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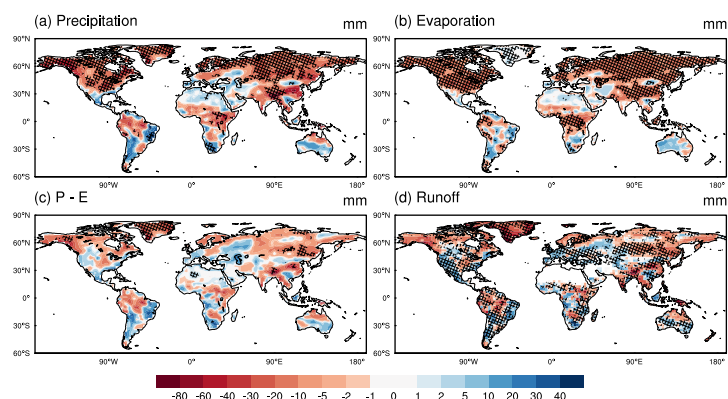
Figure 1: The absolute change (mm) of annual precipitation (a), evaporation (b), precipitation minus evaporation (P–E) (c) and runoff (d) between G4 and rcp45 during the period of 2030-2069.
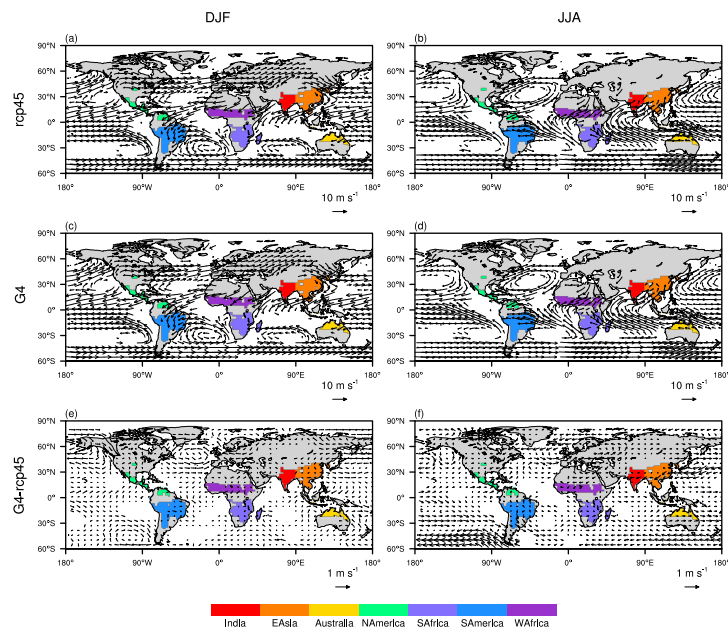

Figure S9: Multi-model ensemble mean of 925hPa wind field during (December-Janurary-Febreary, DJF) and (June-July-Augest, JJA) seasons. Panel (a) and (b) for rcp45, panel (c) and (d) for G4, panel (e) and (f) for the difference between G4 and rcp45. Grids where wind speed less than 1 m s^{-1} are masked 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.
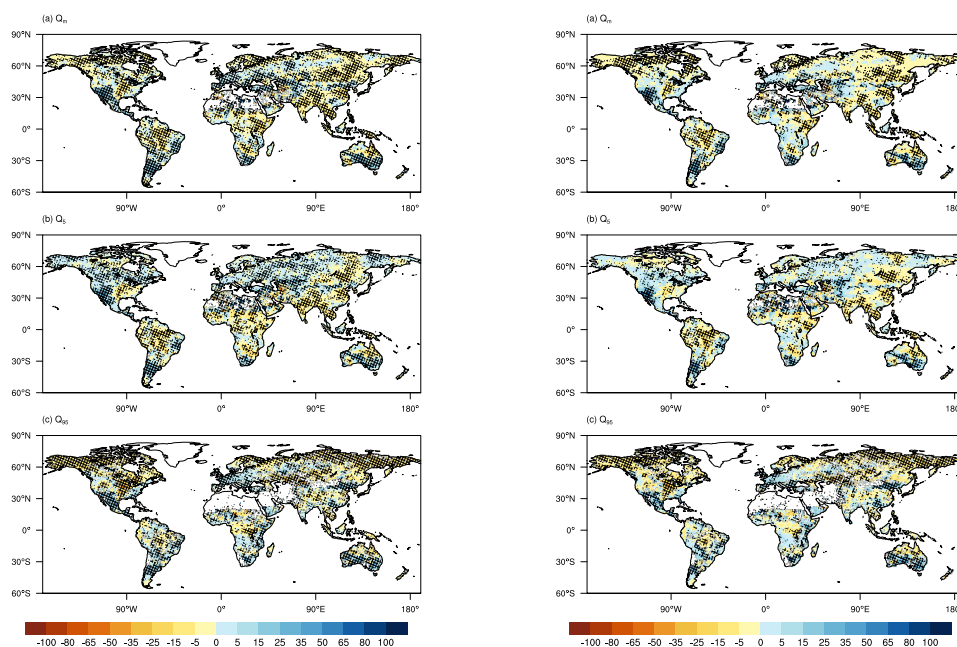
Line 124: How are the different model realizations generated? What is the impact of using a single historical run (MIROC-ESM-CHEM) as reference for multiple experiment simulations?

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.
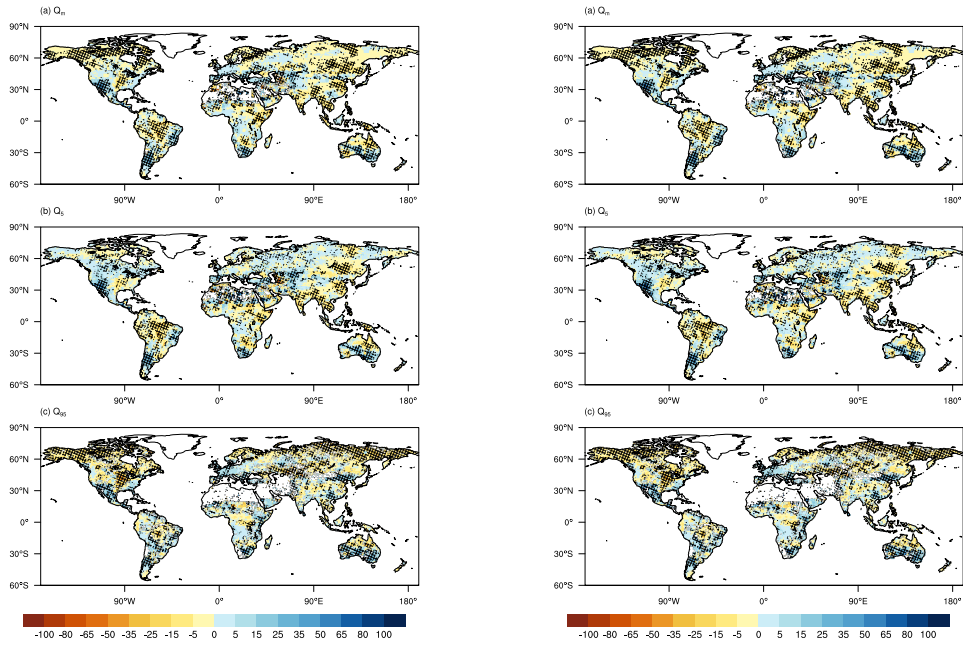
We also use the new metric to show relative changes of rcp45 (left panel) and G4 (right panel) relative to historical climate. The metric we use here is "(rcp45-historical)/historical×100%" for rcp45, "(G4-historical)/historical×100%" for G4.
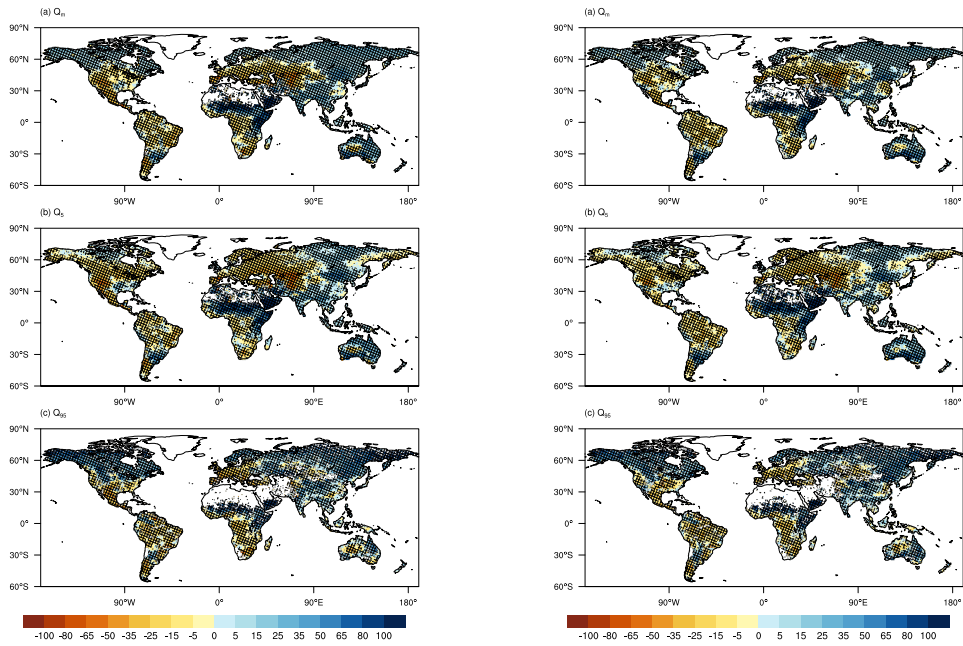
Line 280: I see streamflow increases (blue colors) on the western side of large continents in Fig 1 (Mexico, southern California, Spain, western Europe); please elaborate.

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:
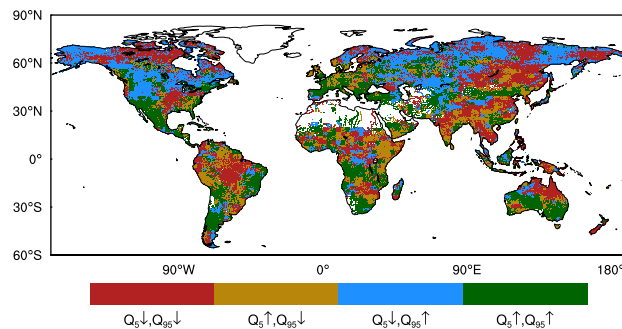

Figure 3: The ensemble mean difference (G4-rcp45) of high (Q5) and low (Q95) streamflow. Color bar is defined such that grids where G4 is less than RCP4.5 for both Q5 and Q95 is in red (Q5↓Q95↓); both Q5 and Q95 greater in G4 than rcp45 is in green (Q5↑Q95↑); Q5 greater in G4 and Q95 greater in rcp45 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.

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 #2

The authors present a suite of simulations from six GCMs that participated in GeoMIP, under two scenarios: RCP4.5 and an SRM scenario from GeoMIP (G4). The authors note that this is the first study to assess how SRM might affect global-scale streamflow. The authors compare the two scenarios to demonstrate what effect SRM might have compared with a non-SRM scenario in a mid-emissions (RCP4.5) future, in terms of high and low flows, mean flow, and return period flows.

General Comments:
1. *Lines 76-80: "Under the Geoengineering Model Intercomparison Project (GeoMIP; Robock et al., 2011; Kravitz et al., 2011, 2012, 2013a), G4 experiment a constant 5Tg per year of SO2 is introduced into the lower tropical stratosphere of climate models during the period of 2020-2069, while greenhouse gas forcing is defined by the RCP4.5 scenario". It would be helpful to readers if a little more information could be said about G4, to provide some context to the results. For example, how does the injection affect the time-series of global-mean temperature (a graph showing global-mean temperature would be quite informative here, or at least a statement of the magnitude of global cooling achieved by G4 relative to RCP4.5).*

   Reply: Yes, we agree that more details are needed on the basic response of G4 compared with RCP4.5. As this is just the introduction, we don't want to introduce figures yet. In fact, we do add a more detailed description of the results from the G4 experiment in a new section 3.1 (Precipitation, evaporation and runoff changes) later. In the Introduction we already state (line 83) "The direct radiative effects mainly result in the sharp reduction of TOA net radiative flux with a significant drop in global surface temperature, and concomitant decrease in global precipitation (Yu et al., 2015).", giving the mechanism why immediately afterwards. We also state a few side-effects of the sulphate injection. We prefer to leave the more detailed analysis to the Results section 3.1 later.

2. *Lines 61-62: "River flood models such as CaMa-Flood (Yamazaki et al., 2011) are important tools for simulating flood hazard." Yes they are, but CaMa-Flood is better described as a "river routing model" as opposed to a "river flood model". The latter implies that hydrological processes are included explicitly but they are not, since routing models take the outputs of hydrological models or climate models and route them through a network.*

Reply: Agreed. We rephrase this sentence as the following:

River routing models, such as CaMa-Flood, are important tools for the future flood hazard projection.

3. *Lines 64-66: "The high-resolution models have contributed to better simulation of river discharge (Yamazaki et al., 2009; Yamazaki et al., 2013 and Mateo et al., 2017)". This is a fair point and it is worth noting that 'offline' (i.e. separate from the hydrological model) routing models, such as CaMa-Flood specifically, have resulted in better agreement between simulated and observed discharge, compared with when the native hydrological model routing methods are used (see Zhao et al., 2017).*

Reply: Agreed. We rephrase this sentence as the following:

The high-resolution offline river-routing models, such as CaMa-Flood, have contributed to better simulation of river discharge (Yamazaki et al., 2009; Yamazaki et al., 2013; Mateo et al., 2017). Zhao et al. (2017) use daily runoff from Global hydrological models (GHMs) driving CaMa-Flood to produce monthly and daily river discharge. Zhao et al. (2017) find that this approach results in better agreement between simulated and observed discharge, compared to using native hydrological model routing. The CaMa-Flood model accounts for floodplain storage and backwater effects that are not represented in most GHM native routing methods (Yamazaki et al., 2014; Zhao et al. 2017; Mateo et al., 2017).

4. *Line 79: "while greenhouse gas forcing is defined by the RCP4.5". I appreciate that this scenario can be used with G4 but it is worth noting in the Discussion section, that the general conclusions drawn from this research are based upon these specific scenarios, i.e. G4 and RCP4.5. There are three other emissions scenarios under the RCPs (2.6, 6 and 8.5), which means the simulated offsetting (or otherwise) effects of SRM, particularly in terms of magnitude, could be different if the underlying emissions scenario was different (i.e. RCP2.6, 6 or 8.5).*

Reply: This relates to the linearity of response of both greenhouse gas and solar radiation management. This has been explored in several studies over the past 5 years, in particular using control theory methods (e.g. MacMartin et al., 2014). At least in the climate models the responses to most variables such as global temperature, precipitation are surprisingly linear, while the response of others, particularly associated with ice/water phase change, e.g. sea ice extent, are not. This general linear response spans the full set of RCP scenarios and their radiative forcing negation by sulphate injection. Several papers discuss the control theory and linearization in terms of GeoMIP experiments, and also with the CESM1 models using a large ensemble (Tilmes et al., 2018). We add the following paragraph in Discussion section to clarify this point:
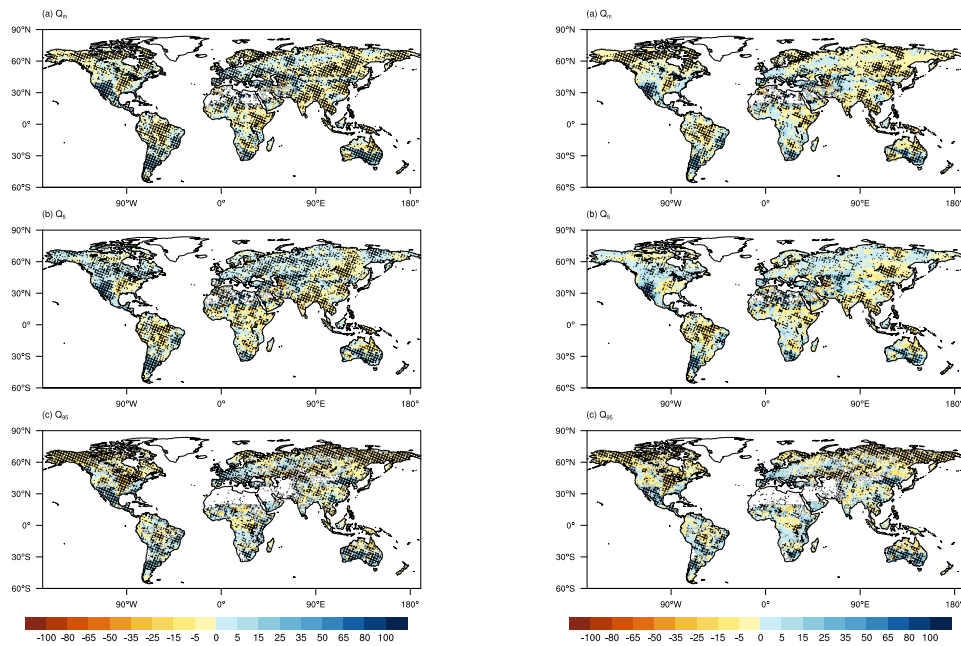
Our results on streamflow and flood response are based on GeoMIP G4 experiment and its reference rcp45 experiment. The generalizations of the work to other types and extents of solar geoengineering depends on the linearity of the streamflow response to both

greenhouse gas and geoengineering. The linearity of response of radiative forcing and global temperatures in particular have been explored in CESM1 stratospheric aerosol Geoengineering Large Ensemble (GLENS, Tilmes et al., 2018). Many climate fields, such as temperature, are surprisingly linear under a very wide range of forcing, potentially allowing standard engineering control theory methods (e.g. MacMartin et al., 2014) to tailor a global response given the freedom to use different latitudinal input locations for the aerosol injection (MacMartin et al., 2018; Kravitz et al., 2018), or combinations of, for example aerosol injection and marine cloud brightening (Cao et al., 2017). Non-linearities are expected for systems that depend on ice/water phase changes, and these could affect global streamflow and flood responses in some regions, especially in the Arctic. Moreover, the type of solar geoengineering might be relevant as well. Ferraro et al. (2014) found that the tropical overturning circulation weakens in response to geoengineering with stratospheric sulfate aerosol injection due to radiative heating from the aerosol layer, but geoengineering simulated as a simple reduction in total solar irradiance does not capture this effect. A larger tropical precipitation perturbation occurs under equatorial injection scenarios (such as G4) than under simple solar dimming geoengineering, or the latitudinal varying injections schemes explored by GLENS, or a mix of different geoengineering strategies (such as aerosol injection and marine cloud brightening, Cao et al., 2017). So the response of streamflow and flood would be expected to differ, to some extent, under different types of solar geoengineering.

5. *Table 1: Several of the GCMs include multiple runs for a single scenario, e.g. 3 runs for the CanESM-2 historical scenario. Can the authors please explain what this means? Is it a perturbed parameter ensemble of three members, or something else? Table 1. In Section 2.1 the authors also need to explain how the multiple runs in this table were dealt with. Was an ensemble mean used where there were three runs for one GCM, or were the calculations performed for each of the three runs in turn? From lines 159-160 it appears as though the runs were averaged, but it would be helpful to clarify this in Section 2.1.*

Reply: 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 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).

We revised Table 1 to reflect these changes:

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

6. *Line 148: "an adaptive time step approach was applied in simulation". Can the authors please explain in detail what this means in practical terms? I had presumed that CaMa-Flood was run at daily temporal resolution for all GCMs, but this text suggests that this is not the case.*

Reply: We rephrase this sentence as the following to clarify:

The spatial resolution of CaMa-Flood is set to 0.25° (~25km at mid-latitudes). An adaptive time step scheme was applied in the model numerical integration leading to a time step of about 10 minutes, while the model outputs at daily temporal resolution.

7. *Line 150: "to the fine resolution hydrological model". I made this point earlier – CaMa-Flood is not a hydrological model, it's a routing model.*

Reply: We rephrase this sentence as the following:

In order to conserve the input runoff mass, an area-weighted averaging method is used in CaMa-Flood to distribute the coarse input to the fine resolution routing model.

8. *Section 2.4. and Figure 4: the authors calculate 30, 50 and 100-year return period levels of flows and then calculate the average across all GCMs. This approach is reasonable but in applying this method the authors overlook two important uncertainties that could influence the results significantly: 1) climate model uncertainty (from using several GCMs); and 2) statistical uncertainty introduced by calculating extreme flows for return periods that are longer than the period used to calculate them (40 years). It is known that climate model uncertainty can result in return period flows that vary more between GCMs than they do between warming scenarios, and that the range in return period flows across GCMs can be significant (Gosling et al., 2017). The authors may therefore like to consider presenting the range across all GCMs (as opposed to just the ensemble median).*

Reply: Thanks for your constructive suggestions, we add the following paragraph in Results section:

Gosling et al. (2017) compared the river runoff output from multiple global and catchment-scale hydrological model under three 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 the shorter return period (i.e. higher flood frequency) from historical to future, show a relatively small range among models (e.g. India and Southeastern Asia). Regions that have the 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 flow 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.
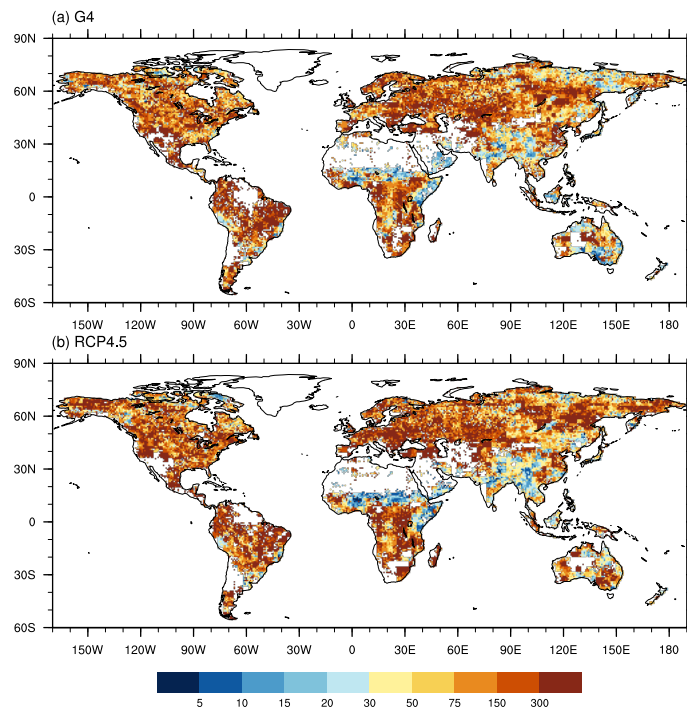
Figure 9: Multi-model ensemble range of return period for discharge that correspond with the 30-year return period in the historical simulation (1960-1999) under (a) G4 and (b) RCP4.5 scenarios, as the difference between maximum and minimum return periods. Grid cells in extremely dry regions, i.e. $Q_m < 0.01$ mm/day and extreme high value of return period regions were masked out.

9. *Lines 399-402: "Under the G4 experiment, some recent studies (Jones et al., 2018; Sonntag et al., 2018) have pointed out that the increased P–E (difference between precipitation and evaporation) in northern Asia caused by global warming could be partly counteracted by solar geoengineering." It is perhaps also worth noting that the way in which evapotranspiration is estimated is quite important, as this can vary significantly between different models (Wartenburger et al., 2018).*

Reply: Agreed. We add following sentences to reflect the uncertainties on estimating evapotranspiration:

The method for calculating potential evapotranspiration (ET) plays a significant role in determining simulated surface runoff changes (Haddeland et al., 2011; Thompson et al., 2013), which would influence the condition of streamflow. A recent study (Wartenburger et al., 2018) compared the ET spatial and temporal patterns simulated by GHMs in second phase of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2a) also confirmed that the ET scheme used affects model ensemble variance. The ET in this study is calculated by the ESMs (Table 1), not GHMs, and any biases in ET would feed into streamflow. For example, Mueller and Seneviratne (2014) found that climate models which

participated in CMIP5 display an overall systematic overestimation of annual average ET over most regions, particularly in Europe, Africa, China, Australia, Western North America, and part of the Amazon region.

10. *Lines 425-427: "The CaMa-Flood river routing model also does not consider anthropogenic effects on rivers (e.g. dams), so the results presented here are for a hypothetical natural condition." This is true but can the authors explain how this may have affected their results? Would the differences be smaller or lager if human impacts were included? This could lead on to an interesting discussion on the relative value of using runoff direct from GCMs compared with inputting precipitation and other variables from GCMs into hydrological models that include human impacts. Recent work with hydrological models shows that including dams etc. within them improves their representation of river flows compared with excluding dams (Veldkamp et al., 2018; Zaherpour et al., 2018), but also that the way human impacts such as dams are presented is quite important (Masaki et al., 2017) – so, does this mean that we should be using hydrological models that include human impacts to assess changes in the hydrological cycle with SRM, or is it reasonable to use naturalised runoff direct from GCMs instead? Clearly there is no straightforward answer but the Discussion chapter could be enhanced by considering this important issue.*

Reply: Thanks for your constructive suggestions. We agree this is an interesting and important topic and so we add the following paragraphs to discuss this issue:

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 infrastructures, 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 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 emission increasing, 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 to an 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 the 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.

References:

Cao, L., Duan, L., Bala, G. and Caldeira, K.: Simultaneous stabilization of global temperature and precipitation through cocktail geoengineering, Geophys. Res. Lett., 44(14), 7429–7437, doi:10.1002/2017GL074281, 2017.

Kravitz, B., MacMartin, D. G., Mills, M. J., Richter, J. H., Tilmes, S., Lamarque, J.-F., Tribbia, J. J. and Vitt, F.: First Simulations of Designing Stratospheric Sulfate Aerosol Geoengineering to Meet Multiple Simultaneous Climate Objectives, J. Geophys. Res. Atmos., 122(23), 12,616-12,634, doi:10.1002/2017JD026874, 2017.

MacMartin, D. G., Kravitz, B., Keith, D. W. and Jarvis, A.: Dynamics of the coupled human-climate system resulting from closed-loop control of solar geoengineering, Clim. Dyn., 43(1–2), 243–258, doi:10.1007/s00382-013-1822-9, 2014.

MacMartin, D. G., Kravitz, B., Tilmes, S., Richter, J. H., Mills, M. J., Lamarque, J.-F., Tribbia, J. J. and Vitt, F.: The climate response to stratospheric aerosol geoengineering can be tailored using multiple injection locations, J. Geophys. Res. Atmos., 1(November), doi:10.1002/2017JD026868, 2017.

Tilmes, S., Richter, J. H., Kravitz, B., MacMartin, D. G., Mills, M. J., Simpson, I. R., Glanville, A. S., Fasullo, J. T., Phillips, A. S., Lamarque, J.-F., Tribbia, J., Edwards, J., Mickelson, S. and Gosh, S.: CESM1(WACCM) Stratospheric Aerosol Geoengineering Large Ensemble (GLENS) Project, Bull. Am. Meteorol. Soc., BAMS-D-17-0267.1, doi:10.1175/BAMS-D-17-0267.1, 2018.

Yamazaki, D., Sato, T., Kanae, S., Hirabayashi, Y. and Bates, P. D.: Regional flood dynamics in a bifurcating mega delta simulated in a global river model, Geophys. Res. Lett., 41(9), 3127–3135, doi:10.1002/2014GL059744, 2014.

1 **Global streamflow and flood response to stratospheric aerosol geoengineering**

2 Liren Wei[1], Duoying Ji[1], Chiyuan Miao[2], Helene Muri[3,4], John C. Moore[1,5,6]

3 [1]College of Global Change and Earth System Science, Beijing Normal University,
4 Beijing 100875, China
5 [2]State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of
6 Geographical Science, Beijing Normal University, Beijing 100875, China
7 [3]Department of Geosciences, University of Oslo, Oslo, Norway
8 [4]Department of Energy and Process Engineering, Norwegian University of Science and
9 Technology, Trondheim, Norway
10 [5]Arctic Centre, University of Lapland, P.O. Box 122, 96101 Rovaniemi, Finland
11 [6]CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing 100101, China

12

13 *Correspondence to:* Duoying Ji (duoyingji@gmail.com) or John C. Moore

14 (john.moore.bnu@gmail.com)

15 **Abstract:**

16 Flood risk is projected to increase under future warming climates due to an enhanced

17 hydrological cycle. Solar geoengineering is known to reduce precipitation and slow

18 down the hydrological cycle, and may be therefore be expected to offset increased flood

19 risk. We examine this hypothesis using streamflow and river discharge responses to the

20 representative concentration pathway RCP4.5 and Geoengineering Model

21 Intercomparison Project (GeoMIP) G4 scenarios. Compared with RCP4.5, streamflow

22 on the western sides of Eurasia and North America are increased under G4, while the

23 eastern sides see a decrease. In the southern hemisphere, northern parts of the

24 landmasses have lower streamflow under G4, and southern parts increases relative to

25 RCP4.5. We furthermore calculate changes in 30, 50, 100-year flood return periods

26 relative to the historical (1960-1999) period under the RCP4.5 and G4 scenarios.

27 Similar spatial patterns are produced for each return period, although those under G4

28 are closer to historical values than under RCP4.5. Hence, in general, solar

29 geoengineering does appear to reduce flood risk in most regions, but the overall effects

30 are largely determined by this large-scale geographic pattern. Although G4

31 stratospheric aerosol geoengineering ameliorates the Amazon drying under RCP4.5,

32 with a weak increase in soil moisture, the decreased runoff and streamflow leads to

33 increased flood return period under G4 compared with RCP4.5.

34 **1. Introduction**

35 Floods cause considerable damage every year (UNISDR, 2013), which increases with

36 economic development and rate of climate change (Ward et al., 2017). Generally,

37 people and assets exposed to extreme hydrology disasters, including flooding, increase

38 under global warming (Alfieri et al., 2017; Arnell and Gosling, 2013; Tanoue et al.,

39 2016; Ward et al., 2013). Previous studies have shown that flood risk co-varies with

40 runoff and streamflow (Arnell and Gosling, 2013; Hirabayashi et al., 2013; Hirabayashi

41 et al., 2008). Hirabayashi et al. (2013) analyzed CMIP5 (Coupled Model

42 Intercomparison Project Phase 5) projections for the RCP4.5 and RCP8.5 scenarios

43 (Meinshausen et al., 2011), and found shortened return periods for floods, especially in

44 Southeast Asia, India and eastern Africa, especially under the RCP8.5 scenario.

45

46 Streamflow is a continuous variable and for convenience 3 quantities are commonly

47 used to measure its distribution: Q_5, the level of streamflow exceeded 5% in a year; Q_{95},

the level of streamflow exceeded 95% in a year; and Q_m the annual mean flow. Koirala et al. (2014) analyzed the changes in streamflow conditions under the different RCP scenarios. Under the RCP8.5 Q_5 increases at high latitudes, Asia and central Africa, while Q_m and Q_{95} decrease in Europe, western parts of North and central America. The spatial pattern under RCP4.5 is similar, and changes of Q_m and Q_5 streamflow are somewhat smaller than those under RCP8.5, while Q_{95} is about the same under both scenarios.

Other hydrologic indicators show similar results under future climate projections. For example, Arnell and Gosling (2013) used a global daily water balance hydrologic model (Mac-PDM.09; Gosling et al., 2010), forced by 21 climate models from the CMIP3 ensemble and analyzed 10-year and 100-year return periods of maximum daily flood under various scenarios. They found that the uncertainty in projecting river streamflow is dominated by across-model differences rather than the climate scenario. Dankers et al. (2014) used 30-year return period of 5-day average peak flows to study the changing patterns of flood hazard under the RCP8.5 scenario. They used nine global hydrology models, together with five coupled climate models from CMIP5 and showed that simulated increases in flood risk occur in Siberia, Southeast Asia and India, while decreases occur in northern and eastern Europe, and northwestern North America.

River routing models such as CaMa-Flood (Yamazaki et al., 2011) are important tools for simulating flood hazard. These models have been combined with high resolution

70   digital elevation models, flow direction maps (e.g. HYDRO1k and HydroSHEDS;

71   Lehner et al., 2008), and hydrological models. Global scale river models (GRMs) are

72   typically structured to use the gridded runoff outputs from Earth system models (ESMs),

73   land surface models (LSMs) or global hydrological models (GHMs) to simulate the

74   lateral movement of water (Trigg et al., 2016). High-resolution, offline river-routing

75   models, such as CaMa-Flood, have contributed to improved simulation of river

76   discharge (Yamazaki et al., 2009; Yamazaki et al., 2013; Mateo et al., 2017). Zhao et al.

77   (2017) used daily runoff from GHMs driving CaMa-Flood to produce monthly and

78   daily river discharge, and found that this approach results in better agreement between

79   simulated and observed discharge compared with using native hydrological model

80   routing. The CaMa-Flood model accounts for floodplain storage and backwater effects

81   that are not represented in most GHM native routing methods, and these effects play a

82   critical role in simulating peak river discharge (Yamazaki et al., 2014; Zhao et al. 2017;

83   Mateo et al., 2017). Vano et al. (2014) analyzed several sources of uncertainty in future

84   flood projections, and suggested inter-model variability in forcing from ESM are the

85   major source of uncertainty in modeling the river discharge, although the model's

86   ability to handle complex channels (e.g. deltas and floodplains) also has an important

87   impact on simulation realism.

88

89   Solar Radiation Management (SRM) is geoengineering designed to reduce the amount

90   of sunlight incident on the surface and so cool the climate. Stratospheric aerosol

91   injection is one SRM method inspired by volcanic eruptions, that utilizes the aerosol

direct effect to scatter incoming solar radiation. 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 (SO_2) per year injection to the tropical lower stratosphere, or the equivalent aerosol burden, for the period of 2020-2069. This mimics about one-fourth of the stratospheric load injected by the 1991 eruption of Mount Pinatubo. Greenhouse gas forcing is specified by the RCP4.5 scenario. Nine ESMs have done the GeoMIP G4 experiment, with sulfate aerosols handled differently by each model. For example, BNU-ESM and MIROC-ESM use the prescribed meridional distribution of aerosol optical depth (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 AOD; NorESM1-M specifies the AOD and effective radius, calculated in previous simulations with the aerosol microphysical model ECHAM5-HAM (Niemeier et al., 2011; Niemeier and Timmreck, 2015). Indirect, potentially undesirable, side-effects of the injected sulfur aerosol include changing ice particle distributions in the upper-troposphere, and the distribution of ozone and water vapor in stratospheric (Visioni et al., 2017). The direct radiative effects mainly result in the sharp reduction of the top of the atmosphere (TOA) net radiative flux with a significant drop in global surface temperature, and concomitant decrease in global precipitation (Yu et al., 2015). The decline of precipitation under SRM is mainly due to increasing atmospheric static stability, together with a reduction of latent heat flux from the land surface to the atmosphere (Bala et al., 2008; Kravitz et al., 2013b; Tilmes et al., 2013).

114 Both the reduction of latent heat flux and precipitation result in a slow-down of the

115 global hydrological cycle (Niemeier et al., 2013; Kalidindi et al., 2014; Ferraro and

116 Griffiths, 2016).

117

118 The spatial pattern of runoff roughly follows that of precipitation. Global spatially

119 continuous and temporally variable observations of runoff are not available (Ukkola et

120 al., 2018). Climate model simulated runoff is usually compared with observed

121 downstream river discharge datasets, with the dataset collected by Dai et al. (2009; 2016)

122 being the most complete. The Dai et al. (2016) dataset represents historical monthly

123 streamflow at the farthest downstream stations for the world's 925 largest ocean-

124 reaching rivers from 1900 to early 2014, lacking of global daily observations. As daily

125 runoff is largely driven by daily precipitation, it is difficult to evaluate how good the

126 runoff outputs from the climate models are at a daily scale. Over longer time scales,

127 Alkama et al. (2013) found the CMIP5 models simulate mean runoff reasonably well

128 (±25% of observed) at the global scale. The CMIP5 models tend to slightly

129 underestimate global runoff, with South American runoff underestimated by all models.

130 Koirala et al. (2014) found more CMIP5 model agreement on streamflow projections

131 under RCP8.5 than under the RCP4.5 scenario, but the projected changes in low flow

132 are robust in both scenarios with strong model agreement. Previous studies have shown

133 that under RCP4.5, precipitation would decrease over southern Africa, the Amazon

134 Basin and central America, and runoff follows these patterns. Over dry continental

135 interiors relatively large evaporation means that runoff does not follow precipitation

(Dai, 2016). SRM affects both precipitation and evaporation and hence global patterns of runoff and thence streamflow. The risk of drought in dry regions under SRM appears to be reduced (Curry et al., 2014; Keith and Irvine, 2016; Ji et al. 2018). While many studies have looked at the impact of solar geoengineering on the hydrologic cycle, none has specifically considered the potential changes of river flow and flood frequency.

We investigate the potential change of streamflow using annual mean and extreme daily discharge, and changes in the pattern of flooding using flood return period. Our study is organized as follows: Section 2 describes the models and methods used in this study; Section 3 presents the results of projected precipitation, evaporation, runoff, streamflow and return period under the G4 and RCP4.5 simulations. Section 4 provides a discussion of mechanisms for the differences between G4 and RCP4.5, and uncertainties in the study. Finally, Section 5 summarized the findings and mentions some social and economic implications from this study.

2. Data and Methods

2.1 GeoMIP experiments

To analyze the potential changes of flood under stratospheric sulfate injection geoengineering, we compare the streamflow patterns under the RCP4.5 and G4 scenarios. Five ESMs were used here due to data availability (Table 1). 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

some parts of the climate system, and analyze the precipitation, evaporation, runoff and streamflow pattern changes between each of model's G4 and RCP4.5 simulations during the period of 2030-2069. Using the last 40 years of G4 simulations is common to several previous studies (e.g. Curry et al., 2014; Ji et al., 2018). The historical simulation covering the period of 1960-1999 is used as the reference for the return period analysis. 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.

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

Model	Resolution (degrees lat × lon, level)	Number of realizations		
		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

2.2 The river routing model

The river routing model used here is the Catchment-based Macro-scale Floodplain Model (CaMa-Flood; Yamazaki et al., 2011). The CaMa-Flood uses a local inertial flow equation (Bates et al., 2010; Yamazaki et al., 2014a) to integrate runoff along a high-resolution river map (HydroSHEDS; Yamazaki et al., 2013). Sub-grid characteristics such as slope, river length, river channel width, river channel depth are parameterized in each grid box by using the innovative up-scaling method: Flexible Location of

176 Waterways (FLOW) (Mateo et al., 2017; Yamazaki et al., 2014b; Zhao et al., 2017). In

177 addition, the CaMa-Flood implements channel bifurcation and accounts for floodplain

178 storage and backwater effects, which are not represented in most global hydrological

179 models (Zhao et al., 2017). CaMa-Flood is able to reproduce relatively realistic flow

180 patterns in complex river regions, such as deltas (Ikeuchi et al., 2015; Yamazaki et al.,

181 2011, 2013). CaMa-Flood has been extensively validated and applied to many regional

182 and global scale hydrological studies (e.g. Pappenberger et al., 2012; Hirabayashi et al.,

183 2013; Mateo et al., 2014; Ikeuchi et al., 2015; Trigg et al., 2016; Zsótér et al., 2016;

184 Emerton et al., 2017; Ikeuchi et al., 2017; Suzuki et al., 2017; Yamazaki et al., 2017).

185

186 We use only the daily runoff outputs from climate models to drive CaMa-Flood v3.6.2,

187 which calculates the river discharge along the global river network. The spatial

188 resolution of CaMa-Flood is set to 0.25° (~25 km at mid-latitudes). An adaptive time

189 step scheme was applied in the model numerical integration leading to a time step of

190 about 10 minutes, while the model outputs at daily temporal resolution. To conserve the

191 input runoff mass, an area-weighted averaging method is used in CaMa-Flood to

192 distribute the coarse input to the fine resolution routing model (Mateo et al., 2017).

193 CaMa-Flood performs a 1-year spin-up before simulating 40-year river discharge in our

194 historical, RCP4.5 and G4 experiments. The runoff and river discharge from Antarctica

195 and Greenland are not included in the simulations. For each streamflow level, grid cells

196 with less than 0.01 mm/day are excluded from the analysis.

2.3 Indicators of streamflow

We analyze the streamflow change under the RCP4.5 and G4 scenarios using three streamflow indicators for the 2030-2069 period; that is annual mean flow (Q_m), and extreme high (Q_5) or low flow (Q_{95}). Q_m, Q_5 and Q_{95} are averaged over 40 years for each model respectively, then averaged between models to get the multi-model mean response under the different scenarios. We compared the multi-model mean and multi-model median responses of the five models used in this study, and found no obvious difference between the two averages.

We employ the two-sample Mann-Whitney U (MW-U) test to measure the significance of streamflow differences between G4 and RCP4.5. The MW-U test is a non-parametric test, which does not need the assumption of normal probability distributions. We use a bootstrap resampling method (Ward et al., 2016), with the MW-U test to increase sample size and to minimize the effects of outliers that can arise from the relatively short study period (Koirala et al., 2014). Specifically, we first apply the MW-U test to the G4 and RCP4.5 annual mean daily streamflow data for each model to get the value of the rank sum statistical value, U_0. Then we generate 1000 random paired series of 40-year streamflow data from RCP4.5 and G4 simulations 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, $j = 1, 2 \cdots 1000$. The rank of U_0 is then used to calculate the non-exceedance probability (Cunnane, 1978):

$$p_0 = \frac{R_0 - 0.4}{N_b + 0.2}$$

10

219   Here p_0 is the non-exceedance probability and R_0 is the rank of U_0, and N_b is the number

220   of the bootstrap samples. Finally, a non-exceedance probability less than 0.025 (or

221   greater than 0.975) indicates a significant increase (or decrease) from RCP4.5 to G4,

222   respectively.

223   **2.4 Changes in flood frequency**

224   The return period of a flood event is as an indicator of flood frequency (e.g. Dankers et

225   al., 2014; Ward et al., 2017). The N-year return period indicates the probability of flood

226   exceeding a given level in any given year of 1/N. For each model, we choose the

227   historical period of 1960-1999 as a reference for the return period calculation based on

228   the annual maximum daily river discharge. We then analyze the return period change

229   under RCP4.5 and G4 scenarios during the period of 2030-2069. In this study, we

230   choose the 30, 50 and 100-year return period levels of river flow at each grid cell to

231   study the change of flood probability. To estimate the return period, the time series of

232   annual maximum daily discharge for historical, RCP4.5 and G4 from each ESM are

233   first arranged in ascending order and then fitted to a Gumbel probability distribution.

234   The Gumbel distribution was used as a statistic of extreme flood events in previous

235   studies (e.g. Hirabayashi et al., 2013; Ward et al., 2014). Using the Gumbel distribution,

236   the cumulative distribution function, F(x), of river discharge (x) can be expressed as

237
$$F(x) = e^{-e^{-(\frac{x-b}{a})}}$$

238   where the two parameters a (scale) and b (location) are the parameters of Gumbel

239   distribution (Gumbel, 1941). The parameters are estimated using an L-moments based

11

240 approach (Rasmussen et al., 2003), where

241 $$L_1 = \frac{1}{N} \sum_{i=1}^{N} X_i$$

242 $$L_2 = \frac{2}{N} \sum_{i=1}^{N} \frac{i-1}{N-1} X_i - L_1$$

243 and X_i is the annual maximum daily river discharge and is sorted in ascending order

244 and N is the number of sample years, then:

245 $$a = \frac{L_2}{\ln 2}$$

246 $$b = L_1 - ac$$

247 where c = 0.57721 is Euler's constant. Changes in return period under SRM are

248 expressed as differences G4 - RCP4.5 relative to the corresponding historical level.

249 **3. Results**

250 **3.1 Projected changes in precipitation, evaporation and runoff**

251 G4 stratospheric aerosol geoengineering lowers net radiation fluxes at TOA by ~ 0.36

252 W m^{-2}, reduces mean global temperature by ~ 0.5 K and slows down of the global

253 hydrological cycle. Global precipitation decreases by 2.3 ± 0.5 % per Kelvin in

254 response to G4 stratospheric aerosol injection (Ji et al., 2018). Precipitation and

255 evaporation rates are strongly influenced by incoming radiation and the water vapor

256 content of the troposphere. Solar geoengineering produces changes in both atmospheric

257 circulation and thermodynamics. Several studies have analyzed changes in large scale

258 circulation under the G1 solar dimming experiment (e.g., Tilmes et al., 2009; Davis et

259 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, 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, the global annual precipitation over land (excluding Greenland and Antarctic) decreases 9.3 mm relative to the reference RCP4.5 scenario. The tropical Africa and south Asia regions suffer large precipitation reduction with values up to 37.1 mm and 52.3 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 6.9 mm in mean global land (excluding Greenland and Antarctic) evaporation relative to RCP4.5.

Under G4, there is large precipitation reduction over the Indian subcontinent and East Asia monsoon regions of 5.4% and 5.0% respectively. Under G1, these reductions have been 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. 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 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. 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

304   G4 stratospheric aerosol geoengineering.

305

306   Increased evaporation forecast under RCP4.5 is suppressed under G4 geoengineering

307   due to reduced downward surface radiation (Kravitz et al. 2013a; Yu et al., 2015).

308   Evaporation decreases over a significantly ($p<0.05$) broader area than precipitation,

309   especially in the Northern Hemisphere (Figure 1b). The change of precipitation minus

310   evaporation (P-E) basically follows the change of precipitation and evaporation, but is

311   of a smaller magnitude (Figure 1c), due to their simultaneous reductions. There are

312   significant reductions in P-E over south Asia, tropical eastern Africa and the Amazon

313   basin, and significant increases over Southern Africa and eastern Brazil. Increased P–E

314   in northern Asia caused by global warming could be partly counteracted by solar

315   geoengineering (Jones et al., 2018; Sonntag et al., 2018). The simulated precipitation

316   and evaporation changes under the G4 implies potentially significant changes in the

317   terrestrial hydrological cycle. P–E can be used as a simplified measure of runoff and

318   water availability. Under the G4 experiment, P–E increases over Europe during summer

319   time, implying more water availability and shortened return period of river discharge.

320   Soil moisture also reflects local water mass balance, i.e. the difference between P-E and

321   runoff. Soil moisture increases over Southern Africa, southwestern North America and

322   several parts of South America, where P-E and runoff both increase. The regions with

323   both significant reductions in P-E and runoff also show decreases soil moisture, such

324   as tropical Africa, south Asia and most of middle Northern America.

325

326 The spatial pattern of runoff change from RCP4.5 to G4 resembles that of P-E with a

327 broader area of significant changes (Figure 1c,1d). The annual runoff decreases by 2.4

328 mm, similar to the change in P-E. There are large runoff decreases over tropical Africa,

329 South Asia, southeastern Northern America, the Amazon basin and Alaska. Runoff

330 slightly increases over Southern Africa, southwestern North America and several

331 regions of South America. Variability in runoff and streamflow is greater than for

332 precipitation and evaporation (Figure 1, 2), due to spatial heterogeneity in soil moisture

333 and because streamflow spatially integrates runoff (Chiew and McMahon, 2002).

334

335 Precipitation, evaporation and runoff changes show that land areas dry slightly,

336 especially around the equator, south Asia and at northern high-latitudes under G4.

337 Increases in P-E are predicted on the western parts of Europe and North America, with

338 their eastern sides becoming drier with decreasing P-E and runoff.
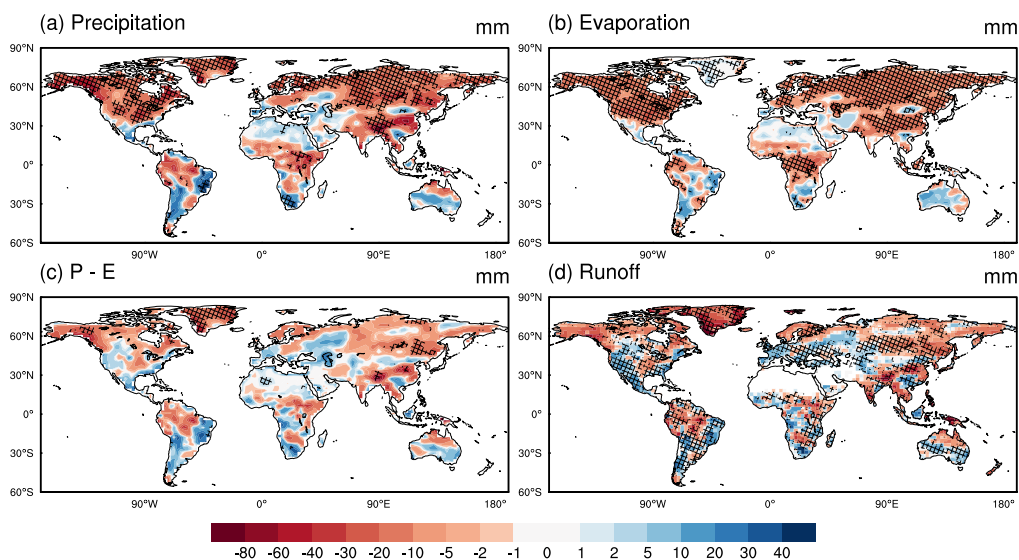
339

340 Figure 1: Changes of annual precipitation (a), evaporation (b), precipitation minus

341 evaporation (P–E, c) and runoff (d) between G4 and RCP4.5 during the period of 2030-

342 2069. Hashed areas indicate locations where the changes are significant at the 95% level

343 using the two-sample MW-U test. For runoff (d), grid cells with less than 0.01 mm/day

344 are masked out.

345 **3.2 Projected changes in streamflow**

346 Figure 2 shows the relative changes of three characteristic indicators of streamflow,

347 while Figure 3 presents the degree of across-model agreement. Figures S1-S5 show the

348 results for each of the models listed in Table 1. Figure S6-S7 show the relative changes

349 of three streamflow indicators under G4 and RCP4.5 relative to the historical period. In

350 general, the streamflow indicators under G4 are less changed from the historical levels

351 than under RCP4.5. In Fig. 2, positive values mean G4 streamflow is larger than

352 RCP4.5 levels. Generally, decreases Q_m occur at high northern latitudes such as Siberia,

353 Northern Europe and the Arctic Ocean coast of North America, along with Southeast

354 Asia, middle and southern Africa. Q_m increases in Western Europe, central Asia,

355 southwestern North America and central America (Fig. 2a). Significant changes are

356 generally distributed around the globe. Based on the ensemble response of the five

357 models analyzed here, 55% of global continental area excluding Greenland, Antarctica

358 and masked cells show decreases in Q_m under G4 compared with RCP4.5, and about

359 45% of global continental area shows increases. Figure 3 shows areas with robust

360 agreement between models and allows the primary regions affected to be seen more

361 clearly. Globally, only 21% of global continental area exhibits robust decreases and 12%

362 increases in Q_m under G4 (Fig. 3a). Despite the few grid cells with robust agreement

363   between models, the general patterns are similar for the mean changes in Fig. 2a.

364   Consistent decreases occur at high northern latitudes and in Papua New Guinea and the

365   semi-arid Sahel. Increases are mainly in the southern hemisphere, but also parts of

366   Western Europe, and the southwestern USA. The MIROC-ESM and NorESM1-M (Fig.

367   S3) contradict the ensemble in having larger areas with increases in Q_m under G4 than
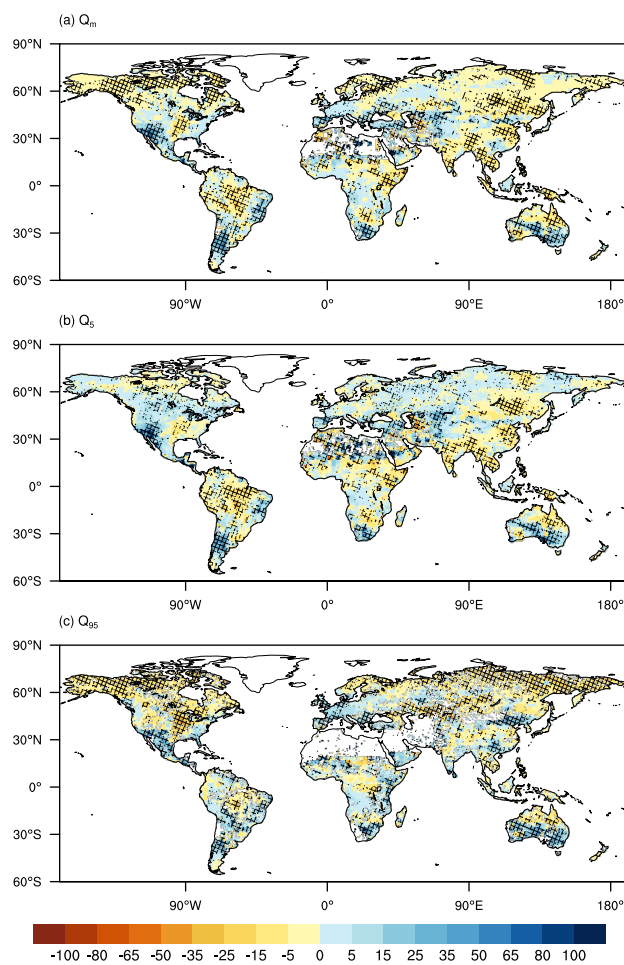
368   RCP4.5.

369

370   Figures 2b and 3b show that under G4, 52% of unmasked land area are projected to

371   increase their high flow Q_5 levels under G4. Europe, western North America, Central

372   Asia and central Australia show increases in Q_5 under G4 compared with RCP4.5.

373   Differences at the 95% significance level are distributed fairly similarly as for Q_m in

374   Figure 2a. The Amazon Basin shows decreases in both Q_5 and Q_m and southwestern

375   USA shows increases in both. Globally, 17% of unmasked land area show robust

376   increases and 17% show decreases in Q_5 under G4 (Fig. 3b). Robust increases generally

377   are confined to the extra-tropics, while decreases are mainly, but not only, in the tropics.

378   The projections of Q_5 from CanESM2 under G4 show largest differences in spatial

379   pattern from the ensemble mean (Fig. S2) and it is the only model with more decreases

380   than increases in Q_5 under G4. Though high flow levels usually correspond with flood

381   events (Ward et al., 2016), changes in flow levels do not necessarily translate into

382   increases in flood frequency. We elaborate further on flood return period in Section 3.3.

383

384   Low flow (Q_{95}, in Figs. 2c and 3c) has a noisier spatial pattern than those for mean and

385   high flow. Low flow shows a relatively uniform decrease around the globe. 49% of

386   global unmasked land area show increases in Q_{95} under G4. Despite the generally

387   noisier pattern, the regions with differences significant at the 95% level are more

388   defined for Q_{95} than either Q_m or Q_5. The high northern latitudes become drier under

389   G4, the southern high latitudes wetter. Robust increases cover about 11% of global

390   unmasked land area, mainly in Europe and South America. Robust decreases appear

391   mainly in northern high-latitude regions, central Africa and northern Asia, and occupy

392   about 20% of global unmasked land area. Projections by NorESM1-M (Fig. S5) show

393   different patterns from the ensemble mean (Fig. 2c) with bigger areas showing increases

394   than decreases in Q_{95} under G4.

395

396   Figure 2: Relative difference of three streamflow indicators between G4 and RCP4.5

397   during the period of 2030-2069, as percentages of RCP4.5: (G4-

398   RCP4.5)/RCP4.5×100%. Top, annual mean flow (Q_m); Middle, annual high flow (Q_5);

399   Bottom, annual low flow (Q_{95}). For each streamflow level, grid cells with less than 0.01

400   mm/day are masked out. Hashed areas indicate locations where the streamflow changes

401   are significant at the 95% level using the two-sample MW-U test.

402
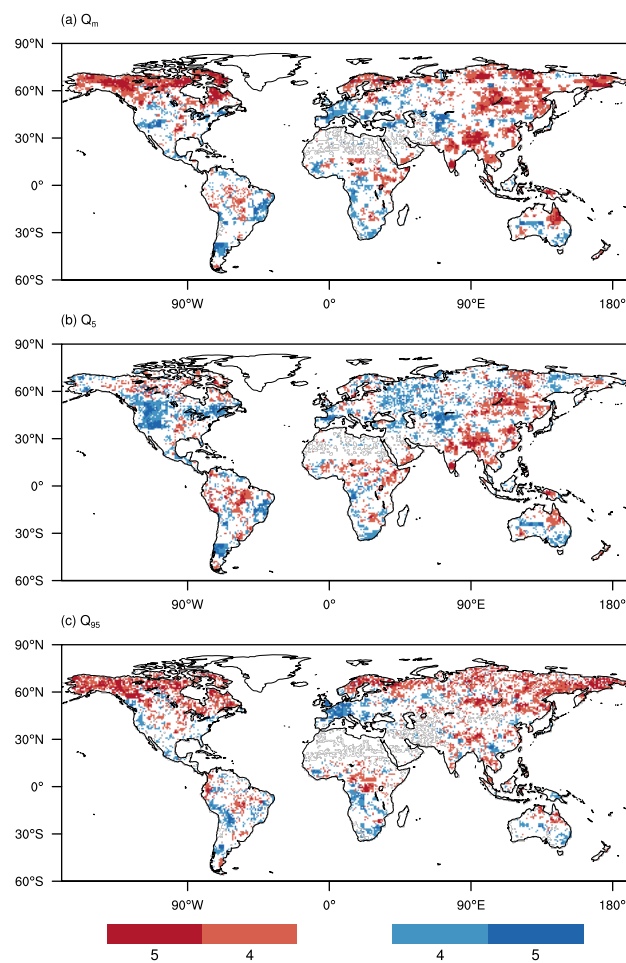
403

404   Figure 3: Number of models agreeing on sign of change (red means G4-RCP4.5<0,

405   blue means G4-RCP4.5>0) of streamflow indicator. Top, annual mean flow (Q_m);

406   Middle, annual high flow (Q_5); Bottom, annual low flow (Q_{95}). Shaded grid cells

407   indicate a relatively robust response (at least 4 models show same direction of change).

408   For each streamflow level, grid cells with less than 0.01 mm/day are masked out.

409

410   Some of the regions show contrasting responses under G4 for high and low streamflow.

411   Figure 4 shows regions where both high and low flow decrease under G4 cover about

412   30% of global unmasked land area (regions in red), mainly in eastern and southeastern

413   Asia, central Africa, and Amazon Basin, together with central and eastern Siberia. In

414   20% of global unmasked land area high flows are projected to increase while low flows

415   decrease (regions in yellow), mainly in the remaining parts of south Asia, central Africa

416   and South America. Increased high flow and simultaneous decrease in low flow

417   suggests the potential for increased flood and drought frequencies. In 21% of global

418   unmasked land area, high flows decrease and low flows increase (regions in blue),

419   which suggests these would see a decline in streamflow extremes, and are mainly at

420   northern mid- and high-latitudes. Areas with both increased high and low flow also

421   cover 29% of the unmasked land surface (regions in green), mainly in Europe, central

422   America and the southern hemisphere mid-latitudes. Perhaps the clearest overall pattern

423   is the streamflow generally increasing under G4 on the western sides of the large

424   continents of Eurasia and North America, especially over Mexico, southern California,

425   Spain and western Europe, while streamflow decreases on the eastern sides of these

426   continents. In the southern hemisphere, the pattern is meridional, with northern, wetter

427   parts of the landmasses having lower streamflow under G4, and southern, drier parts

428   increases.

429
430
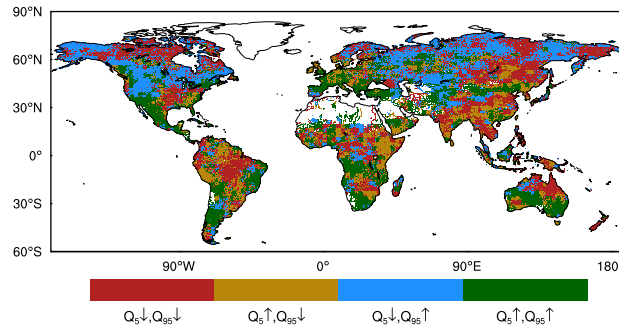
Figure 4: The ensemble mean difference (G4-RCP4.5) of high (Q_5) and low (Q_{95})

streamflow. The color bar is defined such that grid cells where G4 is less than RCP4.5

for both Q_5 and Q_{95} is in red ($Q_5\downarrow Q_{95}\downarrow$); both Q_5 and Q_{95} greater in G4 than RCP4.5 is

in green ($Q_5\uparrow Q_{95}\uparrow$); Q_5 greater in G4 and Q_{95} greater in RCP4.5 in yellow ($Q_5\uparrow Q_{95}\downarrow$)

and vice versa in blue ($Q_5\downarrow Q_{95}\uparrow$). Grid cells with Q_{95} less than 0.01 mm/day are masked

out.

3.3 Projected changes in return period

Changes in flooding between RCP4.5 and G4 scenarios are measured by the changes

in the return period of particular river discharge magnitude. Previous studies have used

30-year return period as a relatively modest indicator of flood frequency (Dankers et

al., 2014). We choose both the same flooding frequency indicator and also the more

extreme 50 or 100-year return levels. The discharge for each model's 30, 50 and 100-

year return periods in the simulated historical period define the reference magnitudes

at each grid cell. The return period of discharge corresponding to those levels are then

found under the RCP4.5 and G4 scenarios. Dry regions, defined as mean annual

streamflow during the historical period (1960-1999) less than 0.01 mm/day, are masked

447   out. The 40-year time series of the historical period (1960-1999) and 40-year future

448   projections (2030-2069) then are fitted to the Gumbel probability distribution for each

449   grid cell.

450

451   Figure 5a and 5b show the global distribution of multi-model ensemble median return

452   period of the historical 30-year return level under the RCP4.5 and G4 scenarios. Figs.

453   S8 and S9 show the relevant patterns for 50 and 100-year return periods. The elongation

454   of return period in some regions (such as central Asia and the Amazon basin) indicates

455   relatively less frequent flooding events compared with the past. Very close to half the

456   global unmasked land area (49%) show increases in return period under RCP4.5

457   scenario, while the other half experience decreases. Increases of return period are

458   mainly in Asia and eastern Africa while decreases occur in Europe and North America.

459   Our results agree with similar previous studies for RCP4.5 (e.g., Hirabayashi et al.,

460   2013). Under G4 the spatial pattern is very similar as RCP4.5, with comparable large

461   differences from the historical levels.

462

463   Figure 5c shows the difference of return period between the G4 and RCP4.5 scenarios.

464   A negative value means a shorter return period under G4 than RCP4.5, which indicates

465   an increase of flood frequency under G4. Decreasing flood frequency appears in India,

466   China, Siberia, parts of the Amazon basin, and northern Australia. Increasing flood

467   frequencies are projected mainly in Europe, the southwestern USA and much of

468   Australia. The regions which are projected to experience an increased flood frequency

469   under the RCP4.5 scenario (Fig. 5a; Dankers et al., 2014; Hirabayashi et al., 2013)

470   would experience a consistent decline of the flood frequency under G4, such as

471   southern and southeastern Asia. In general, the G4 return periods are less changed from

472   the historical levels than under RCP4.5.

473

474   Figure 6 shows the regions of robust agreement between models in changes of 30-year

475   return period under RCP4.5 and G4. Slightly fewer grid cells show robust responses

476   under G4 than RCP4.5. As with Fig. 5, there is close agreement in spatial pattern of

477   return period under the RCP4.5 and G4 scenarios. The spatial pattern of the changes in

478   50 and 100-year return levels shown in Figs. S8 and S9 are similar to those for the 30-

479   year return level (Fig. 5), while the spread between two different return period levels is

480   slightly different from the 30-year levels. These results suggest a consistent changing

481   pattern of flood frequency as defined by the three return levels, but with different

482   magnitudes of differences between RCP4.5 and G4, with G4 being closer to the

483   historical levels.

484
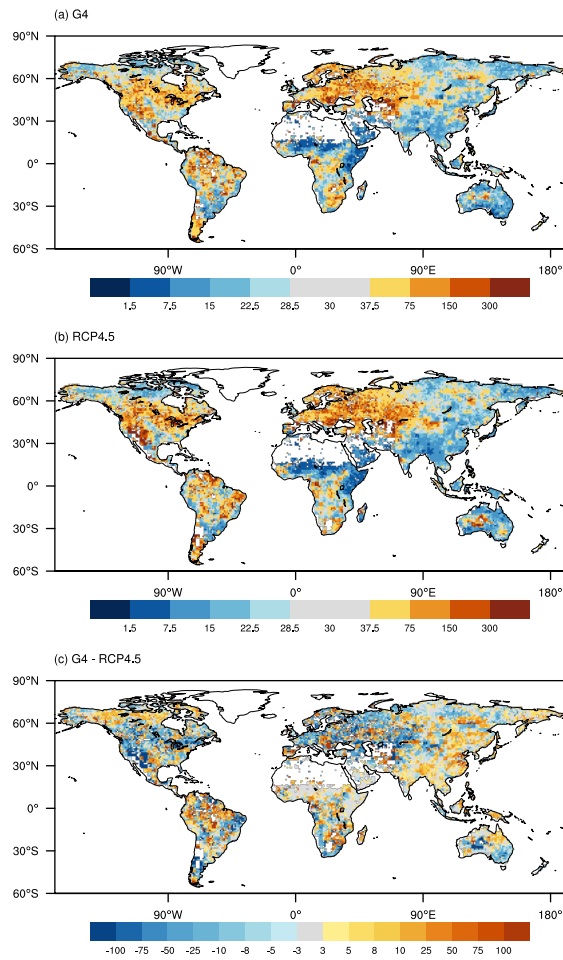
Figure 5: Multi-model ensemble median of return periods for discharge which correspond to 30-

year return period level in the historical simulation (1960-1999) under (a) G4, (b) RCP4.5 and (c)

the difference of G4 and RCP4.5. Grid cells in extremely dry regions in historical simulation, i.e.

$Q_m < 0.01$ mm/day are masked out.
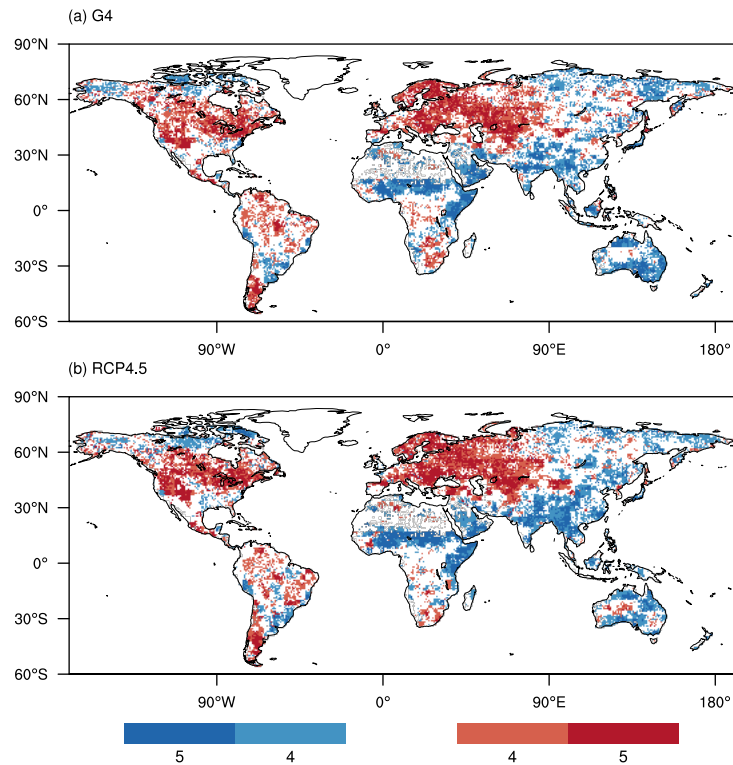
Figure 6: The number of models agreeing on the sign of change in 30-year return period under G4 (top panel) and RCP4.5 (bottom panel). Blue colors indicate decreases and red colors indicate increases relative to the historical simulation. Grid cells in extremely dry regions in historical simulation, i.e. Q_m<0.01 mm/day are masked out.

4. Discussion

4.1 G4 changes relative to RCP4.5

G4 weakens the streamflow changes expected under RCP4.5 relative to the historical period (Koirala et al., 2014). For example, in southeastern Asia and India, both high flows and low flows are projected to increase under the RCP4.5 scenario, while both of them would increase less under G4. In contrast, southern Europe is projected to see decreases in both high and low flow under RCP4.5, while the projected streamflow

501 shows less decreases under G4. However, in the Amazon basin, both high and low

502 streamflow decreases in under both RCP4.5 and G4 relative to the historical period. In

503 Siberia both high and low streamflow increases under RCP4.5 relative to historical,

504 while the pattern is mixed under G4. This means that G4 offsets the impact introduced

505 by anthropogenic climate warming in some regions, while in other regions such as the

506 Amazon basin and Siberia, it further enhances the decreasing trend of streamflow under

507 the RCP4.5 scenario. The pattern seen is suggestive of the role of large-scale circulation

508 patterns (Fig. 7), westerly flows over the northern hemisphere continents and the Asian

509 monsoon systems, with relative increases in mid-latitude storm systems and decreases

510 in monsoons under G4 compared with RCP4.5. These circulation changes result in, for

511 example, more moist maritime air flowing into the Mediterranean region, and weakened

512 summertime monsoonal circulation under G4 in India and East Asia (Fig. 7 e,f). Similar

513 mechanisms may also account for the north-south pattern seen in Australia and South

514 America. Monsoonal indicators do decrease under the much more extreme G1

515 experiment, in which solar dimming is designed to offset quadrupled CO_2 levels

516 (Tilmes et al., 2013).

517

518
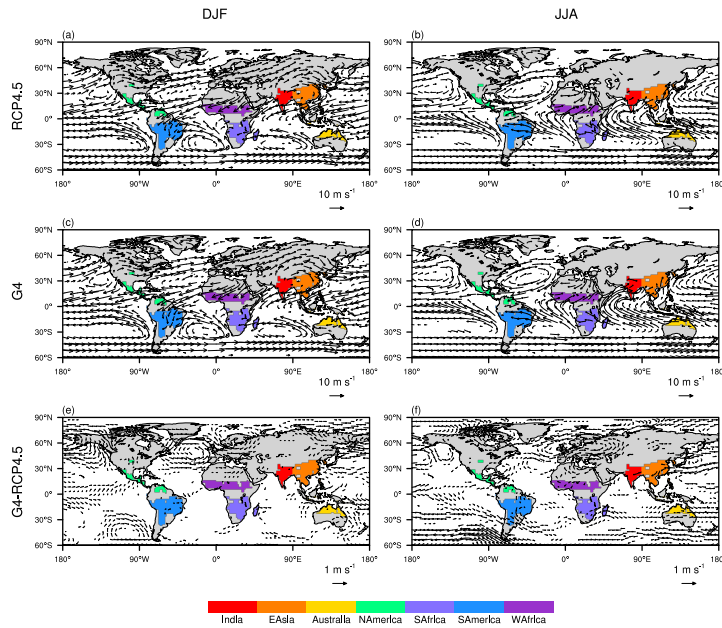
Figure 7: Multi-model ensemble mean of 925hPa wind field during December-January-February (DJF) and June-July-August (JJA) seasons. Panel (a) and (b) for RCP4.5, panel (c) and (d) for G4, panel (e) and (f) for the difference between G4 and RCP4.5. Grid cells where wind speed less than 2.0 m s^{-1} are masked out in panel (a), (b), (c) and (d), grids cells where wind speed less than 0.1 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 (Adler et al., 2003).

There is a latitudinal dependence for streamflow: generally, the Q_m decreases across all latitudes; high flow, Q_5, decreases most in tropical regions; low flow, Q_{95}, decreases most at high-latitudes. The high-latitudes display a complicated streamflow pattern with weakly increasing Q_5 and significant decreasing Q_{95}. The decrease in the lower probability tail of streamflow is indicative of hydrological droughts, while the increases

28

533   in the high streamflow tail indicates hydrological flooding (Keyantash and Dracup,

534   2002). Previous studies (Dankers et al., 2014; Hirabayashi et al., 2008) have noted that

535   the flood frequency for rivers at high latitude (e.g. Alaska and Siberia) decreases under

536   global warming, even in areas where the frequency, intensity of precipitation, or both,

537   are projected to increase. The annual hydrograph of these rivers is dominated by snow

538   melt, so changes of peak flow reflect the balance between length and temperature of

539   winter season, and the total amount of winter precipitation. The thawing of permafrost

540   and changes in evapotranspiration also play an important role in the increasing of runoff

541   and streamflow (Dai, 2016). The combined effect of atmospheric circulation and land

542   surface processes results in the complex change pattern in this cold region.

543

544   Under the G4 experiment, recent studies (Jones et al., 2018; Sonntag et al., 2018) have

545   pointed out that the increased P–E in northern Asia caused by global warming could be

546   partly counteracted by solar geoengineering. At the same time, solar geoengineering

547   reduces polar temperatures and precipitation (Berdahl et al., 2014; Ji et al., 2018). The

548   balance among precipitation, evaporation and temperature accounts for the complex

549   spatial pattern of streamflow and flood frequency under solar geoengineering, that has

550   been previously related to soil moisture content (Dagon and Schrag, 2017). It is worth

551   noting that the method for calculating potential evapotranspiration (ET) plays a

552   significant role in determining simulated surface runoff changes (Haddeland et al., 2011;

553   Thompson et al., 2013), which would influence the condition of streamflow. A recent

554   study (Wartenburger et al., 2018) compared the ET spatial and temporal patterns

simulated by GHMs in the second phase of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2a) which also confirmed that the ET scheme used affects model ensemble variance. The ET in this study is calculated by the ESMs (Table 1), not GHMs, and any biases in ET would feed into streamflow. For example, Mueller and Seneviratne (2014) found that climate models which participated in CMIP5 display an overall systematic overestimation of annual average ET over most regions, particularly in Europe, Africa, China, Australia, Western North America, and part of the Amazon region.

The relatively drier streamflow pattern in the Amazon basin under G4 is notable and consistent with changes in P-E (e.g. Jones et al., 2018). This drying pattern would increase the risk of a decline of the Amazon tropical rainforest (Boisier et al., 2015). 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 such changes occurring under SRM (Gabriel and Robock, 2015); and iii) changes to 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; Muri et al., 2018).

4.2 Uncertainties

Previous studies suggest that the river routing model CaMa-Flood can realistically reproduce peak river discharge because of the floodplain storage and backwater effects are implemented (e.g. Zhao et al., 2017). In this study, the CaMa-Flood is driven by the runoff output directly from ESMs to simulate streamflow and flood response. Therefore, the uncertainty in runoff from the ESMs is also important. To drive the high-resolution CaMa-Flood model, the coarse resolution runoff from ESMs 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 the 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

31

599   natural processes which could change river routes, and river network silt-up over time,

600   these changes would impact local runoff and streamflow (Chezik et al., 2017), and we

601   do not account for them in this study.

602

603   Gosling et al. (2017) compared the river runoff output from multiple global and

604   catchment-scale hydrological model under three warming scenarios simulated by ESMs

605   finding that the across-model uncertainty overwhelmed the ensemble median

606   differences between the scenarios. Yu et al. (2016) suggested model internal variability

607   may be larger than across-model spread in eastern and southeastern Asia. In this study

608   we use the offline hydrological model driven by runoff outputs from ESMs to calculate

609   the streamflow, the uncertainty between ESMs is reflected in the range of return period

610   based on streamflow change. Figure S10 shows the multi-model ensemble range of the

611   30-year return period level. Regions that have the shorter return period (i.e. higher flood

612   frequency) from historical to future, show a relatively small range among models (e.g.

613   India and Southeastern Asia). Regions that have the longer return period show a large

614   range (e.g. Europe and North America). This reflects larger inter-model uncertainty

615   over dry zones than for wetter ones. The return period change over dry zones is more

616   meaningful when interpreted as the change of drought tendency. 50- and 100-year

617   return period levels flow show larger uncertainty than 30-year return period level,

618   which is expected when estimating the low probability extreme tails of the flow

619   probability density function from relatively short (40 year) sets of results.

620

5. Summary and Implications

We analyzed the streamflow response under the stratospheric aerosol injection geoengineering, G4, and the RCP4.5 scenario using the daily total runoff from five climate models that participated in GeoMIP. We investigated the mean change patterns of annual mean, extreme high and low streamflow, and analyzed the global flood frequency change in terms of return period. There is pattern of generally increasing streamflow under G4 on the western sides of the major continents of Eurasia and North America, with decreasing streamflow on their eastern sides. In the southern hemisphere, the pattern is meridional, with northern parts of the landmasses having lower streamflow under G4, and southern parts increases. We further investigated the change of flooding corresponding to the magnitudes of the historical 30, 50 and 100-year return period levels; the flooding frequencies change dramatically from historical levels under both RCP4.5 and G4, and show similar spatial patterns. The projected return period pattern under RCP4.5 scenario agrees well with previous studies, such as Dankers et al. (2014) and Hirabayashi et al. (2013). Generally, stratospheric aerosol injection geoengineering as simulated by G4 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. The Amazon Basin shows a relative elongation of flood return period, while Europe shows shortening of return period under G4, and this was also implicit in streamflow characteristics in these regions.

643 CaMa-Flood does not consider anthropogenic infrastructure, such as dams or reservoirs,

644 which some hydrological models do include. However, estimating future changes in

645 human intervention on the natural system is highly uncertain. Technological advances

646 over the century that may affect anthropogenic changes are by their nature entirely

647 unknown at present. Hence integrating the human dimension into a model of the

648 physical system is fraught with difficulty and uncertainty. Several studies can be used

649 as a guide to the possible effects of anthropogenic impacts compared with natural

650 changes that are captured in CaMa-Flood. Dai et al. (2009) argued that the direct human

651 influence on the major global river streamflow is relatively small compared with

652 climate forcing during the historical period. Mateo et al. (2014) suggested that dams

653 regulate streamflow consistently in a basin study using CaMa-Flood combined with

654 integrated water resources and reservoir operation models. Wang et al. (2017) shows

655 that the reservoir would effectively suppress the flood magnitude and frequency.

656 Recently, analyses of the role of human impact parameterizations (HIP) in five

657 hydrological models found that the inclusion of HIP improves the performance of

658 GHMs, both in managed and near-natural catchments, and simulates fewer hydrological

659 extremes by decreasing the simulated high-flows (Veldkamp et al., 2018; Zaherpour et

660 al., 2018). These studies suggest that the high-flows and flood response under G4

661 relative to RCP4.5 might be smaller when human intervention is considered, and

662 indicate the importance of considering human impacts in future hydrological response

663 studies under geoengineering.

664

665   The accurate assessment of human impacts on flood frequency and magnitude depends

666   not only on how anthropogenic effects are parameterized in hydrological models

667   (Masaki et al., 2017), but also on how human activities are represented in

668   geoengineering scenarios. As anthropogenic GHG emissions increase, human society

669   would continually adapt to climate change and mitigate the related risk, including

670   building new dams and reservoirs to withstand a strengthened global hydrological cycle.

671   How society would response to future streamflow and flood risk is an important topic

672   both scientifically and in policy making. This is especially true for the developing world,

673   where many cities are experiencing subsidence due to unsustainable rates of ground

674   water extraction. Subsidence accounted for up to 1/3 of 20th century relative sea level

675   rise in around China (Chen, 1991; Ren, 1993). Subsidence and sea level rise both

676   increase flooding risks. However, in densely populated regions with long experience of

677   irrigation management, such as Southeast Asia and India, reduced flood frequency

678   under G4 stratospheric aerosol geoengineering might be further ameliorated.

679

680   Our results on streamflow and flood response are based on GeoMIP G4 simulation and

681   its reference RCP4.5 simulation. The generalizations of the work to other types and

682   extents of solar geoengineering depends on the linearity of the streamflow response to

683   both greenhouse gas and geoengineering. The linearity of response of radiative forcing

684   and global temperatures in particular have been explored in CESM1 stratospheric

685   aerosol Geoengineering Large Ensemble (GLENS, Tilmes et al., 2018). Many climate

686   fields, such as temperature, are surprisingly linear under a very wide range of forcing,

687   potentially allowing standard engineering control theory methods (e.g. MacMartin et

688   al., 2014) to tailor a global response given the freedom to use different latitudinal input

689   locations for the aerosol injection (MacMartin et al., 2018; Kravitz et al., 2018), or

690   combinations of, for example aerosol injection and marine cloud brightening (Cao et

691   al., 2017). Non-linearities are expected for systems that depend on ice/water phase

692   changes, and these could affect global streamflow and flood responses in some regions,

693   especially in the Arctic. Moreover, the type of solar geoengineering might be relevant

694   as well. Ferraro et al. (2014) found that the tropical overturning circulation weakens in

695   response to geoengineering with stratospheric sulfate aerosol injection due to radiative

696   heating from the aerosol layer, but geoengineering simulated as a simple reduction in

697   total solar irradiance does not capture this effect. A larger tropical precipitation

698   perturbation occurs under equatorial injection scenarios (such as G4) than under simple

699   solar dimming geoengineering, or the latitudinal varying injections schemes explored

700   by GLENS, or a mix of different geoengineering strategies (such as aerosol injection

701   and marine cloud brightening, Cao et al., 2017). So the response of streamflow and

702   flood would be expected to differ, to some extent, under different types of solar

703   geoengineering.

704

705   Floods are among the most costly natural disasters around the world, especially for

706   more vulnerable developing countries (e.g. Bangladesh, India and China). Our study

707   suggests that solar geoengineering would exert non-uniform impacts on global flooding

708   risk and hence local hydraulic infrastructure needs would vary if solar geoengineering

of the G4-type were undertaken. Changes in flooding are strongly connected with the economic cost of damage due to climate change and sea level rise (Jevrejeva et al., 2016; Hinkel et al., 2014) and thorough studies should be made for further policy and decision-making, especially applied to high value economic or ecological entities. This may be done in the framework of specific impact models applied to local cities or regions, and would hence benefit from local knowledge, especially in the developing world where resources for adaptation measures are scarce. Linkages between the developing world climate impacts researchers and the GeoMIP community will be encouraged and funded by the Developing Country Impacts Modelling Analysis for SRM (DECIMALS) project (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.

Acknowledgments

732

References

734 Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R. and Seibert, J.: Robust

735 changes and sources of uncertainty in the projected hydrological regimes of Swiss

736 catchments, Water Resour. Res., 50(10), 7541–7562, doi:10.1002/2014WR015549,

737 2014.

738 Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B.,

739 Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P. and Nelkin, E.:

740 The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation

741 Analysis (1979–Present), J. Hydrometeorol., 4(6), 1147–1167, doi:10.1175/1525-

742 7541(2003)004<1147:TVGPCP>2.0.CO;2, 2003.

743 Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K.

744 and Feyen, L.: Global projections of river flood risk in a warmer world, Earth's Futur.,

745 5(2), 171–182, doi:10.1002/2016EF000485, 2017.

746 Alkama, R., Marchand, L., Ribes, A. and Decharme, B.: Detection of global runoff

747 changes: Results from observations and CMIP5 experiments, Hydrol. Earth Syst. Sci.,

748 17(7), 2967–2979, doi:10.5194/hess-17-2967-2013, 2013.

749 Arnell, N. W. and Gosling, S. N.: The impacts of climate change on river flow regimes

750 at the global scale, J. Hydrol., 486, 351–364, doi:10.1016/j.jhydrol.2013.02.010, 2013.

751 Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M.,

752 Kharin, V. V, Lee, W. G. and Merryfield, W. J.: Carbon emission limits required to

753 satisfy future representative concentration pathways of greenhouse gases, Geophys.

754   Res. Lett., 38(5), 2011.

755   Bala, G., Duffy, P. B. and Taylor, K. E.: Impact of geoengineering schemes on the global

756   hydrological cycle., Proc. Natl. Acad. Sci. U. S. A., 105(22), 7664–9,

757   doi:10.1073/pnas.0711648105, 2008.

758   Bates, P. D., Horritt, M. S. and Fewtrell, T. J.: A simple inertial formulation of the

759   shallow water equations for efficient two-dimensional flood inundation modelling, J.

760   Hydrol., 387(1–2), 33–45, doi:10.1016/j.jhydrol.2010.03.027, 2010.

761   Bentsen, M., Bethke, I., Debernard, J. B., Iversen, T., Kirkevåg, A., Seland, Ø., Drange,

762   H., Roelandt, C., Seierstad, I. A., Hoose, C. and Kristjánsson, J. E.: The Norwegian

763   Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the

764   physical climate, Geosci. Model Dev., 6(3), 687–720, doi:10.5194/gmd-6-687-2013,

765   2013.

766   Berdahl, M., Robock, A., Ji, D., Moore, J. C., Jones, A., Kravitz, B. and Watanabe, S.:

767   Arctic cryosphere response in the Geoengineering Model Intercomparison Project G3

768   and G4 scenarios, J. Geophys. Res. Atmos., 119(3), 1308–1321,

769   doi:10.1002/2013JD020627, 2014.

770   Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B., Aronica, G.

771   T., Bilibashi, A., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B.,

772   Claps, P., Fiala, K., Frolova, N., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S.,

773   Kireeva, M., Kiss, A., Kjeldsen, T. R., Kohnová, S., Koskela, J. J., Ledvinka, O.,

774   Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P.,

775   Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Rogger, M., Salinas,

J. L., Sauquet, E., Šraj, M., Szolgay, J., Viglione, A., Volpi, E., Wilson, D., Zaimi, K. and Živković, N.: Changing climate shifts timing of European floods, Science (80-.)., 357(6351), 588–590, doi:10.1126/science.aan2506, 2017.

Boisier, J. P., Ciais, P., Ducharne, A. and Guimberteau, M.: Projected strengthening of Amazonian dry season by constrained climate model simulations, Nat. Clim. Chang., 5(7), 656–660, doi:10.1038/nclimate2658, 2015.

Cao, L., Duan, L., Bala, G. and Caldeira, K.: Simultaneous stabilization of global temperature and precipitation through cocktail geoengineering, Geophys. Res. Lett., 44(14), 7429–7437, doi:10.1002/2017GL074281, 2017.

Chadwick, R., Douville, H. and Skinner, C. B.: Timeslice experiments for understanding regional climate projections: applications to the tropical hydrological cycle and European winter circulation, Clim. Dyn., 49(9–10), 3011–3029, doi:10.1007/s00382-016-3488-6, 2017.

Chen, J., Brissette, F. P., Poulin, A. and Leconte, R.: Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed, Water Resour. Res., 47(12), 1–16, doi:10.1029/2011WR010602, 2011.

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.

Chezik, K. A., Anderson, S. C. and Moore, J. W.: River networks dampen long-term hydrological signals of climate change, Geophys. Res. Lett., doi:10.1002/2017GL074376, 2017.

Chiew, F. H. S. and McMahon, T. A.: Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability, Hydrol. Sci. J., 47(3), 505–522, doi:10.1080/02626660209492950, 2002.

Chou, C., Chiang, J. C. H., Lan, C. W., Chung, C. H., Liao, Y. C. and Lee, C. J.: Increase in the range between wet and dry season precipitation, Nat. Geosci., 6(4), 263–267, doi:10.1038/ngeo1744, 2013.

Chylek, P., Li, J., Dubey, M. K., Wang, M. and Lesins, G.: Observed and model simulated 20th century Arctic temperature variability: Canadian Earth System Model CanESM2, Atmos. Chem. Phys. Discuss., 11(8), 22893–22907, doi:10.5194/acpd-11-22893-2011, 2011.

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., Fowler, H. J., Prudhomme, C., Arnold, J. R. and Brekke, L. D.: Characterizing Uncertainty of the Hydrologic Impacts of Climate Change, Curr. Clim. Chang. Reports, 2(2), 55–64, doi:10.1007/s40641-016-0034-x, 2016.

Cunnane, C.: Unbiased Plotting Position - A Review, J. Hydrol., 37(3), 205–222, 1978.

Curry, C. L., Sillmann, J., Bronaugh, D., Alterskjaer, K., Cole, J. N. S., Ji, D., Kravitz, B., Kristjánsson, J. E., Moore, J. C., Muri, H., Niemeier, U., Robock, A., Tilmes, S. and Yang, S.: A multimodel examination of climate extremes in an idealized geoengineering experiment, J. Geophys. Res. G Biogeosciences, 119(7), 3900–3923, doi:10.1002/2013JD020648, 2014.

Dai, A., Qian, T., Trenberth, K. E. and Milliman, J. D.: Changes in continental freshwater discharge from 1948 to 2004, J. Clim., 22(10), 2773–2792,

820   doi:10.1175/2008JCLI2592.1, 2009.

821   Dai, A.: Historical and Future Changes in Streamflow and Continental Runoff: A

822   Review, Terr. Water Cycle Clim. Chang. Nat. Human-Induced Impacts, 17–37,

823   doi:10.1002/9781118971772.ch2, 2016.

824   Dagon, K. and Schrag, D. P.: Regional Climate Variability Under Model Simulations

825   of Solar Geoengineering, J. Geophys. Res. Atmos., 1–16, doi:10.1002/2017JD027110,

826   2017.

827   Dankers, R., Arnell, N. W., Clark, D. B., Falloon, P. D., Fekete, B. M., Gosling, S. N.,

828   Heinke, J., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D.: First

829   look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison

830   Project   ensemble,   Proc.   Natl.   Acad.   Sci.,   111(9),   3257–3261,

831   doi:10.1073/pnas.1302078110, 2014.

832   Davis, N. A., Seidel, D. J., Birner, T., Davis, S. M. and Tilmes, S.: Changes in the width

833   of the tropical belt due to simple radiative forcing changes in the GeoMIP simulations,

834   Atmos. Chem. Phys., 16(15), 10083–10095, doi:10.5194/acp-16-10083-2016, 2016.

835   Ehsani, N., Vörösmarty, C. J., Fekete, B. M. and Stakhiv, E. Z.: Reservoir Operations

836   Under Climate Change: Storage Capacity Options to Mitigate Risk, J. Hydrol., 555,

837   435–446, doi:10.1016/j.jhydrol.2017.09.008, 2017.

838   Emerton, R., Cloke, H. L., Stephens, E. M., Zsoter, E., Woolnough, S. J. and

839   Pappenberger, F.: Complex picture for likelihood of ENSO-driven flood hazard, Nat.

840   Commun., 8, 14796, doi:10.1038/ncomms14796, 2017.

841   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(1), doi:10.1088/1748-9326/9/1/014001, 2014.

Ferraro, A. J., Charlton-Perez, A. J. and Highwood, E. J.: Stratospheric dynamics and midlatitude jets under geoengineering with space mirrors and sulfate and titania aerosols, J. Geophys. Res. Atmos., 120(2), 414–429, doi:10.1002/2014JD022734, 2015.

Ferraro, A. J. and Griffiths, H. G.: Quantifying the temperature-independent effect of stratospheric aerosol geoengineering on global-mean precipitation in a multi-model ensemble, Environ. Res. Lett., 11(3), 34012, doi:10.1088/1748-9326/11/3/034012, 2016.

Gabriel, C. J. and Robock, A.: Stratospheric geoengineering impacts on El Niño/Southern Oscillation, Atmos. Chem. Phys., 15(6), 9173–9202, 2015.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z. L. and Zhang, M.: The community climate system model version 4, J. Clim., 24(19), 4973–4991, doi:10.1175/2011JCLI4083.1, 2011.

Gosling, S. N., Bretherton, D., Haines, K. and Arnell, N. W.: Global hydrology modelling and uncertainty: running multiple ensembles with a campus grid, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 368(1926), 4005–4021, doi:10.1098/rsta.2010.0164, 2010.

Gosling, S. N., Zaherpour, J., Mount, N. J., Hattermann, F. F., Dankers, R., Arheimer, B., Breuer, L., Ding, J., Haddeland, I., Kumar, R., Kundu, D., Liu, J., van Griensven, A., Veldkamp, T. I. E., Vetter, T., Wang, X. and Zhang, X.: A comparison of changes in

864 river runoff from multiple global and catchment-scale hydrological models under

865 global warming scenarios of 1 °C, 2 °C and 3 °C, Clim. Change, 141(3), 577–595,

866 doi:10.1007/s10584-016-1773-3, 2017.

867 Gumbel, E. J.: The Return Period of Flood Flows, Ann. Math. Stat., 12(2), 163–190,

868 doi:10.1214/aoms/1177731747, 1941.

869 Guo, A., Moore, J. C. and Ji, D.: Tropical atmospheric circulation response to the G1

870 sunshade geoengineering radiative forcing experiment, Atmos. Chem. Phys., 18(12),

871 8689–8706, doi:10.5194/acp-18-8689-2018, 2018.

872 Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand,

873 N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S.,

874 Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke,

875 T., Viterbo, P., Weedon, G. P. and Yeh, P.: Multimodel Estimate of the Global Terrestrial

876 Water Balance: Setup and First Results, J. Hydrometeorol., 12(5), 869–884,

877 doi:10.1175/2011JHM1324.1, 2011.

878 Halladay, K. and Good, P.: Non-linear interactions between CO2 radiative and

879 physiological effects on Amazonian evapotranspiration in an Earth system model, Clim.

880 Dyn., 49(7–8), 2471–2490, doi:10.1007/s00382-016-3449-0, 2017.

881 Harris, P. P., Huntingford, C. and Cox, P. M.: Amazon Basin climate under global

882 warming: the role of the sea surface temperature, Philos. Trans. R. Soc. B Biol. Sci.,

883 363(1498), 1753–1759, doi:10.1098/rstb.2007.0037, 2008.

884 Hattermann, F. F., Krysanova, V., Gosling, S. N., Dankers, R., Daggupati, P., Donnelly,

885 C., Flörke, M., Huang, S., Motovilov, Y., Buda, S., Yang, T., Müller, C., Leng, G., Tang,

886　Q., Portmann, F. T., Hagemann, S., Gerten, D., Wada, Y., Masaki, Y., Alemayehu, T.,

887　Satoh, Y. and Samaniego, L.: Cross-scale intercomparison of climate change impacts

888　simulated by regional and global hydrological models in eleven large river basins, Clim.

889　Change, 141(3), 561–576, doi:10.1007/s10584-016-1829-4, 2017.

890　Haywood, J. M., Jones, A., Bellouin, N. and Stephenson, D.: Asymmetric forcing from

891　stratospheric aerosols impacts Sahelian rainfall, Nat. Clim. Chang., 3(7), 660–665,

892　doi:10.1038/nclimate1857, 2013.

893　Hinkel, J., Lincke, D., Vafeidis, A. T., Perrette, M., Nicholls, R. J., Tol, R. S. J.,

894　Marzeion, B., Fettweis, X., Ionescu, C. and Levermannet, A.: Coastal flood damage

895　and adaptation costs under 21st century sea-level rise, Proc. Natl. Acad. Sci. USA, 111,

896　3292–3297, 2014.

897　Hirabayashi, Y., Kanae, S., Emori, S., Oki, T. and Kimoto, M.: Global projections of

898　changing risks of floods and droughts in a changing climate, Hydrol. Sci. J., 53(4), 754–

899　772, doi:10.1623/hysj.53.4.754, 2008.

900　Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe,

901　S., Kim, H. and Kanae, S.: Global flood risk under climate change, Nat. Clim. Chang.,

902　3(9), 816–821, doi:10.1038/nclimate1911, 2013.

903　Ikeuchi, H., Hirabayashi, Y., Yamazaki, D., Kiguchi, M., Koirala, S., Nagano, T., Kotera,

904　A. and Kanae, S.: Modeling complex flow dynamics of fluvial floods exacerbated by

905　sea level rise in the Ganges–Brahmaputra–Meghna Delta, Environ. Res. Lett., 10(12),

906　124011, doi:10.1088/1748-9326/10/12/124011, 2015.

907　Ikeuchi, H., Hirabayashi, Y., Yamazaki, D., Muis, S., Ward, P. J., Winsemius, H. C.,

908   Verlaan, M. and Kanae, S.: Compound simulation of fluvial floods and storm surges in

909   a global coupled river-coast flood model: Model development and its application to

910   2007 Cyclone Sidr in Bangladesh, J. Adv. Model. Earth Syst., 9(4), 1847–1862,

911   doi:10.1002/2017MS000943, 2017.

912   Jevrejeva, S., Jackson, L. P., Riva, R. E. M., Grinsted, A. and Moore, J. C.: Coastal sea

913   level rise with warming above 2°C, Proc. Natl Acad. Sci. USA, 113, 13342–13347,

914   2016.

915   Ji, D., Wang, L., Feng, J., Wu, Q., Cheng, H., Zhang, Q., Yang, J., Dong, W., Dai, Y.,

916   Gong, D., Zhang, R. H., Wang, X., Liu, J., Moore, J. C., Chen, D. and Zhou, M.:

917   Description and basic evaluation of Beijing Normal University Earth System Model

918   (BNU-ESM) version 1, Geosci. Model Dev., 7(5), 2039–2064, doi:10.5194/gmd-7-

919   2039-2014, 2014.

920   Ji, D., Fang, S., Curry, C. L., Kashimura, H., Watanabe, S., Cole, J., Lenton, A., Muri,

921   H., Kravitz, B. and Moore, J. C.: Extreme temperature and precipitation response to

922   solar dimming and stratospheric aerosol geoengineering, Atmos. Chem. Phys., 18(14),

923   10133–10156, doi:10.5194/acp-18-10133-2018, 2018.

924   Jiménez-Muñoz, J. C., Mattar, C., Barichivich, J., Santamaría-Artigas, A., Takahashi,

925   K., Malhi, Y., Sobrino, J. A. and Schrier, G. Van Der: Record-breaking warming and

926   extreme drought in the Amazon rainforest during the course of El Niño 2015-2016, Sci.

927   Rep., 6(September), 1–7, doi:10.1038/srep33130, 2016.

928   Jones, A. C., Hawcroft, M. K., Haywood, J. M., Jones, A., Guo, X. and Moore, J. C.:

929   Regional climate impacts of stabilizing global warming at 1.5 K using solar

930   geoengineering, Earth's Futur., 1–22, doi:10.1002/2017EF000720, 2018.

931   Kalidindi, S., Bala, G., Modak, A. and Caldeira, K.: Modeling of solar radiation

932   management: a comparison of simulations using reduced solar constant and

933   stratospheric sulphate aerosols, Clim. Dyn., 44(9–10), 2909–2925,

934   doi:10.1007/s00382-014-2240-3, 2014.

935   Kashimura, H., Abe, M., Watanabe, S., Sekiya, T., Ji, D., Moore, J. C., Cole, J. N. S.

936   and Kravitz, B.: Shortwave radiative forcing, rapid adjustment, and feedback to the

937   surface by sulfate geoengineering: Analysis of the Geoengineering Model

938   Intercomparison Project G4 scenario, Atmos. Chem. Phys., 17(5), 3339–3356,

939   doi:10.5194/acp-17-3339-2017, 2017.

940   Keith, D. W. and Irvine, P. J.: Solar geoengineering could substantially reduce climate

941   risks—A research hypothesis for the next decade, Earth's Futur., 4(11), 549–559,

942   doi:10.1002/2016EF000465, 2016.

943   Keyantash, J. and Dracup, J. A.: The quantification of drought: an evaluation of drought

944   indices, Bull. Am. Meteorol. Soc., 83(8), 1167–1180, 2002.

945   Koirala, S., Hirabayashi, Y., Mahendran, R. and Kanae, S.: Global assessment of

946   agreement among streamflow projections using CMIP5 model outputs, Environ. Res.

947   Lett., 9(6), 64017, doi:10.1088/1748-9326/9/6/064017, 2014.

948   Kravitz, B., Robock, A., Boucher, O., Schmidt, H., Taylor, K. E., Stenchikov, G. and

949   Schulz, M.: The Geoengineering Model Intercomparison Project (GeoMIP), Atmos. Sci.

950   Lett., 12(2), 162–167, doi:10.1002/asl.316, 2011.

951   Kravitz, B., Robock, A. and Haywood, J. M.: Progress in climate model simulations of

952 geoengineering, Eos (Washington. DC)., 93(35), 340, doi:10.1029/2012EO350009,

953 2012.

954 Kravitz, B., Robock, A., Forster, P. M., Haywood, J. M., Lawrence, M. G. and Schmidt,

955 H.: An overview of the Geoengineering Model Intercomparison Project (GeoMIP), J.

956 Geophys. Res. Atmos., 118(23), 13103–13107, doi:10.1002/2013JD020569, 2013a.

957 Kravitz, B., Caldeira, K., Boucher, O., Robock, A., Rasch, P. J., Alterskjær, K., Karam,

958 D. B., Cole, J. N. S., Curry, C. L., Haywood, J. M., Irvine, P. J., Ji, D., Jones, A.,

959 Kristjánsson, J. E., Lunt, D. J., Moore, J. C., Niemeier, U., Schmidt, H., Schulz, M.,

960 Singh, B., Tilmes, S., Watanabe, S., Yang, S. and Yoon, J. H.: Climate model response

961 from the Geoengineering Model Intercomparison Project (GeoMIP), J. Geophys. Res.

962 Atmos., 118(15), 8320–8332, doi:10.1002/jgrd.50646, 2013b.

963 MacMartin, D. G., Ricke, K. L. and Keith, D. W.: Solar geoengineering as part of an

964 overall strategy for meeting the 1.5°C Paris target, Philos. Trans. R. Soc. A Math. Phys.

965 Eng. Sci., 376(2119), doi:10.1098/rsta.2016.0454, 2018.

966 May, W., Rummukainen, M., Chéruy, F., Hagemann, S. and Meier, A.: Contributions of

967 soil moisture interactions to future precipitation changes in the GLACE-CMIP5

968 experiment, Clim. Dyn., 49(5–6), 1681–1704, doi:10.1007/s00382-016-3408-9, 2017.

969 Lehner, B., Verdin, K. and Jarvis, K.: New global hydrograghy derived from spaceborne

970 elevation data, Eos, Trans. AGU, 89(10), 93–94, doi:10.1029/2008EO100001, 2008.

971 Lohmann, U. and Feichter, J.: Global indirect aerosol effects: a review, Atmos. Chem.

972 Phys., 5(3), 715–737, doi:10.5194/acp-5-715-2005, 2005.

973 MacMartin, D. G., Caldeira, K. and Keith, D. W.: Solar geoengineering to limit the rate

of temperature change, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 372(2031), doi:10.1098/rsta.2014.0134, 2014.

MacMartin, D. G., Ricke, K. L. and Keith, D. W.: Solar geoengineering as part of an overall strategy for meeting the 1.5°C Paris target, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 376(2119), doi:10.1098/rsta.2016.0454, 2018.

Masaki, Y., Hanasaki, N., Biemans, H., Schmied, H. M., Tang, Q., Wada, Y., Gosling, S. N., Takahashi, K. and Hijioka, Y.: Intercomparison of global river discharge simulations focusing on dam operation - Multiple models analysis in two case-study river basins, Missouri-Mississippi and Green-Colorado, Environ. Res. Lett., 12(5), doi:10.1088/1748-9326/aa57a8, 2017.

Mateo, C. M., Hanasaki, N., Komori, D., Tanaka, K., Kiguchi, M., Champathong, A., Sukhapunnaphan, T., Yamazaki, D. and Oki, T.: Assessing the impacts of reservoir operation to floodplain inundation by combining hydrological, reservoir management, and hydrodynamic models, Water Resour. Res., 50(9), 7245–7266, doi:10.1002/2013WR014845, 2014.

Mateo, C. M. R., Yamazaki, D., Kim, H., Champathong, A., Vaze, J. and Oki, T.: Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods, Hydrol. Earth Syst. Sci., 21(10), 5143–5163, doi:10.5194/hess-21-5143-2017, 2017.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., van Vuuren, D. P. P.: The RCP greenhouse gas concentrations and their

extensions from 1765 to 2300, Climatic Change, 109:213, doi:10.1007/s10584-011-0156-z, 2011.

Mueller, B. and Seneviratne, S. I.: Systematic land climate and evapotranspiration biases in CMIP5 simulations, Geophys. Res. Lett., 41(1), 128–134, doi:10.1002/2013GL058055, 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, 31, 6319-6340, doi:10.1175/JCLI-D-17-0620.1, 2018.

Niemeier, U., Schmidt, H. and Timmreck, C.: The dependency of geoengineered sulfate aerosol on the emission strategy, Atmos. Sci. Lett., 12(2), 189–194, doi:10.1002/asl.304, 2011.

Niemeier, U., Schmidt, H., Alterskjær, K. and Kristjánsson, J. E.: Solar irradiance reduction via climate engineering: Impact of different techniques on the energy balance and the hydrological cycle, J. Geophys. Res. Atmos., 118(21), 11905–11917, doi:10.1002/2013JD020445, 2013.

Niemeier, U. and Timmreck, C.: What is the limit of climate engineering by stratospheric injection of SO2?, Atmos. Chem. Phys., 15(16), 9129–9141, doi:10.5194/acp-15-9129-2015, 2015.

Niemeier, U. and Tilmes, S.: Sulfur injections for a cooler planet, Science (80-.)., 357(6348), 246–248, doi:10.1126/science.aan3317, 2017.

Oman, L. D., Ziemke, J. R., Douglass, A. R., Waugh, D. W., Lang, C., Rodriguez, J. M.

and Nielsen, J. E.: The response of tropical tropospheric ozone to ENSO, Geophys. Res. Lett., 38(13), 2–7, doi:10.1029/2011GL047865, 2011.

Pappenberger, F., Dutra, E., Wetterhall, F. and Cloke, H. L.: Deriving global flood hazard maps of fluvial floods through a physical model cascade, Hydrol. Earth Syst. Sci., 16(11), 4143–4156, doi:10.5194/hess-16-4143-2012, 2012.

Pawson, S., Stolarski, R. S., Douglass, A. R., Newman, P. A., Nielsen, J. E., Frith, S. M. and Gupta, M. L.: Goddard earth observing system chemistry-climate model simulations of stratospheric ozone-temperature coupling between 1950 and 2005, J. Geophys. Res. Atmos., 113(12), 1–16, doi:10.1029/2007JD009511, 2008.

Pitari, G., Aquila, V., Kravitz, B., Robock, A., Watanabe, S., Cionni, I., Luca, N. De, Genova, G. Di, Mancini, E. and Tilmes, S.: Stratospheric ozone response to sulfate geoengineering: Results from the Geoengineering Model Intercomparison Project (GeoMIP), J. Geophys. Res. Atmos., 119(5), 2629–2653, doi:10.1002/2013JD020566, 2014.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D.: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment, Proc. Natl. Acad. Sci., 111(9), 3262–3267, doi:10.1073/pnas.1222473110, 2014.

Rahman, A. A., Artaxo, P., Asrat, A. and Parker, A.: Developing countries must lead on solar geoengineering research, Nature, 556(7699), 22–24, doi:10.1038/d41586-018-

03917-8, 2018.

Rasmussen, P. F. and Gautam, N.: Alternative PWM-estimators of the gumbel distribution, J. Hydrol., 280(1–4), 265–271, doi:10.1016/S0022-1694(03)00241-5, 2003.

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.

Rienecker, M. M. and Coauthors: The GEOS-5 Data Assimilation System—Documentation of versions 5.0.1 and 5.1.0, and 5.2.0, NASA Tech. Rep. Ser. Glob. Model. Data Assim. NASA/TM-2008-104606, 27(December), 92pp, doi:10.2759/32049, 2008.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L. and Kim, G.-K.: MERRA: NASA's modern-era retrospective analysis for research and applications, J. Clim., 24(14), 3624–3648, 2011.

Robock, A., Kravitz, B. and Boucher, O.: Standardizing experiments in geoengineering, Eos (Washington. DC)., 92(23), 197, doi:10.1029/2011EO230008, 2011.

Smyth, J. E., Russotto, R. D. and Storelvmo, T.: Thermodynamic and dynamic responses of the hydrological cycle to solar dimming, Atmos. Chem. Phys., 17(10), 6439–6453, doi:10.5194/acp-17-6439-2017, 2017.

Schmidt, H., Alterskjær, K., Alterskjær, K., Bou Karam, D., Boucher, O., Jones, A., Kristjánsson, J. E., Niemeier, U., Schulz, M., Aaheim, A., Benduhn, F., Lawrence, M. and Timmreck, C.: Solar irradiance reduction to counteract radiative forcing from a

1062   quadrupling of CO2: Climate responses simulated by four earth system models, Earth

1063   Syst. Dyn., 3(1), 63–78, doi:10.5194/esd-3-63-2012, 2012.

1064   Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A.,

1065   Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A.,

1066   Dufresne, J.-L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S.,

1067   Rummukainen, M. and Smith, B.: Impact of soil moisture-climate feedbacks on CMIP5

1068   projections: First results from the GLACE-CMIP5 experiment, Geophys. Res. Lett.,

1069   40(19), 5212–5217, doi:10.1002/grl.50956, 2013.

1070   Sonntag, S., Ferrer González, M., Ilyina, T., Kracher, D., Nabel, J. E. M. S., Niemeier,

1071   U., Pongratz, J., Reick, C. H. and Schmidt, H.: Quantifying and Comparing Effects of

1072   Climate Engineering Methods on the Earth System, Earth's Futur.,

1073   doi:10.1002/2017EF000620, 2018.

1074   Suzuki, T., Yamazaki, D., Tsujino, H., Komuro, Y., Nakano, H. and Urakawa, S.: A

1075   dataset of continental river discharge based on JRA-55 for use in a global ocean

1076   circulation model, J. Oceanogr., (123456789), doi:10.1007/s10872-017-0458-5, 2017.

1077   Tanoue, M., Hirabayashi, Y. and Ikeuchi, H.: Global-scale river flood vulnerability in

1078   the last 50 years, Sci. Rep., 6(1), 36021, doi:10.1038/srep36021, 2016.

1079   Thompson, J. R., Green, A. J., Kingston, D. G. and Gosling, S. N.: Assessment of

1080   uncertainty in river flow projections for the Mekong River using multiple GCMs and

1081   hydrological models, J. Hydrol., 486, 1–30, 2013.

1082   Tilmes, S., Garcia, R. R., Kinnison, D. E., Gettelman, A. and Rasch, P. J.: Impact of

1083   geoengineered aerosols on the troposphere and stratosphere, J. Geophys. Res. Atmos.,

1084   114(12), 1–22, doi:10.1029/2008JD011420, 2009.

1085   Tilmes, S., Fasullo, J., Lamarque, J. F., Marsh, D. R., Mills, M., Alterskjær, K., Muri,

1086   H., Kristjánsson, J. E., Boucher, O., Schulz, M., Cole, J. N. S., Curry, C. L., Jones, A.,

1087   Haywood, J., Irvine, P. J., Ji, D., Moore, J. C., Karam, D. B., Kravitz, B., Rasch, P. J.,

1088   Singh, B., Yoon, J. H., Niemeier, U., Schmidt, H., Robock, A., Yang, S. and Watanabe,

1089   S.: The hydrological impact of geoengineering in the Geoengineering Model

1090   Intercomparison Project (GeoMIP), J. Geophys. Res. Atmos., 118(19), 11036–11058,

1091   doi:10.1002/jgrd.50868, 2013.

1092   Tilmes, S., Richter, J. H., Kravitz, B., MacMartin, D. G., Mills, M. J., Simpson, I. R.,

1093   Glanville, A. S., Fasullo, J. T., Phillips, A. S., Lamarque, J.-F., Tribbia, J., Edwards, J.,

1094   Mickelson, S. and Gosh, S.: CESM1(WACCM) Stratospheric Aerosol Geoengineering

1095   Large Ensemble (GLENS) Project, Bull. Am. Meteorol. Soc., doi:10.1175/BAMS-D-

1096   17-0267.1, 2018.

1097   Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger,

1098   J., Seland and Heinze, C.: Evaluation of the carbon cycle components in the Norwegian

1099   Earth System Model (NorESM), Geosci. Model Dev., 6(2), 301–325, doi:10.5194/gmd-

1100   6-301-2013, 2013.

1101   Trigg, M. A., Birch, C. E., Neal, J. C., Bates, P. D., Smith, A., Sampson, C. C.,

1102   Yamazaki, D., Hirabayashi, Y., Pappenberger, F., Dutra, E., Ward, P. J., Winsemius, H.

1103   C., Salamon, P., Dottori, F., Rudari, R., Kappes, M. S., Simpson, A. L., Hadzilacos, G.

1104   and Fewtrell, T. J.: The credibility challenge for global fluvial flood risk analysis,

1105   Environ. Res. Lett., 11(9), 94014, doi:10.1088/1748-9326/11/9/094014, 2016.

1106 Ukkola, A. M., Pitman, A. J., De Kauwe, M. G., Abramowitz, G., Herger, N., Evans, J.

1107 P. and Decker, M.: Evaluating CMIP5 Model Agreement for Multiple Drought Metrics,

1108 J. Hydrometeorol., 19(6), 969–988, doi:10.1175/JHM-D-17-0099.1, 2018.

1109 UNISDR (The United Nations International Strategy for Disaster Reduction): Global

1110 Assessment Report on Disaster Risk Reduction: From Shared Risk to Shared Value –

1111 The Business Case for Disaster Risk Reduction, Glob. Assess. Rep. Disaster Risk

1112 Reduct., 246, 2013.

1113 Vano, J. A., Udall, B., Cayan, D. R., Overpeck, J. T., Brekke, L. D., Das, T., Hartmann,

1114 H. C., Hidalgo, H. G., Hoerling, M., McCabe, G. J., Morino, K., Webb, R. S., Werner,

1115 K. and Lettenmaier, D. P.: Understanding uncertainties in future Colorado River

1116 streamflow, Bull. Am. Meteorol. Soc., 95(1), 59–78, doi:10.1175/BAMS-D-12-

1117 00228.1, 2014.

1118 Veldkamp, T. I. E., Zhao, F., Ward, P. J., de Moel, H., Aerts, J. C. J. H., Schmied, H. M.,

1119 Portmann, F. T., Masaki, Y., Pokhrel, Y., Liu, X., Satoh, Y., Gerten, D., Gosling, S. N.,

1120 Zaherpour, J. and Wada, Y.: Human impact parameterizations in global hydrological

1121 models improve estimates of monthly discharges and hydrological extremes: a multi-

1122 model validation study, Environ. Res. Lett., 13(5), 055008, doi:10.1088/1748-

1123 9326/aab96f, 2018.

1124 Visioni, D., Pitari, G. and Aquila, V.: Sulfate geoengineering: A review of the factors

1125 controlling the needed injection of sulfur dioxide, Atmos. Chem. Phys., 17(6), 3879–

1126 3889, doi:10.5194/acp-17-3879-2017, 2017.

1127 Wada, Y., de Graaf, I. E. M. and van Beek, L. P. H.: High-resolution modeling of human

and climate impacts on global water resources, J. Adv. Model. Earth Syst., 8(2), 735–763, doi:10.1002/2015MS000618, 2016.

Wang, B. and Ding, Q.: Changes in global monsoon precipitation over the past 56 years, Geophys. Res. Lett., 33(6), 1–4, doi:10.1029/2005GL025347, 2006.

Wang, W., Lu, H., Ruby Leung, L., Li, H., Zhao, J., Tian, F., Yang, K. and Sothea, K.: Dam construction in Lancang-Mekong River Basin could mitigate future flood risk from warming-induced intensified rainfall, Geophys. Res. Lett., 378–386, doi:10.1002/2017GL075037, 2017.

Wang, Q., Moore, J. C. and Ji, D.: A statistical examination of the effects of stratospheric sulfate geoengineering on tropical storm genesis, Atmos. Chem. Phys., 18(13), 9173–9188, doi:10.5194/acp-18-9173-2018, 2018.

Ward, P. J., Jongman, B., Weiland, F. S., Bouwman, A., van Beek, R., Bierkens, M. F. P., Ligtvoet, W. and Winsemius, H. C.: Assessing flood risk at the global scale: model setup, results, and sensitivity, Environ. Res. Lett., 8(4), 44019, doi:10.1088/1748-9326/8/4/044019, 2013.

Ward, P. J., Eisner, S., Flörke, M., Dettinger, M. D. and Kummu, M.: Annual flood sensitivities to El Niño–Southern Oscillation at the global scale, Hydrol. Earth Syst. Sci., 18(1), 47–66, doi:10.5194/hess-18-47-2014, 2014.

Ward, P. J., Kummu, M. and Lall, U.: Flood frequencies and durations and their response to El Niño Southern Oscillation: Global analysis, J. Hydrol., 539, 358–378, doi:10.1016/j.jhydrol.2016.05.045, 2016.

Ward, P. J., Jongman, B., Aerts, J. C. J. H., Bates, P. D., Botzen, W. J. W., DIaz Loaiza,

A., Hallegatte, S., Kind, J. M., Kwadijk, J., Scussolini, P. and Winsemius, H. C.: A global framework for future costs and benefits of river-flood protection in urban areas, Nat. Clim. Chang., 7(9), 642–646, doi:10.1038/nclimate3350, 2017.

Wartenburger, R., Seneviratne, S. I., Masaki, Y., Morfopoulos, C. and Christoph, M.: Evapotranspiration simulations in ISIMIP2a - Evaluation of spatio-temporal characteristics with a comprehensive ensemble of independent datasets, Environ. Res. Lett., 13(7), 1–49, doi:10.1088/1748-9326/aac4bb, 2018.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S. and Kawamiya, M.: MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments, Geosci. Model Dev., 4(4), 845–872, doi:10.5194/gmd-4-845-2011, 2011.

Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J. and Bouwman, A.: A framework for global river flood risk assessments, Hydrol. Earth Syst. Sci., 17(5), 1871–1892, doi:10.5194/hess-17-1871-2013, 2013.

Yamazaki, D., Oki, T. and Kanae, S.: Deriving a global river network map and its sub-grid topographic characteristics from a fine-resolution flow direction map, Hydrol. Earth Syst. Sci., 13(11), 2241–2251, doi:10.5194/hess-13-2241-2009, 2009.

Yamazaki, D., Kanae, S., Kim, H. and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resour. Res., 47(4), 1–21, doi:10.1029/2010WR009726, 2011.

Yamazaki, D., De Almeida, G. A. M. and Bates, P. D.: Improving computational

efficiency in global river models by implementing the local inertial flow equation and a vector-based river network map, Water Resour. Res., 49(11), 7221–7235, doi:10.1002/wrcr.20552, 2013.

Yamazaki, D., O'Loughlin, F., Trigg, M. A., Miller, Z. F., Pavelsky, T. M. and Bates, P. D.: Development of the Global Width Database for Large Rivers, Water Resour. Res., 50(4), 3467–3480, doi:10.1002/2013WR014664, 2014a.

Yamazaki, D., Sato, T., Kanae, S., Hirabayashi, Y. and Bates, P. D.: Regional flood dynamics in a bifurcating mega delta simulated in a global river model, Geophys. Res. Lett., 41(9), 3127–3135, doi:10.1002/2014GL059744, 2014b.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S. and Bates, P. D.: A high-accuracy map of global terrain elevations, Geophys. Res. Lett., 44(11), 5844–5853, doi:10.1002/2017GL072874, 2017.

Yu, X., Moore, J. C., Cui, X., Rinke, A., Ji, D., Kravitz, B. and Yoon, J. H.: Impacts, effectiveness and regional inequalities of the GeoMIP G1 to G4 solar radiation management scenarios, Glob. Planet. Change, 129(March), 10–22, doi:10.1016/j.gloplacha.2015.02.010, 2015.

Yu, M., Wang, G. and Chen, H.: Quantifying the impacts of land surface schemes and dynamic vegetation on the model dependency of projected changes in surface energy and water budgets, J. Adv. Model. Earth Syst., 8(1), 370–386, doi:10.1002/2015MS000492, 2016.

Zaherpour, J., Gosling, S. N., Mount, N., Schmied, H. M., Veldkamp, T. I. E., Dankers, R., Eisner, S., Gerten, D., Gudmundsson, L., Haddeland, I., Hanasaki, N., Kim, H.,

1194   Leng, G., Liu, J., Masaki, Y., Oki, T., Pokhrel, Y., Satoh, Y., Schewe, J. and Wada, Y.:

1195   Worldwide evaluation of mean and extreme runoff from six global-scale hydrological

1196   models that account for human impacts, Environ. Res. Lett., 13(6), 1–23,

1197   doi:10.1088/1748-9326/aac547, 2018.

1198   Zhao, F., Veldkamp, T. I. E., Frieler, K., Schewe, J., Ostberg, S., Willner, S.,

1199   Schauberger, B., Gosling, S. N., Schmied, H. M., Portmann, F. T., Leng, G., Huang, M.,

1200   Liu, X., Tang, Q., Hanasaki, N., Biemans, H., Gerten, D., Satoh, Y., Pokhrel, Y., Stacke,

1201   T., Ciais, P., Chang, J., Ducharne, A., Guimberteau, M., Wada, Y., Kim, H. and

1202   Yamazaki, D.: The critical role of the routing scheme in simulating peak river discharge

1203   in global hydrological models, Environ. Res. Lett., 12(7), 75003, doi:10.1088/1748-

1204   9326/aa7250, 2017.

1205   Zsótér, E., Pappenberger, F., Smith, P., Emerton, R. E., Dutra, E., Wetterhall, F.,

1206   Richardson, D., Bogner, K. and Balsamo, G.: Building a Multimodel Flood Prediction

1207   System with the TIGGE Archive, J. Hydrometeorol., 17(11), 2923–2940,

1208   doi:10.1175/JHM-D-15-0130.1, 2016.