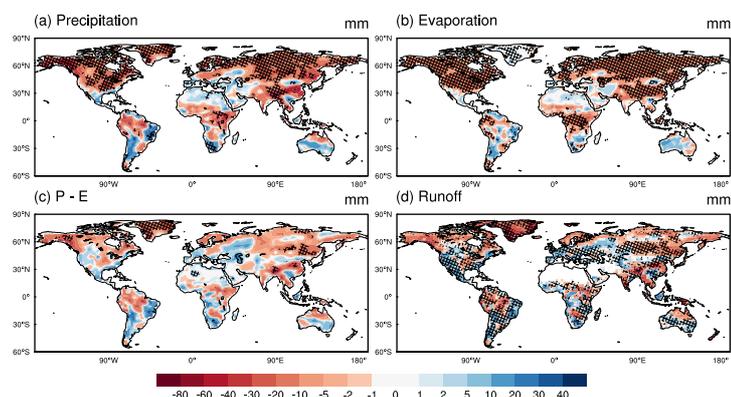



Figure 1: The absolute change (mm) of annual precipitation (a), evaporation (b), precipitation minus evaporation (P–E) (c) and runoff (d) between G4 and rcp45 during the period of 2030-2069.

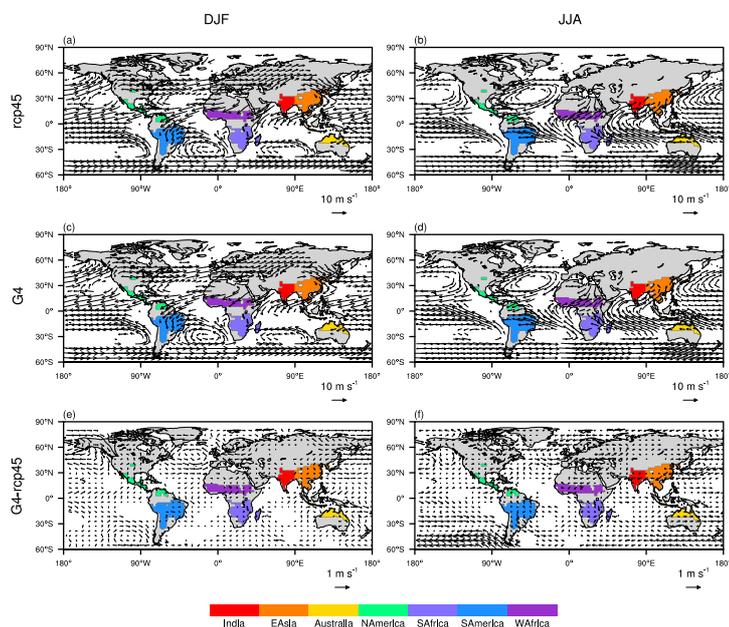



Figure S9: Multi-model ensemble mean of 925hPa wind field during (December-Janurary-Febreary, DJF) and (June-July-Augest, JJA) seasons. Panel (a) and (b) for rcp45, panel (c) and (d) for G4, panel (e) and (f) for the difference between G4 and rcp45. Grids where wind speed less than 1 m s⁻¹ are masked out in panel (a), (b), (c) and (d), grids where wind speed less than 0.025 m s⁻¹ are masked out in panel (e) and (f). Shaded monsoonal regions are derived using the criteria of Wang and Ding (2006) with the Global Precipitation Climatology Project (GPCP) data set covering the years 1979–2010.

2. *Show more results to back up some of the mechanistic claims in the discussion section (large scale circulation changes, monsoonal flows in different regions, flooding vs. terrestrial water availability). Right now, this section reads more like conjecture because the figures show only runoff-derived streamflow and flood return period.*

Reply: We add related analysis on large-scale circulation, monsoonal flow and water availability changes in the new section "3.1 Precipitation, evaporation and runoff changes" and add new Figure 1 and Figure S9. Please refer to our replies to your first point.

We also expand on the discussion of high latitude process in Section 4 (Discussion and Conclusions):

There is a latitudinal dependence for streamflow: generally, the Qm decreases across

all latitudes; high flow, Q5, decreases most in tropical regions; low flow, Q95, decreases most at high-latitudes. The high-latitudes display a complicated streamflow pattern with weakly increasing Q5 and significant decreasing Q95. The decrease in the lower probability tail of streamflow is indicative of hydrological droughts, while the increases in the high streamflow tail indicates hydrological flooding (Keyantash and Dracup, 2002). The balance among precipitation, evaporation and temperature accounts for the complex spatial pattern of streamflow and flood frequency under solar geoengineering, and has been related to soil moisture content (Dagon and Schrag, 2017). Previous studies (Dankers et al., 2014; Hirabayashi et al., 2008) have noted that the flood frequency for rivers at high latitude (e.g. Alaska and Siberia) decreases under global warming, even in areas where the frequency, intensity of precipitation, or both, are projected to increase. The annual hydrograph of these rivers is dominated by snow melt, so changes of peak flow reflect the balance between length and temperature of winter season, and the total amount of winter precipitation. The thawing of permafrost and changes in evapotranspiration also play an important role in the increasing of runoff and streamflow (Dai, 2016). The combined effect of atmospheric circulation and land surface processes results in the complex change pattern in this cold region.

3. *More discussion on model uncertainties (both in the climate models, and the river routing model) in the discussion section.*

Reply: This question is similar as that raised by referee#2 point 8 and we agree that this is an important topic. We address that as follows in Section 4 (Discussion and Conclusions):

Gosling et al. (2017) compared the river runoff output from multiple global and catchment-scale hydrological model under three global warming scenarios simulated by global climate models (GCMs), finding that the across-model uncertainty overwhelmed the ensemble median differences between the scenarios. In this study we use the offline hydrological model driven by runoff outputs from GCMs to calculate the streamflow, the uncertainty between GCMs is reflected in the range of return period based on streamflow change. Figure 9 shows the multi-model ensemble range of the 30-year return period level. Regions that have a shorter return period (i.e. higher flood frequency) from historical to the future, show a relatively small range among models (e.g. India and Southeastern Asia). Regions with a longer return period, show a large range (e.g. Europe and North America). This reflects larger inter-model uncertainty over dry zones than for wetter ones. The return period change over dry zones is more meaningful when interpreted as the change of drought tendency. 50- and 100-year return period levels show larger uncertainty than 30-year return period level, which is expected when estimating the low probability extreme tails of the flow probability density function from relatively short (40 year) sets of results.

In Section 4 (Discussion and Conclusions) we addressed limitations of the approach we use:

Previous studies suggest that the CaMa-Flood model can realistically reproduce peak river discharge because of the floodplain storage and backwater effects implemented in the routing model (e.g. Zhao et al., 2017). The river routing model CaMa-Flood is driven by the runoff output from ESMs to simulate streamflow. Therefore, the uncertainty in runoff in this study is also implied in ESMs. To drive high-resolution CaMa-Flood model, the coarse resolution runoff from ESM were regridded using a first-order conservation method. Although the regridding method conserves the mass of runoff, distributing the runoff from coarse climate model grids to fine river routing model grids introduces unavoidable errors. The relative magnitudes of this kind of error are dependent on the regional terrain and river routing map. The uncertainty in runoff might be transformed by the river routing model and overlap with the in-built bias of the river routing model itself. Comparing the ratio between inter-model spread and multi-model ensemble mean, we find that runoff usually has large inter-model spread in arid regions, and streamflow has large inter-model spread over a broader area than that of runoff. This is due to the streamflow integrating the runoff spatially along river routing map, therefore it carries the uncertainties of runoff to a relatively large extent.

Several studies have identified the uncertainty introduced by hydrological models (e.g., Chen et al., 2011; Prudhomme et al., 2014). We assume that systematic river routing model bias relative to observations can be alleviated by subtracting historical simulations, and simulated runoff biases are not expected to change significantly under future scenarios. In addition to model inherent biases, there are natural processes which could change river routes, and river network silt-up over time, these changes would impact local runoff and streamflow (Chezik et al., 2017), and we do not account for them in this study.

4. *Lack of anthropogenic effects in the river routing model (e.g. dams) seems like a big uncertainty in the results. What sort of effect might this have? Furthermore, how useful is analysis of a "hypothetical natural condition" (line 427) when the premise of the study implies large-scale human intervention in the climate system?*

Reply: As we also respond to Referee #2 point 10, we expand this interesting discussion point:

In this study we use runoff direct from ESMs to drive the river routing model CaMa-Flood to study streamflow and flood response. CaMa-Flood does not consider anthropogenic effects, such as dams or reservoirs, which some hydrological models do include. However, estimating future changes in human intervention on the natural system is highly uncertain. Technological advances over the century, that may affect anthropogenic changes, are by their nature entirely unknown at present. Hence integrating the human dimension into a model of the physical system is fraught with difficulty and uncertainty.

Several studies can be used as a guide to the possible effects of anthropogenic impacts compared with natural changes that are captured in CaMa-Flood. Dai et al. (2009) argued that the direct human influence on the major global river streamflow is relatively small compared with climate forcing during the historical period. Mateo et al. (2014) suggested that dams regulate streamflow consistently in a basin study, using CaMa-Flood combined

with integrated water resources and reservoir operation models. Wang et al. (2017) shows that the reservoir would effectively suppress the flood magnitude and frequency. Recently, analysis of the role of human impact parameterizations (HIP) in five hydrological models and found the inclusion of HIP improves the performance of GHMs, both in managed and near-natural catchments, and simulates fewer hydrological extremes by decreasing the simulated high-flows (Veldkamp et al., 2018; Zaherpour et al., 2018). These studies suggest that the high-flows and flood response under G4 relative to rcp45 might be smaller when human intervention is considered.

As anthropogenic GHG emissions increase, human society would continually adapt to climate change and mitigate the related risk, including building new dams and reservoirs to withstand enhanced strength of global hydrological cycle. How the society would response to future streamflow and flood risk becomes an increasingly important topic in both science research and policy-making. This is especially true for the developing world, where many cities are experiencing subsidence due to unsustainable rates of ground water extraction. Subsidence accounted for up to 1/3 of 20th century relative sea level rise in China (Chen, 1991). Sea level has risen fastest in deltas and coastal plains around the coastline of the China Seas largely due to the local subsidence (Chen, 1991; Ren, 1993). Subsidence and sea level rise both increase flooding risks. In particular, in densely populated regions with long experience of irrigation management, such as Southeast Asia and India, the reduced flood frequency under G4 stratospheric aerosol geoengineering might be further ameliorated.

The accurate assessment of human impacts on flood frequency and magnitude depends not only how human activities are represented in geoengineering scenarios, but also on how anthropogenic effects are parameterized in hydrological models (Masaki et al., 2017). Using the outputs from climate models to drive river routing models or hydrological models is a reasonable way to study how the streamflow and flood response under different climate changing scenarios. River routing models driven by runoff directly from GCMs and hydrological models considering human impacts both contribute to better our understanding of how the hydrological cycle would change under solar geoengineering.

Specific comments:

Line 22: Lower/higher relative to which simulation?

Reply: The lower/higher flow under G4 is compared with RCP4.5 scenario.
We rephrased this sentence as the following:
Compared with rcp45, streamflows on the western sides of Eurasia and North America are increased under G4, while the eastern sides see a decrease.

Line 27: How does the return period show increased drying?

Reply: We analyzed the precipitation, evaporation, runoff and soil moisture change between G4 and rcp45, and find there is a weak increase of soil moisture and a

significant decrease in runoff at Amazon basin under G4. Therefore, we rephrase this sentence as the following:

Although G4 stratospheric aerosol geoengineering ameliorates the Amazon drying under rcp45, with a weak increase in soil moisture, the decreased runoff and streamflow leads to increased flood return period under G4 compared to rcp45.

Line 42: Connect the text descriptions to flow abbreviations (I assume percentiles are used to define high/low flow but that is not clear from the text.) I see these are defined later in the methods (lines 158-159) but should be defined at first mention. Alternatively, don't use the abbreviations in the introduction.

Reply: Thanks. We've corrected this error and defined the flow abbreviations at first mention.

Line 45: How do the effects on streamflow scale with the emissions scenario (or the amount of CO2)?

Reply: This sentence describes the relevant result from Koirala et al. (2014), in which they mostly focused on the changes of streamflow indicators under RCP8.5 scenario relative to historical period, and gives less details on relative changes in the RCP4.5 scenario. We rephrase this sentence as the following to make it clearer:

Under the RCP4.5 scenario, the spatial distributions of changes are similar to those under RCP8.5. The changes of mean and high streamflow are smaller under RCP4.5 than those under RCP8.5, while the change in low flow are similar under both scenarios.

Line 70: Change model to model's

Reply: Done.

Line 78: Define SO2; how many models?

Reply: Thanks. We rephrase this sentence and combine our reply to your next comment in the following to clarify the definition of the GeoMIP G4 experiment:

Under the Geoengineering Model Intercomparison Project (GeoMIP; Robock et al., 2011; Kravitz et al., 2011, 2012, 2013a), the G4 experiment specifies a constant 5Tg sulfur dioxide (SO2) per year injection to the tropical lower stratosphere, or equivalent aerosol burden, during the period of 2020-2069 to mimic one-fourth of the 1991 eruption of Mount Pinatubo. At the same time, the greenhouse gas forcing is defined by the RCP4.5 scenario. There are nine coupled climate models that take part in the GeoMIP G4 experiment, and sulfate aerosols are handled differently among the participating models. For example, BNU-ESM and MIROC-ESM use the prescribed meridional distribution of AOD recommended by the GeoMIP protocol; CanESM2 specifies a uniform sulfate AOD (Kashimura et al., 2017); GISS-E2-R and HadGEM2-ES adopt stratospheric aerosol schemes to simulate the sulfate aerosol optical depth (AOD); NorESM1-M specifies the AOD and effective radius, which were calculated in

previous simulations with the aerosol microphysical model ECHAM5-HAM (Niemeier et al., 2011; Niemeier and Timmreck, 2015).

Line 80: Is aerosol injected or SO2?

Reply: Please refer to our reply to your previous comment.

Line 103: Is scenario the right word here?

Reply: Thanks. We corrected the word "scenario" to "change".

Line 115: Why is it important to use the same 40 years?

Reply: The first decade (2020-2029) of the G4 experiment follows the abrupt increase in stratospheric aerosol forcing, which likely exerts a large and fast perturbation to the climate system with various possible system transients. To minimize the effects of possible transients, we use the last 40 years to analyze the streamflow changes. Several previous studies make the same choice. We rephrased this sentence to clarify this point:

We exclude the first decade of the G4 simulation from our analysis because it follows the abrupt increase in stratospheric aerosol forcing, which likely exerts a large perturbation to the climate system, and analyze the streamflow pattern changes between each of model's G4 and rcp45 simulations during the period 2030-2069. Using the last 40 years of G4 simulations is common in previous studies (e.g. Curry et al. 2014; Ji et al. 2018). The historical simulation covering the period 1960-1999 is used as the reference for the return period analysis.

Line 124: How are the different model realizations generated? What is the impact of using a single historical run (MIROC-ESM-CHEM) as reference for multiple experiment simulations?

Reply: This is essentially the same question as Referee #2 point 5. And we address both points in the same way.

The multiple runs in Table 1 are the number of realizations of the experiment that each model made. BNU-ESM, CanESM2, MIROC-ESM and NorESM1-M all have the same number of realizations for its historical, rcp45 and G4 experiments. Their rcp45 and G4 runs are branched from the end of corresponding historical runs, while their historical runs are branched from each model's pre-industrial control runs, that were started with different initial states. MIROC-ESM-CHEM has only one historical run and its three rcp45 and G4 runs are branched from this same historical run with different initialization perturbations. GEOSCCM has no historical run, its rcp45 and G4 runs are forced with sea surface temperature and sea ice concentrations, as simulated by the CESM rcp45 runs. Equal weight is given to each model in the analysis, and streamflow and flood response are calculated for each model before multi-model ensemble

averaging is done. For models with multiple realizations, streamflow and flood response are calculated for individual realization and then averaged for each model.

To ensure our analysis is consistent on streamflow and flooding return period changes under rcp45 and G4 scenarios, we now remove the GEOSCCM model due to it lacking corresponding historical runs and also the 2nd and 3rd rcp45 and G4 realizations of MIROC-ESM-CHEM which also have no corresponding historical run. We find our conclusions hold with the reduced ensemble members. The following two figures show the streamflow changes with all models and their realizations included as in the submitted manuscript (left panel) and the streamflow changes with GEOSCCM model and the 2nd and 3rd rcp45 G4 realizations of MIROC-ESM-CHEM excluded (right panel).

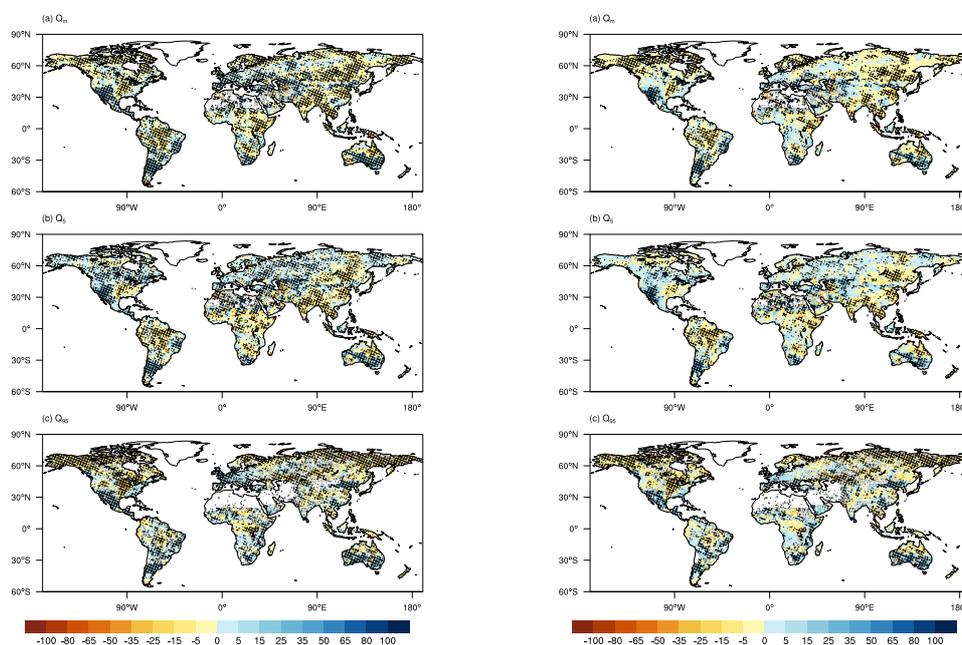



Table 1: For the models with multiple experiments for each type of simulation, are these ensemble members? Specify this in table caption or text. Also, better to define horizontal resolution as degrees lat/lon.

Reply: They are the number of ensembles of each model included in our study. We revised the table and change the definition of horizontal resolution as degrees lat/lon according to your suggestion:

Table 1: GeoMIP models and experiments used in this study.

Model	Resolution (degrees lat × lon, level)	Number of ensembles		
		Historical	RCP4.5	G4
BNU-ESM (Ji et al., 2014)	2.8 × 2.8, L26	1	1	1
CanESM2 (Arora et al., 2011; Chylek et al., 2011)	2.8 × 2.8, L35	3	3	3

MIROC-ESM (Watanabe et al., 2011)	2.8 × 2.8, L80	1	1	1
MIROC-ESM-CHEM (Watanabe et al., 2011)	2.8 × 2.8, L80	1	1	1
NorESM1-M (Bentsen et al. 2013, Tjiputra et al. 2013)	1.9 × 2.5, L26	1	1	1

Line 136: Is "FLOW" an acronym? If so, please define.

Reply: yes, it's an acronym for "Flexible Location of Waterways". We revise this sentence as the following to clarify:

… in each grid box by using the innovative up-scaling method, Flexible Location of Waterways (FLOW), (Mateo et al., 2017; Yamazaki et al., 2014b; Zhao et al., 2017).

Line 141: Need a sentence break here.

Reply: Thanks. Done.

Line 145: How good are the daily runoff outputs from the climate models? Is this discussed anywhere (e.g., in the discussion section)?

Reply: The daily runoff from the climate models are largely affected by the daily precipitation and displays significant variability. Therefore, the runoff is usually evaluated over a longer time scale. We add the following paragraph in the Discussion section:

Global spatially continuous and temporally variable observations of runoff are not available (Ukkola et al., 2018). Model simulated runoff is usually compared with observed downstream river discharge datasets, with the dataset collected by Dai et al. (2009; 2016) being the most complete. The Dai et al. (2016) dataset represents historical monthly streamflow at the farthest downstream stations for the world's 925 largest ocean-reaching rivers from 1900 to early 2014. However, the length and reliability of the available time series vary greatly from one river basin to another and contain gaps. As daily runoff is largely affected by daily precipitation, it is difficult to evaluate how good the runoff outputs from the climate models are at a daily scale. Over longer time scales, Alkama et al. (2013) found the CMIP5 models simulate mean runoff reasonably well (±25% of observed) at the global scale. The CMIP5 models tend to slightly underestimate global runoff, with South American runoff underestimated by all models. Koirala et al. (2014) found more CMIP5 model agreement on runoff projections under RCP8.5 than under RCP4.5 scenario, but the projected changes in low flow are robust in both scenarios with strong model agreement.

Line 174: Does "generated data" refer to runoff output from the climate model or streamflow output from the river routing model?

Reply: The "generated data" refer to the resampled streamflow output from the river routing model. We rephrase this sentence and its context as the following to clarify this point:

Specifically, we first apply the MW-U test to the G4 and rcp45 annual mean daily streamflow data for each model to get the value of the rank sum statistical value, U0. Then we generate 1000 random paired series of 40-year streamflow data from rcp45 and G4 using the bootstrap resampling method, and apply the MW-U test to each sample pair of generated streamflow data to get a series of statistical values: U_j, $j = 1,2 \cdots 1000$.

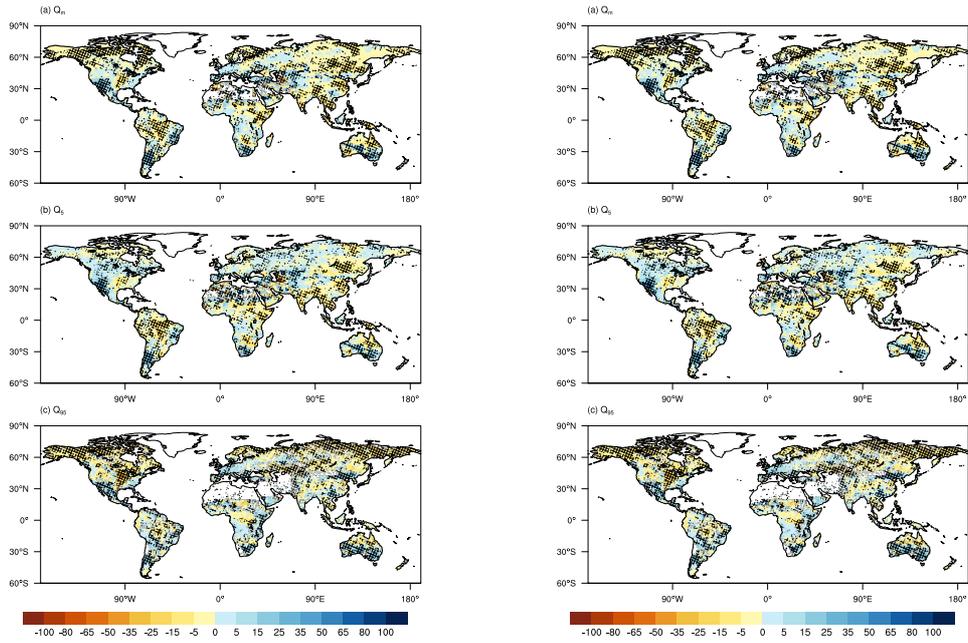
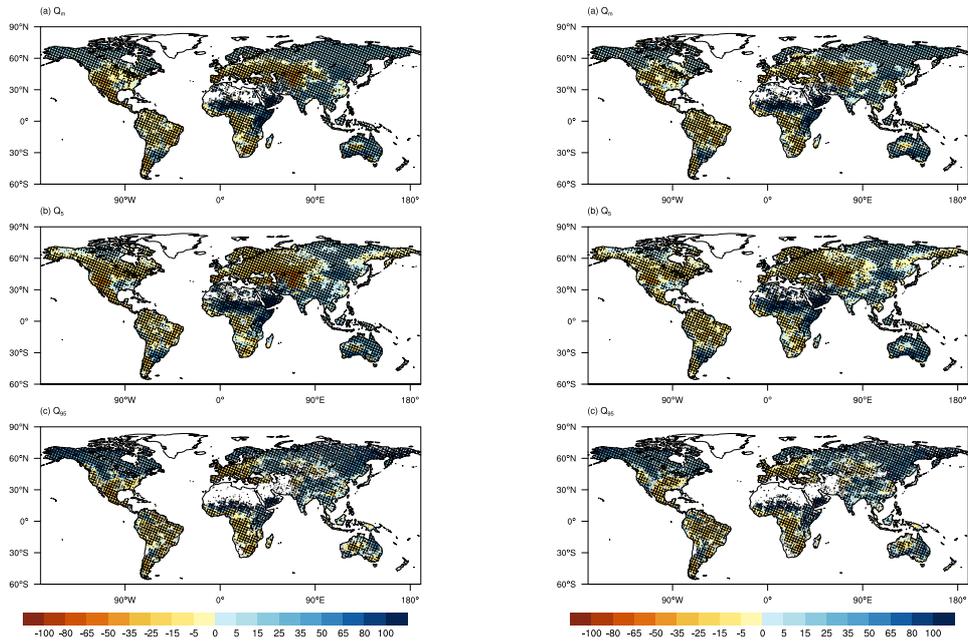
Line 184: Should that be 1:N instead of 1/N?

Reply: No. The probability in an N year return period flood is 1/N in any single year.

Line 240: This is an important point, distinguishing changes in flow level with changes in flood frequency.

Reply: Thanks. We revise this sentence a bit to make it more accurate:
Though high flow levels usually correspond with flood events, changes in flow levels do not necessarily translate into increases in flood frequency (Ward et al., 2016).

Figure 1: Can you explain more about the metric plotted here? Why are you using the mean of G4 and RCP4.5? It would also be useful to see this figure for the difference between G4 and historical/present day climate.

Reply: In our previous Figure 1, we used the metric "(G4−rcp45)/(G4+rcp45)/2×100%" to measure the relative streamflow changes and avoid the values near zero in denominator. Perhaps this metric is not particularly intuitive, so now we use the metric "(G4-rcp45)/rcp45×100%" and filter out grids with streamflow smaller than 0.01mm/day. We get very similar spatial patterns as the previous metric. As the new metric is more straightforward, we use the new metric in revised manuscript. The following figures show the streamflow indicators change using two metrics, left panel uses the metric "(G4−rcp45)/(G4+rcp45)/2×100%", right panel uses the metric "(G4-rcp45)/rcp45×100%". Main visual differences between two metrics occur over North Africa.

We also use the new metric to show relative changes of rcp45 (left panel) and G4 (right panel) relative to historical climate. The metric we use here is "(rcp45-historical)/historical×100%" for rcp45, "(G4-historical)/historical×100%" for G4.

Line 280: I see streamflow increases (blue colors) on the western side of large continents in Fig 1 (Mexico, southern California, Spain, western Europe); please elaborate.

Reply: Thanks. Yes, we mistakenly reversed the meaning in the text. We revised this sentence as the following:

Perhaps the clearest overall pattern is the streamflow generally increasing under G4 on the western sides of the large continents of Eurasia and North America, especially over Mexico, southern California, Spain and western Europe, while streamflow decreases on the eastern sides of these continents.

Figure 3: The color bar labels are confusing and should at least be in larger font.

Reply: Yes, agreed. We improve the figure by using up or down arrow to represent the increase or decrease tendency. The revised figure is:

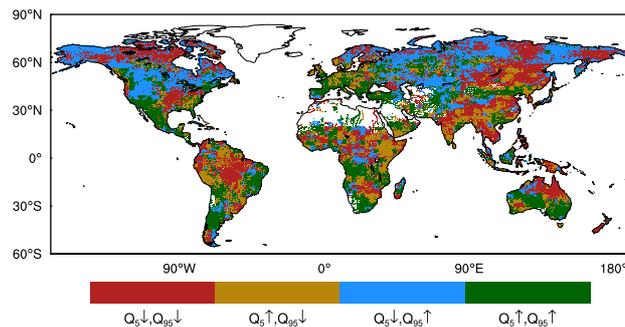



Figure 3: The ensemble mean difference (G4-rcp45) of high (Q5) and low (Q95) streamflow. Color bar is defined such that grids where G4 is less than RCP4.5 for both Q5 and Q95 is in red (Q5↓Q95↓); both Q5 and Q95 greater in G4 than rcp45 is in green (Q5↑Q95↑); Q5 greater in G4 and Q95 greater in rcp45 in yellow (Q5↑Q95↓) and vice versa in blue (Q5↓Q95↑). Grid cells with Q95 less than 0.01 mm/day are masked.

Line 374: Can you show some results that back up these mechanistic claims?

Reply: Yes, we produce new Figure 1 and Figure S9 showing differences in wind patterns between G4 and rcp45. Please refer to our reply to the first of your general comments.

Line 390: Not sure anyone would "benefit from increased flooding", perhaps increased water availability? Do you have the results to show that?

Reply: Figure 1 in our reply to the first of your general comments shows that precipitation-evaporation and runoff increase over the southwestern USA, Mexico and much of Australia, which means the soil moisture increase there as well. We rephrase this sentence as the following:

Generally, stratospheric aerosol injection geoengineering relieves flood stress, especially for Southeast Asia, and in turn increases the probability of flooding in the southwestern USA, Mexico and much of Australia – which are drought-prone places that might benefit from increased soil moisture and streamflow.

Line 406: Change "G4" to "solar geoengineering"

Reply: Thanks. Done.

Lines 411-418: This sentence is too long and convoluted. It needs to be cleaned up or broken into multiple sentences to clarify the important points.

Reply: Thanks. We rephrased this sentence and its context as the following:

Amazon basin drying is complicated by various factors that are dependent on solar geoengineering. These include i) the reduced seasonal movement of Intertropical Convergence Zone (ITCZ) under solar geoengineering (Smyth et al., 2017; Guo et al., 2018); ii) Changes in SST reflecting changes in frequency of El Niño Southern Oscillation (Harris et al., 2008; Jiménez-Muñoz et al., 2016), although there is no evidence of changes occurring under SRM (Gabriel and Robock, 2015); and iii) changes in carbon cycle feedbacks (Chadwick et al., 2017; Halladay and Good, 2017), which would certainly be affected by changes in diffuse radiation under SRM (Bala et al., 2008).

Line 420: Why only this region?

Reply: We greatly revised this section, please refer to our reply to the third of your general comments.

Line 437: I like the reference to the DECIMALS project. I wonder if you could include a bit more detail here about potential connections to socioeconomic research, based on the results of this study. I can see this discussion being very useful to researchers in climate change adaptation, urban design, and hydrologists (among others).

Reply: Thanks for your constructive comment. We add the following paragraph in the Discussion section:

Floods are among the most costly natural disasters around the world, especially for more vulnerable developing countries (e.g. Bangladesh, India and China). Our study suggests that solar geoengineering would exert non-uniform impacts on global flooding risk and hence local hydraulic infrastructure needs would vary if solar geoengineering of the G4-type were undertaken. This highlights the importance of carrying regional impact studies of solar geoengineering. Recently, a fund called Developing Country Impacts Modelling Analysis for SRM (DECIMALS) was announced (Rahman et al., 2018). Developing-country scientists are encouraged to apply DECIMALS to model the solar-geoengineering impacts that matter most to their regions. DECIMALS promotes wider discussion of the implications of regional impacts studies of solar geoengineering. These studies will be a helpful initial step in future decision making related to climate change adaptation and urban infrastructure design.

References:

Chen, X. Q.: Sea level changes since the early 1920's from the long records of two tidal gauges in Shanghai, China. J. Coastal Res. 7(3), 787-799. http://www.jstor.org/stable/4297894, 1991.

Ren, M. E.: Relative sea level changes in China over the last 80 years. J. Coastal. Res. 9(1),

229-241. http://www.jstor.org/stable/4298080, 1993.

Schmidt, H., Alterskjær, K., Bou Karam, D., Boucher, O., Jones, A., Kristjánsson, J. E., Niemeier, U., Schulz, M., Aaheim, A., Benduhn, F., Lawrence, M., Timmreck, C.: Solar irradiance reduction to counteract radiative forcing from a quadrupling of CO2: climate responses simulated by four earth system models. Earth Syst. Dyn. 3(1), 63–78, 2012.

Ferraro, A. J., Highwood, E. J., and Charlton-Perez, A. J.: Weakened tropical circulation and reduced precipitation in response to geoengineering, Environ. Res. Lett., 9, 014001, https://doi.org/10.1088/1748-9326/9/1/014001, 2014.

Muri, H., Tjiputra, J., Otterå, O.H., Adakudlu, M., Lauvset, S. K., Grini, A., Schulz, M., Niemeier, U. and Kristjánsson, J. E.: Climate response to aerosol geoengineering: a multi-method comparison. Journal of Climate, https://doi.org/10.1175/JCLI-D-17-0620.1, 2018.