

## **Review of: "Stratospheric ozone projections under sulfur-based stratospheric aerosol injection: Insights from the multi-model G6-1.5K-SAI experiment"**

By: Ewa M. Bednarz, Amy H. Butler, James M. Haywood, Matthew Henry, Andy Jones, Ben Kravitz, Walker R. Lee, Douglas G. MacMartin, Amanda C. Maycock, Takashi Sekiya, Shingo Watanabe, Daniele Visioni

### **Paper overview**

The manuscript presents work on estimating the stratospheric ozone impacts from the stratospheric aerosol injection (SAI) model intercomparison project GeoMIP, experiment G6-1.5K-SAI. The study shows results from three different Earth system models for future scenario modelling using the SSP2-4.5 scenario. The results on ozone impacts are discussed for two different future time periods and the analysis is split into different regions of the Earth as well as focused on the global impacts on ozone.

I think the manuscript provides a usual set of results for the scientific community and new information on the ozone impacts from SAI, necessary before widespread use of these suggested technologies. However, I have several points that need to be addressed before the manuscript is ready for publication, with one point being the clarity of writing. The presentation of the results lacks clarity in the main messages, which would enhance the importance of the results further. I recommend major revisions as described below in the overall comments followed by line-by-line comments.

### **Overall major comments**

Only MIROC includes an ensemble of 10 members, UKESM and CESM have 3 or 5 members, which is a very low number to obtain robust climate varying temperature scenarios. The temperature changes are central in the analysis of the resulting ozone reductions in the future scenarios as presented in the paper. Please comment on this and address the issue in the introduction and discussion of the results.

The UKESM model runs with heterogeneous chemistry only on background or SAI aerosols make the comparison between the three models difficult as they're not compared like for like. Also, the three models have chemical mechanisms for halogen heterogeneous chemistry that are significantly different, which needs to be highlighted in the beginning of the paper as well as in the discussion of the results, as this could potentially be a main driver of the observed model differences.

What would this analysis show if you focused on stratospheric ozone and not the total column? Are there any differences between the models that would be exaggerated or diminished? How much do these injections influence tropospheric ozone concentrations? i.e. through stratosphere to troposphere transport? Discussion of this would benefit the study.

The writing is very convoluted in places, with a lot of inserted sentences in the main sentence. This makes it difficult to ascertain the key messages, and the findings of the study are obscured. Please go through the text

and check the longer sentences for clarity and possible ambiguity. Examples are highlighted in the line-by-line comments.

The figures need to be more streamlined and described in full. I have several points for you to go through the figures in the main text and in the supplementary material: 1) All figures need to be described precisely and in detail and discussed in the main text or moved to the supplementary material. 2) Each figure must be able to stand alone, to be assessed without previous figures or additional publications. 3) In some figures and parts of figures you refer to/label figures as annual mean data, where you are presenting multi-year averages. Please be precise in how you refer to this and avoid labelling things as annual means if it is not that. 4) Include a secondary vertical axis of altitudes in units of km to all vertical data as well as a line highlighting the tropopause height (if this changes in the SAI vs SSP model run, choose which one and describe this in the caption/text where needed). 5) Add tick marks on the right-hand side of your figures to make it easier for the reader to analyze the figures. 6) The y-axis is plotted vs. the x-axis, not the other way around. Please change this in how you refer to the data plotted on the figures.

### **Line-by-line comments**

Line 50: "... that impact concentrations of species directly relevant for chemical ozone production and loss." This should be specified in more detail. It reads as overly vague as it is, the details would be helpful to have here in the introduction as you discuss the specific species in the results.

Line 53: "Any local changes..." are these in addition to and/or separate from the aerosol-induced temperature changes? In the line above you discuss temperature changes from additional aerosol, this sentence is not clear if this is from the same aerosol change, please clarify. I think you mean to say something more generally about impacts of changing temperatures in the stratosphere than only from sulfate aerosols in lines 53-55, however, this needs to be clearer.

Line 85: "30N+30S" – missing degree signs, please add them and replace the plus-sign with "and". Adding the note here to remind you to go over the manuscript and check for inconsistencies in how you refer to injection locations.

Line 98-101: Two comments here: 1) Figure 1b is mentioned before 1a, this would suggest that you should switch the order of those two panels in your figure, so they're labelled in the order they appear in the text. 2) I'm unsure how the injections were done reading this sentence. Did you use the resulting surface temperature to adjust the amount of SO<sub>2</sub> to inject? The baseline that you are adjusting to include years where SO<sub>2</sub> is injected – why was this chosen as the reference point? Can you please add a bit more information and re-phrase to argue the way the injection has been set up?

Figure 1: How can you show an annual mean "YM" for panels c, d, and e? I think you are showing a mean of all the years 2045-2064. Please remove the YM label on the panels and see comments on editing the figure caption below. Formatting: Almost all text on the figure is too small, the labels on the colorbars are too small to read and you're missing degree signs, subscripts on units as well as showing mixed styles of panels in the sense of

font types and sizes. For this figure you only need one colorbar for the bottom row of panels as you show three identical colorbars here. Please check this and the other figures in the manuscript to make sure they are legible and have correct units.

Figure 1 caption: Please add in the caption text exactly what is shown in each panel. It is important to note the model names in the caption. Line 105: “annual mean global mean”, please rephrase to the more commonly used “global annual mean” or similar. Bottom row: “annual mean 2045-2064 changes”, so you are showing changes between the two years or the mean of all those years? Or did you take the mean of each year and then do something else? In line 109-110, you say “between G6-1.5K-SAI and SSP2-4.5 over the same period in each model”. Reading this I conclude that you have taken a mean of 2045-2064 (not “annual” as these are multiple years) for two model ensembles, see comment above about “YM”: 1) SSP-reference run and 2) SAI model run. The “same time period” refers to what exactly? I’m not convinced you should have a range of years to begin with and then also add this. This caption needs to be clarified.

Line 115: Here you refer to the model as “CESM2(WACCM)”. This results in two brackets next to each other; this should be avoided by rephrasing. I.e., this could be achieved by the common way of referring to the model “CESM2/WACCM” using a slash instead of bracket.

Line 116: Why has the “ES2H” part of MIROC abbreviation not been described? You describe version numbers for the other two models, leaving it out here is inconsistent.

Line 132-145: I am not following how these four runs in total relate to each other from this text. The caption in SI is a lot cleaner and easier to follow. Can you please rephrase this paragraph to resemble more the SI caption? The mention of original simulations (from the Lee-study) and discussing how different simulations are used instead is confusing – it is not clear why you would necessarily want to repeat the published work by Lee et al. How are you using their temperature finding here if you are using a different set of model simulations?

Line 142: “our revised simulations” which simulations are you referring to here?

Line 142-146: “the same absolute target global mean near surface temperature from the original SSP2-4.5 experiment was used in the revised G6-1.5K-SAI simulations, leading to a slight overcooling (~0.2 K) compared to the reference 2020-2039 mean surface temperature in the revised SSP2-4.5 runs (Fig. 1a). We note this slight overcooling occurs under smaller SO<sub>2</sub> injection rates (Fig. S1c), as the SAI-induced ozone reduction (Section 3) acts to amplify the aerosol-induced surface cooling.” From the referred figures, I don’t see this overcooling being very significant until the end of the century – if this is what you want to highlight, say that this is the timeframe for comparison. However, why did you not adjust the SO<sub>2</sub> injected to account for the differences between the NO\_HET and WITH\_HET runs? And is the difference in SO<sub>2</sub> injections in S1c really significant? I see the resulting difference in S1a, but related to that: please explain why the difference between the two SSP runs in S1a diverge more as time passes?

Figure 2: I am concerned regarding panel a that the model differences are largely due to the different chemical mechanisms as described in the supplementary material as UKESM and CESM are showing the largest increases in column ozone, particularly UKESM is showing recovery of global column ozone over time in panel a, and not the GHG emissions differences as you highlight in line 186-188. Can you add a description/discussion on why this is not the case, please. And/or evaluate/assess the reasons for the differences from different chemistry vs different GHG emissions? How are you determining the uncertainties represented as error bars? I am missing a description of this in the text. Finally, same as for the previous figure: YM shouldn't be used in panel c as this is a multi-year mean not the annual mean of one year.

Figure 2 caption: line 173: CESM2(WACCM-MA) is not described before here, and abbreviation is not spelled out in the main text. Please add this information.

Line 183-184: "As shown in Figure S2(a-c), while models generally agree on the magnitude of upper stratospheric ozone increase" the way this is phrased it sounds like you are showing a model comparison of the upper stratosphere, which cannot be seen in figure 2, please rephrase to clarify what you are showing in the figure and what is a separate statement.

Line 192-194: "The latter occurs despite the increasing SO<sub>2</sub> injection rates - and hence increasing stratospheric aerosol burden - over the course of the 21st century in order to offset progressively higher surface temperatures under rising GHGs in SSP2-4.5 (Fig. 1a-b)." The meaning of this sentence gets lost with the inserted sentence referring to aerosols burden, can you rephrase this to make it clear that you are increasing the injected SO<sub>2</sub> over time to account for increasing temperatures due to increasing GHGs in the SSP scenario, please. Highlighting that what you adjusting the models to achieve is the  $\Delta T/\Delta SO_2$  as described in your table 1.

Line 203: all these numbers seem very detailed; do you really have enough confidence to claim two significant digits for the DU ozone changes? Please check the statistical significance of these numbers. Also, these changes are all very small – what are they equivalent to in percentage? You refer to percentages earlier in this paragraph, please make these numbers easily comparable to that.

Line 208: The years in the bracket refers to the literature data from Bednarz et al 2023a, the bracket is misplaced here as it suggests that you are showing new data in this time range. Please make it clear that you are comparing the present study's two averages to the literature 2050-2069 values for the different strategies. Maybe talking the years out of the bracket and rephrasing to accommodate this action or moving the bracketed years to a later place in the sentence (i.e., line 209) would work to clarify this along with adding the time ranges for the present study. Cf. your writing in line 226.

Line 213: please add altitudes and/or pressure levels for your definition of "middle atmosphere"

Line 221: please add altitude and/or pressure levels for your definition of "middle-stratosphere", also remove the hyphen to be consistent with the previous text, see comment on line 213. Highlighting this here to bring to your attention to check the manuscript for this and similar inconsistencies.

Line 222 and 223: "heterogenous N<sub>2</sub>O<sub>5</sub> hydrolysis on sulfate" and "halogen activation on sulfate", I think you mean to say "sulfate aerosols" here, please correct this.

Line 224: 1) “lower stratosphere”, see previous comments, add altitude to your definition. 2) “at 30N+30S” this is not a location when written this way, but a model simulation label. Please be precise and consistent in how you refer to the model simulations, in this case it should be “in 30N+30S” to refer to a model simulation or replace with the proper latitudes with degree signs.

Figure 3: Please add a proper caption for this figure It needs to be able to stand alone from figure 2. Panel a has a legend that covers some of the data, please correct this to allow inspection of that missing data in the ensemble means. Scale the graph to cover all of your data, missing the yellow UKESM peak around 2058 in both panels a and b and maybe also for 2035 in panel b

Figure 4 caption, line 246: remove “over the same period for each model” you are already saying the this is the changes for your two time periods in line 245, repeating it after mentioning the SSP scenario is confusing and unnecessary.

Line 250-251: “the tropical ozone column is projected to decrease slightly under the SSP2-4.5 scenario over the 21st century (Fig. 3a)” I disagree on this part of the sentence: Inspecting figure 3a MIROC and CESM start decreasing after ~2050 and UKESM only after ~2070 if taken as all values are below the zero-line in your figure.

Line 251-252: “This arises because of the GHG-induced acceleration of the BDC increasing transport of ozone-poor tropospheric air into the lower stratosphere” this statement needs a reference, please add this here.

Line 252-253: “the resulting lower stratospheric ozone reduction offsets the ozone increase in the upper stratosphere (Fig. S2a-c)” I do not see the upper stratospheric increase in the S2 figure, please expand your argument for this statement. Also, reading your vertical figures would be easier if you include information on the tropopause height in all vertical figures.

Line 265-266: “These differences in tropical total column ozone responses amongst the three models reflect the complex pattern of tropical ozone changes at different altitudes.” I think you need to add more information to this statement, as you are not including the complexity of the chemical schemes being different in each of the models as highlighted in your own Table S1. You need to add this complexity to the discussion, already here in the first sentence of this discussion or in the beginning of the paragraph. Your current chemistry discussion start in line 287.

Line 273: typo “insolation”, should be “insulation”

Line 307-309: add “vapor” in a few places in this paragraph to make it clear that you are talking about the gas-phase.

Line 320: figure panel labels: even when CESM doesn’t have a panel, the bottom to panels should still be d and e, not skipping the letter d. Please update this accordingly in the figure caption and main text.

Line 330-331: In your figure caption: “(B): As in (A) but for changes in the shallow branch of the BDC, defined as in Bednarz et al. (2023a), i.e. 15°S-15°N mean TEM  $w^*$  changes at 100 hPa.” This text needs to explain what is shown in panel B, referring to Bednarz et al 2023a is necessary, but this figure and the paper should be able

to stand alone. Therefore, you should add more details to this caption, i.e., explain the definition of the shallow BDC branch, “TEM” abbreviation, and describe “w\*”

Line 335: Add “vapor” to the mention of water or add “gas-phase” somewhere in the first line to describe all three species. I am missing the precision in the descriptions of what the figures show, see overall comment as well.

Figure 8: All panels, but in particular panels c and f need to be explained in the figure caption. It should be stressed that some panels are showing changes in DU and c and f in DU/deg K

Line 350-361: discusses figure 8 panels e-f, the following paragraph discusses figure 8 a-c. Please move and re-phrase the text to discuss the panels in order (or re-organize the figure panels to match the order they are mentioned). This will help the reader follow the logic of the text and figure together better.

Line 351-352: You write: “... reductions of up to approx. -10 DU (with the 2045-2064 average of -7 to -9 DU SH mid-latitude ozone loss depending on the model) under SAI compared to SSP2-4.5 throughout the length of the simulation (Fig. 8e).” However, in panel 8e, I see values as low as ~-11 DU for UKESM and CESM in that time range (values below -10 DU). The panel is very difficult to read accurately, making evaluating the numbers difficult, however, please update these numbers in your range to match the data.

Line 363-364: Your text: “Here, higher climatological lower stratospheric temperatures (Fig. 5) slow down the rates of heterogenous halogen activation on sulfate...”. Can you please add discussion on reactions that increase with lower temperatures? The ClONO<sub>2</sub> hydrolysis reaction is a key example of where the reaction becomes less important at higher temperatures, which is the opposite to your overall statement. This information needs to be included in the discussion if you want to give a statement about halogen chemistry.

Line 370-371: In your text: “These differences are likely a manifestation of large differences amongst the models in their projected chemical and dynamical responses to SAI,” please add the differences in halogen chemical mechanisms as an additional explanation

Line 385-388: Please comment here on the fact the MIROC has the most complete heterogeneous chemistry of all three models and the impact that has on the following discussions of transport vs. chemistry importance in MIROC and the other two models. This is necessary and has not been addressed sufficiently in the current manuscript, see overall comments.

Line 409-410 and following paragraphs: Here you start with discussing the SH referring to panels 10d-f, please re-organize the figure or the text so the panels appear in the order they are discussed, starting with panel a.

Line 418: “...at the same time...” please add the years to make it clearer that you refer to 2045-2064 averages

Line 419-421: please add what the ozone loss ranges are for the later period to contrast the 2045-2064 averages with the 2065-2084 ones.

Line 422-425: please comment on the large difference between the CESM 2045-2064 average in the main panel of figure 10f and the past CESM2 run with the 30S+30N strategy. In this figure they look very different, which hasn't been the case for the previous figures of the same setup.

Figure 11: 1) Clarify what the deltas refer to for all panels in this figure – I assume you mean (SAI – SSP2-4.5), but this should be explicit in the figure, text, and caption. 2) panels c and d are delta O<sub>3</sub> vs. the delta U, not as the titles say, 3) why does panel d include MAM for O<sub>3</sub> and DJFMAM for delta U? This discrepancy is not described as it should.

Line 453: please add reference to figure 10 to help the reader understand what figure you are describing – I see no reference to figure 10 panels a-c in the text; this needs to be rectified.

Line 453-462: In this paragraph, please add more description of figure 10a-c and of the relationship described in figure 11b and 11d. The two panels in figure 11 for the NH are referenced in the paragraph, but more description and discussion of the results are needed for background of the conclusion drawn in the paragraph. Please rephrase the text accordingly.

Line 480: "...dominant differences will be driven by the presence or absence of heterogeneous chemistry on SAI aerosol..." Even if you refer to section 2, you need to make a short statement of the main differences of heterogeneous chemistry in the two runs – the one used in the other parts of the paper doesn't have sulfate heterogeneous chemistry, but only on SAI, whereas both MIROC and CESM have heterogeneous reactions occurring on both types of particles at the same time. The caveats about using the two versions of UKESM along with the other two models and using the UKESM only to analyze chemistry vs. transport impacts need to be addressed in the discussion. See overall comment as well.

Line 487: "...driven partly by..." is this a typo? I think you mean to say "partly" or "clearly". Please fix this so the meaning becomes clear.

Line 491-493: these three lines are the only mention of figure 13, if you want to include it in the main text, please describe the information and analysis of it. At the moment it feels like it could be a supplementary figure.

Line 492: "...here approximated as altitudes above 12 km..." Please add the approximate pressure level that this refers to and add altitudes in km as a secondary y-axis on your vertical figures to aid the inspection of the data, see overall comment.

Line 498: This statement is not backed up with the data as it stands: "...ozone poor-air from the lowermost stratosphere (as also suggested by the corresponding changes in model age-of-air pointing towards younger air in these regions, Fig. S6)..." The main feature seen in figure S6 in all four panels is an increase in age-of-air in the polar regions below 300 hPa and in the tropics around 200 hPa and above. Below these altitudes you have this younger air (seen as a decrease). At around 30 hPa, some of the panels show decreases in age-of-air, but if this is what you mean to highlight it is not clear from the text. Please re-phrase this to match your data. The increases cannot be overlooked in your analysis, even if you want to focus on the decreases in age-of-air.

Figure 13: caption: please take the “here taken as altitudes above 12 km” out of the panel b label and add the information as text in the caption, including it with the “(b)” hides the panel designation. Include which pressure levels (approximately) the >12 km refers to in the caption as well.

Line 512-521: Please delete this paragraph as it is repeating points already made in the introduction with no new added information. Replace by one sentence introducing the summary and conclusion section instead, or simply start the section from line 523.

Line 528: “...increased heterogenous halogen activation on sulfate and the resulting enhancement of the halogen catalyzed ozone loss...” When you highlight this, you need to discuss the fact that the three models have very different halogen chemistry included, which complicates the comparison and analysis of the three model outputs. Somewhere close to this sentence, I would suggest to add that information.

Line 541: wrong use of “albeit”, it is synonymous to “though”, “while”, “but”, etc., using it here gives the idea that it is “in contrast to”, which suggests that the models don’t agree on the global ozone reductions, which I don’t think is what you intended. You mean to say that the regional SH mid-latitude ozone loss is larger than the global ozone loss, correct? Please remove the brackets, delete “albeit” and add the comparison to global ozone reductions as part of the main text. Even if you just delete “albeit” and keep the rest of the sentence, it will improve the sentence and understanding for the reader.

Line 554-555: “...becomes a factor of ~2 smaller than earlier in the century...” Again, can you please add the numbers in the 2065-2084 range? Are the models still showing a spread of the same magnitude or do they converge? Is CESM still showing the largest ozone reduction? The numbers to remind us of the figures would be helpful.

Line 566-577: You should discuss here the bias from using the UKESM chemistry versus using the more explicit one in MIROC and how this could potentially underestimate the importance of chemistry driving the changes. Particularly in relation to the discussion below of the model versus observations, where you solely point to transport changes. How would the discussion change if you add the complexity of varying the halogen heterogeneous chemical mechanisms in the models? Please comment and add this to the discussion in this paragraph.

Figure S3: add x-axis label to panel a, the ticks are missing degree signs and something to say “latitude” or similar for the axis. Please add a description for the terms mentioned in the figure. I.e., “ $\Delta T_{AS}$ ”, “YM”, even when these are described in the main text.