Review for EGUsphere

Title: Future changes in atmospheric rivers over East Asia under stratospheric aerosol intervention

Authors: Ju Liang, Jim Haywood

Liang and Haywood analyze a series of climate change and geoengineering simulations from the UK ESM to diagnose impacts on atmospheric rivers (ARs) in the East Asia region. Both SAI and solar dimming experiments are used and compared to SSP5-8.5 and SSP2-4.5. The paper looks at both dynamical differences, precipitation impacts, and geometric AR structure. They find that ARs increase in frequency with related precipitation in warming scenarios, especially for southern China. SAI implementation reduces these impacts over the lower mid-latitudes and subtropics, but potentially exacerbates AR activity and impacts over upper midlatitudes, essentially due to dynamical shifts.

This paper is very well written and organized with clear figures and reasoning. The dynamical arguments are robust regarding connections to ARs. My main issue is with the AR metrics/features section, specifically inferences that are made in the text that are based on a box and whisker figure that does not necessarily support the conclusions stated. The differences between most of the simulation types are small with the spread often overlapping, and all based on a limited ensemble size. I recommend presenting this information in a more convincing manner, or removing this figure. Given only 1 ARDT is used, more information should be provided (see line comments below). Also, more information on UK ESM is needed, specifically for global, climate base states for SSPs compared to Geo simulations (again, see line comments below).

Lines 44-46: ARTMIP Tier 2, and especially O'Brien et al. 2022, have robustly shown that ARDT uncertainty far outweighs model uncertainty. It is also true that ARDT uncertainty outweighs uncertainty across Reanalysis products as well. The current sentence doesn't really communicate this finding, please adjust.
 https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021JD036155.

Line 52: Did you mean this reference? I don't see it in the reference list:
https://doi.org/10.1029/2019GL085565.

Line 110: What is the impact of adjusting the injection period? Also, the injection site? You touch on this in the discussion, but I have these additional questions). I realize that many models use a controller that bases locations based on temperature gradients which is determined by the climate base state. Is this true here as well? My understanding is that injection sites can make a big difference in terms of where aerosols circulate, and thus climate. Maybe a sentence or two acknowledging uncertainty due to injection site (or computation of injection site) can be added.

Line 112: Just to confirm, G6Solar, is dimming imposed on the base state of SSP5-8.5?

Line 120: The relatively low resolution used here compared to many AR studies should be addressed. For synoptics, this resolution is absolutely fine, but for regional precipitation, maybe

not, as seen in your supplemental Figure S2. Therefore, these limitations need to be discussed in the main paper.

Line 129: Why only 3 members? This seems a bit small to capture the internal variability, which is the purpose of ensembles.

Lines 134-145: Thank you for this supplemental information on model biases as they pertain to ARs! Other model biases important for SAI should also be included in supplemental, as well as some base state information to provide context as to how UKESM1 performs with and without SAI. For those not regularly following GeoMIP and SAI literature, this would be very helpful information. (Maybe global maps to see impact on other areas)? Also, in Figure S1, are the gray regions those with PS below 850mb?

Line 131:AR tracking was performed for each ensemble member, then the means were computed, correct?

Lines 150-170, AR detection: Thank you for this description. I did go to the supplemental material of Liang et al. 2022. I am assuming that the ARDT used in this work is the same ARIA-Asia described in this paper? If not, what are the differences and how do they compare to Figure S1 in that paper, i.e. other ARTMIP ARDTs? I would highly recommend a version of that figure to be included in the supplemental material here given it is directly relevant. Because not all readers have the same level of accessibility to papers, it is best to have everything in one spot, especially when that information is so closely tied to the interpretation of results. Given this, I also recommend a more detailed description (including geometry, see Figure 7 comments) should be provided in the supplemental rather than relying on the Liang references.

Figure 1: Here and elsewhere, please darken the continental outlines, they are really hard to see.

Line 193: Seems like there are some words missing at the end of the sentence?

Line 197: Is there more monsoonal precip here as well? The different SSP responses are interesting! For the aerosol feedbacks, is there a reference (such as the Jones et al 2021 you use later), or is this current work?

Line 205: Shields et al., 2022 also shows that the dynamical AR climate/SAI change response is dictated by the jets.

Line 219: with magnitudes (of weakening) greater than any other experiment?

Figures 5,6,7: Unit should be labeled as (fraction of 6-hourly time steps) and not %

Line 256/257/324/345/356 and maybe elsewhere): Not sure "high latitude" is appropriate here, perhaps, upper-midlatitudes.

Line 261: I am confused on the unit… % of time period, or fraction of time period? Percent is usually given in units 0-100%, whereas fraction is 0 to 1.

Figure 7 and discussion:

- Are some of the geometrical qualities somewhat "baked into" the ARDT by definition (line 155)? The geometrical conditions are not specified in the AR detection section, but given this figure, I think they should be stated. Perhaps more detail in the ARDT in the supplemental material is needed, rather than relying on the given references.

- Although the discussion (I think) is using mostly the mean values and movement of the max and min whiskers, I am a little uncomfortable with these statements given that the *range* of possibilities within each category are quite similar for all panels except 7f. Some of the metrics have increased, or decreased, *range envelopes*, but that is about all you can say, especially given the small sample/ ensemble size. Please rewrite this section keeping in mind the uncertainties of each metric. The fact that many of these whiskers (for each metric) completely overlap each other means that there is high uncertainty.

Line 287-288: AR frequency and size (as shown by the O'Brien paper) are very much tied to ARDT specifications. This, plus Fig 7, does not convince this statement can be made for anything over than moisture content for certain experiments.

Line 292-293: Same as above

Figure 9 and 10: Because this area also experiences tropical cyclones, again, a brief sentence on how TCs are filtered out of the ARDT would be useful in the AR detection section (or supplemental). This may be for a different study, but it would be interesting to understand how much of the total precipitation is TC-related vs AR-related.

Line 351-352: An increase in wind speed along the AR core is not supported by Figure 7.