

Author's Response

We thank the Reviewers for their comments. On this document, texts in italics are Reviewers' comments. Author's replies are preceded by "⇒".

In the response below, unless otherwise specified, all page, line numbers, and figure numbers are referring to the marked-up version of the manuscript.

Author's response to RC1

This paper needs major revisions. It reports unusual responses to black carbon injections, but does not diagnose the climate response to explain why. It seems that your model does not loft the BC into the stratosphere nor rain it out, and so it persists somehow in the upper troposphere, heating it and probably evaporating clouds. Is this physically realistic? Please show the vertical distribution of the BC and clouds over the regions that experience surface warming in the first JJA. Without looking at the distribution of clouds, precipitation, and BC, how can we figure out why there is warming. Is it downward emission of longwave from the BC cloud? How is this possible if there are water clouds? The paper calls for future work on that, but this paper needs that analysis. Otherwise, the strange result is not explained.

⇒ We thank the Reviewer for emphasizing the need to make the physical mechanisms more explicit. While our original submission did include diagnostics of the climate response (Sections 3.2 and 3.3), we agree that the explanation of the warming mechanism needed to be clearer and more directly linked to the spatial structure of the response. In the replies below, we have worked with the Reviewer's comments in mind, to clarify and further explore the climate response. For example, we have emphasized changes in the cloud distribution that are of interest as well as changes in column water vapor and clear-sky surface DLRF.

⇒ In the original manuscript, we diagnosed the climate response to black carbon (BC) injection in detail in Section 3. The surface climate impacts were diagnosed in Section 3.2, and the surface energy fluxes were diagnosed in Section 3.3.

We showed that the model did loft the BC into the stratosphere in the simulations and discussed the removal of BC in the troposphere through both dry and wet deposition. For example, as we discussed in Section 3.1, the BC was lofted into the lower stratosphere (the so-called "overworld" term used in the literature on stratosphere troposphere exchange). The overworld burden for the BC24 case attained its maximum at 5.2 Tg in

first year June, while the overworld burden for BC12 attained its maximum at 2.0 Tg in the same month (Fig. 1a inset).

In addition, we showed how the model did remove the BC in the simulations. As can be seen in Fig. 1b, 70-80% of the injected BC was removed from the atmosphere in the first year. The model can do this via either dry or wet deposition (not only via “rain out”). The treatment of dry and wet deposition in CanESM5 was described in Section 2.1.

We showed that the BC does not persist in the upper troposphere. Part of the BC was lofted into the stratospheric overworld, with this portion shown in the inset of Fig. 1a.

Another part of our description of the climate response relates to the character of the heating induced by the BC injection. While the maximum heating occurs in the stratosphere (Fig. 3b), it is indeed the case that some of the radiative heating response does descend into the upper troposphere. This downward extension of the region of radiative heating into the upper troposphere is physically realistic in our view.

⇒ One new point introduced by the Reviewer, which we had not included in the original submission, relates to the possibility of a cloud response, in particular a decline in cloud fraction in the upper troposphere. We appreciate the Reviewer’s encouragement to focus more of our analysis on the regions where surface warming occurs. This has led to what we will propose as the most significant revision to the paper. As suggested, to show the vertical distribution of the BC and clouds over the regions that experience surface warming in the first JJA, we have highlighted those regions in Fig. 6 and plotted the corresponding distributions in Fig. 7 and Fig. A11. The precipitation change is plotted in Fig. A12.

One insight we gained is that the intense downward emission of longwave from the intense warming in the lower stratosphere (Fig. 3), drives warming in the first year’s boreal summer (JJA) in regions that have generally low cloud fraction to start with (Fig. A2). Therefore, the changes in clouds in these relatively clear and arid regions do not appear to be a primary driver of the warming signal. We did, however, examine the cloud fraction response (Fig. A3), and it reveals significant increases in cloud fraction in northern Africa and some significant decreases in western North America and Tibet. The mechanism of the warming is not being primarily related to cloud responses.

To deepen our understanding of the clear-sky warming effects, we examined the clear-sky surface DLRF and water vapor changes. We find that the global mean clear-sky surface DLRF response is similar to that of the all-sky surface DLRF response, with the same peak in the first JJA (Fig. A4 and 2c). Both the clear-sky surface DLRF response and the column water vapor response are spatially similar to the increase in surface DLRF (Fig. A5, A6, and 6a). All three of the responses in water vapor, clear-sky surface DLRF, and all-sky surface DLRF over these land regions that warm in the first JJA are greater than

those over other land regions (Fig. A5, A6, and 6a), suggesting that downward emission of longwave from BC aloft is being amplified by water vapor feedback for these regions in the first JJA.

⇒ The vertical distributions of BC, cloud, and water vapor are physically consistent with one another. There is BC enhancement in the troposphere over the warming regions in the first JJA, but the BC enhancement is much greater in the stratosphere (Fig. 7 first row). With upper tropospheric warming, the cloud fractions in the upper troposphere do decline for these regions, but the cloud fraction in the middle to lower troposphere increases for the northern African region (Fig. A11 first row). The specific humidity has a great percentage increase in the UTLS (Fig. A11 second row), which is consistent with the UTLS warming. The relative humidity decreases in the UTLS due to intense warming, but it increases in the middle to lower troposphere, which is consistent with the suggestion of water vapor feedback.

Thus, we find that overall, for these warming regions in the first JJA, the increased surface DLRF is associated with increased column water vapor and BC burden. This effect is consistent with the nighttime warming (Fig. A9) that we discussed in the original manuscript.

While we did find a mix of increases and decreases of precipitation over the regions that warm in the first JJA (Fig. A12), the contribution of precipitating processes to the transient year-1 warming signal is not clear.

⇒ Once again, we appreciate the Reviewer's encouragement to look at new aspects of the response. With the Reviewer's comments in mind, we have investigated additional aspects of the climate response beyond what was shown in the original submission. Based on the additional diagnostics now included, we find no evidence that the simulated response is inconsistent with our physical understanding. Therefore, we have no concern with physical realism of these simulations beyond those that generally apply to coarse resolution global models used in this field.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

We have revised the manuscript to add Fig. 7 (Line 331-334), Fig. A2-A6 (Line 424-441), A11 (Line 461-464), A12 (Line 466-468), and relevant analysis in the Results (Line 298-309).

We have included a discussion of the regional water vapor responses in the Discussions and Conclusions (Line 349-354).

Why does the paper choose the BC injections that they did? They call for model intercomparison studies, but do not use the same injection scenarios as other studies. It seems they are not aware

of Toon et al. (2019) which examined scenarios with the same range as this paper, 5 Tg, 15 Tg, 27 Tg, and 37 Tg. This paper should have used the same ones so they could be compared.

Toon, Owen B., Charles G. Bardeen, Alan Robock, Lili Xia, Hans Kristensen, Matthew McKinzie, R. J. Peterson, Cheryl Harrison, Nicole S. Lovenduski, and Richard P. Turco, 2019: Rapid expansion of nuclear arsenals by Pakistan and India portends regional and global catastrophe. *Science Advances*, 5, eaay5478, doi:10.1126/sciadv.aay5478.

⇒ We appreciate this important point. We chose the BC injections that we did since Turco et al. (1990) estimated 20 Tg to be the lower bound of the plausible range of BC surface emissions for the full-scale nuclear exchange scenario, as mentioned in the Introduction. The BC injections of this size are extremely large compared to climatological BC burden and could cause unprecedented climate impacts, albeit relatively smaller than 150 Tg, the mean of the estimated range.

⇒ Our focus here is the lower-end US-Russia exchange scenario motivated by Turco et al. (1990), whereas Toon et al. (2019) examines India-Pakistan exchange scenarios.

⇒ Our suggestion for intercomparison is motivated by how this relatively underexplored regime, from a hypothetical limited nuclear exchange, leads to strong, disruptive, and unexpected transient impacts. We believe this would be valuable to check in other models.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

We have clarified the distinction between US-Russia exchange scenario which we consider and India-Pakistan exchange scenarios more explicitly in the Introduction (Line 63-66, 73-75).

They also should be aware of Oman et al. (2006). They found summer warming over Africa and Asia after a volcanic eruption, and the explanation was a weaker summer monsoon, which produced fewer clouds. This paper needs to show the distribution and anomalies of clouds and precipitation in the first JJA, and diagnose whether the response they found was due to a dynamic response, or simply an evaporation of the clouds due to upper troposphere heating.

Oman, Luke, Alan Robock, Georgiy L. Stenchikov, and Thorvaldur Thordarson, 2006: High-latitude eruptions cast shadow over the African monsoon and the flow of the Nile. *Geophys. Res. Lett.*, 33, L18711, doi:10.1029/2006GL027665.

⇒ Thank you for the suggestion. Our assessment of the role of clouds and water vapor has been discussed above and has determined that a clear-sky effect is common for all the regions while the cloud response does not provide a simple explanation. For example, in our model simulations, cloud fraction increases over the region of northern Africa that experiences warming in the first JJA (Fig. A3), so the suggested effect should not be relevant. The surface DLRF increase is likely more associated with water vapor feedback, as suggested above.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

The addition of Fig. 7 (Line 331-334), Fig. A2-A6 (Line 424-441), A11 (Line 461-464), A12 (Line 466-468), and relevant analysis in the Results (Line 298-309)

The inclusion of a discussion of the regional water vapor responses in the Discussions and Conclusions (Line 349-354)

The BC lifetime in this simulation is much shorter than in previous work. This has to be explained, with figures showing the BC distribution vertically. What was not washed out immediately should have been transported to the upper stratosphere and lasted for years. Previous work had e-folding lifetimes of 5-7 years. Why is it so short here?

⇒ Our reported findings of BC lifetimes are not very different from previous work. From previous work, the reported stratospheric BC mass e-folding time ranges from 2-8 years based on the previous literature we cited in the original manuscript (Coupe et al., 2019; Mills et al., 2014; Pausata et al., 2016; Robock et al., 2007a, 2007b; Wagman et al., 2020; see their full citations in original manuscript). The stratospheric BC mass e-folding time found in our simulations ranges from 3-4 years, which falls within previous reported values. While it is at the short end of the range, we believe that it is reasonable, given the limitation that CanESM5.1 is a low top model. While the BC lofting effect is present in the model, it may be limited. As stated above, we have included vertical distributions of BC.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

The addition of Fig. 7 (Line 331-334)

The terminology of “low-moderate” injections, “standard high-end cases,” and “relatively weak” injection needs to be changed. There is no standard case, and all incidents of nuclear war, including 5 Tg injections would be horrific, and would not be low or moderate.

175

176 ⇒ We agree that all incidents of nuclear war would be horrific. Indeed, this is what
177 motivated us to study the impacts of what would be even a limited nuclear exchange,
178 compared to the available arsenal, between these two states. But the term “low-to-
179 moderate” inadvertently connotes a lack of harm; that was not our intention.

180 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

181 We have replaced the terminology of “low-to-moderate” and “relatively weak” by
182 “relatively limited” (Line 1-2, 13-17, 58, 74, 76, 88, 92, 144, 185, 338, 341, 359-360,
183 367-368, 408).

184

185 *The caption for Fig. 4 says sea ice is shown by orange lines, but I don’t see any.*

186

187 ⇒ Thank you for catching this. In Line 249-250 in the first version of the manuscript “The
188 solid and dashed orange lines indicate the sea ice lines for the scenario and for the
189 climatology, respectively. It is defined as the grids with 15% sea ice area.” should belong
190 to the caption of Fig. 5.

191 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

192 We have moved the sentence “The solid and dashed orange lines indicate the sea ice lines
193 for the scenario and for the climatology, respectively. It is defined as the grids with 15%
194 sea ice area.” from the caption of Fig. 4 (Line 273-274) to the caption of Fig. 5 (Line
195 292-293).

196

197 *There are 23 additional comments on the attached annotated manuscript which also should be*
198 *addressed.*

199

200 *For comments on Line 1, 12, 14, 56, 64 on the annotated manuscript attached with RCI:*

201

202 ⇒ These are related to the comment on the use of the terminologies above.

203 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

204 We have replaced the terminology of “low-to-moderate” and “relatively weak” by
205 “relatively limited” (Line 1-2, 13-17, 58, 74, 76, 88, 92, 144, 185, 338, 341, 359-360,
206 367-368, 408).

207

208 *For comments on Line 27, 28 on the annotated manuscript attached with RCI:*

209

210 ⇒ We have included the suggested literature, which are relevant and recent.

211 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

212 Line 34-35, 606-608

213

214 *For comments on Line 50, 72, 331 on the annotated manuscript attached with RCI:*

215

216 ⇒ These are related to the comment on the choice of the BC injections, which has been
217 responded above.

218 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

219 We have clarified the distinction between US-Russia exchange scenario which we
220 consider and India-Pakistan exchange scenarios more explicitly in the Introduction (Line
221 63-66, 73-75).

222

223 *For comments on Line 143 on the annotated manuscript attached with RCI:*

224

225 ⇒ The first half of the comment is regarding stratospheric BC lifetime and has been
226 responded above. As for the second half of the comment on the use of CanESM5 for
227 studying the Mt. Pinatubo volcanic eruption, we note that CanESM5 has not been used to
228 actively simulate stratospheric aerosols from the Mt. Pinatubo eruption although it has
229 been used to study the climate response to the radiative forcing from the eruption, for
230 example in Pauling et al. (2023) and Rieger et al. (2020).

231 Pauling, A. G., Bitz, C. M., and Armour, K. C., 2023: The climate response to the Mt. Pinatubo
232 eruption does not constrain climate sensitivity. Geophysical Research Letters, 50,
233 <https://doi.org/10.1029/2023GL102946>.

234 Rieger, L. A., Cole, J. N. S., Fyfe, J. C., Po-Chedley, S., Cameron-Smith, P. J., Durack, P. J.,
235 Gillett, N. P., and Tang, Q., 2020: Quantifying CanESM5 and EAMv1 sensitivities to Mt.
236 Pinatubo volcanic forcing for the CMIP6 historical experiment. Geosci. Model Dev., 13, 4831–
237 4843, <https://doi.org/10.5194/gmd-13-4831-2020>.

238

239 *For comments on Line 176, 177 on the annotated manuscript attached with RCI:*

240

241 ⇒ Thank you for the corrections.

242 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

243 We have replaced “brackets” with “parentheses” (Line 195).

244

245 *For comments on Line 235 on the annotated manuscript attached with RCI:*

246

247 ⇒ This is related to the first comment above and has been responded.

248

249 *For comments on Line 249 on the annotated manuscript attached with RCI:*

250

251 ⇒ This is related to the question about the orange lines and has been answered above. Thank
252 you again for catching this.

253 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

254 We have moved the sentence “The solid and dashed orange lines indicate the sea ice lines
255 for the scenario and for the climatology, respectively. It is defined as the grids with 15%
256 sea ice area.” from the caption of Fig. 4 (Line 273-274) to the caption of Fig. 5 (Line
257 292-293).

258

259 *For comments on Line 270 on the annotated manuscript attached with RCI:*

260

261 ⇒ “Following the fire” was referring to following the fire emission.

262 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

We have replaced “following the fire” by “following the fire emissions” (Line 295).

For comments on Line 310, 311 on the annotated manuscript attached with RC1:

⇒ As suggested above, in our revision, we have included the vertical distributions of BC mass mixing ratio, cloud fraction, specific humidity, and relative. We have also added relevant discussion on water vapor feedback in Discussions and Conclusions.

⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

The addition of Fig. 7 (Line 331-334), A11 (Line 461-464), relevant analysis in the Results (Line 298-309), and a discussion of the regional water vapor responses in the Discussions and Conclusions (Line 349-354)

For comments on Line 409 on the annotated manuscript attached with RC1:

⇒ We use standard Copernicus formatting in the revised text.

For comments on Line 490, 493 on the annotated manuscript attached with RC1:

⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

We have specified 2007a and 2007b for the two references (Line 576, 579).

Author’s response to RC2

Major comments:

(i) *The authors claim they find a non-linear response across the different injection rates. However, 3 out of 4 simulations (including the 12 and 24 Tg injection cases used to produce the results in Table 1) consist of only 3 ensemble members and the analysis is done for single individual time periods (e.g. JF mean in year 1, JJA mean in year 1, etc), leading to 3 data*

points in total contributing to each response. While this should be sufficient to estimate the changes in BC burden and related variables (e.g. AOD) with relative confidence, I'm not convinced this is sufficient to estimate the response in global mean SW and LW fluxes or surface temperature with confidence sufficient to confidently claim whether linearity or non-linearity is found.

⇒ To address this comment, we have run additional simulations to increase the number of ensemble members to 9 and 8 for the BC12 and BC24 cases, respectively. The expanded ensembles confirm the original conclusions, with nearly identical central estimates and overall reduced uncertainty ranges. For example, please see the original and the updated values of DSRF, DLRF, and TS (Table 1 in original manuscript and revised manuscript); all results are consistent within their uncertainty ranges. In addition, please see the updated timeseries (Fig. 2), which is very similar to the original figure (Fig. 2 in the original manuscript).

⇒ After increasing the ensemble size, the surface DSRF reduction per Tg burden is still comparable across different cases, indicating a linear dependence of global mean surface DSRF response on BC burden after the first year (comparing inset of Fig. 2a in revised manuscript to inset of Fig. 2a in revised manuscript).

⇒ Regarding the nonlinear dependence of global mean surface DLRF response on BC burden in first year JJA, after increasing the ensemble size, it is still clear from Table 1 that the global mean surface DLRF response for BC12 is comparable to that for BC24. In other words, the response per unit burden for BC12 is much greater than that for BC24. This nonlinear dependence is also clear from the inset of Fig. 2c. During first year JJA, the separation of the global mean surface DLRF responses per unit burden is large between the different scenarios, with the color shading indicating the standard error of the ensemble mean.

⇒ Therefore, increasing the number of members does not change the discussion and conclusion in the manuscript.

Mills, M. J., Toon, O. B., Lee-Taylor, J., and Robock, A.: Multidecadal global cooling and unprecedented ozone loss following a regional nuclear conflict, *Earth's Future*, 2(4), 161-176, <https://doi.org/10.1002/2013EF000205>, 2014.

Robock, A., Oman, L., Stenchikov, G. L., Toon, O. B., Bardeen, C., and Turco, R. P.: Climatic consequences of regional nuclear conflicts, *Atmospheric Chemistry and Physics*, 7(8), 2003-2012, <https://doi.org/10.5194/acp-7-2003-2007>, 2007.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

We have updated the number of members for BC12 and BC24 in Methods (Line 140-141).

We have updated Fig. 1, 2, 4 and Table 1, as well as the corresponding values in the main text.

As is also evident in Fig. 2 (a,c,d), the interannual variability on most of these variables is not negligible. While Table 1 reports 95% confidence intervals, it is not clear to me how are these calculated.

⇒ The 95% confidence intervals are calculated by first computing the standard deviation of the aggregated timeseries. Then, the standard deviation is divided by the square root of the number of samples and is multiplied by 1.96, which corresponds to the 95% confidence interval. For first year JF (or first year JJA), the aggregated timeseries is the aggregation of first year JF average (or first year JJA average) of all ensemble members, so the number of samples is the number of ensemble members.

⇒ For year 6-15, the aggregated timeseries is the aggregation of all monthly output in year 6-15 of all ensemble members. The number of samples is the ensemble-average effective sample size after applying a lag-1 correlation to the monthly timeseries of each ensemble member. Thus, the effective sample size would be smaller than the total number of months in the aggregated timeseries.

⇒ On a related note, during the revision, we identified and corrected an issue in the implementation of the 95% confidence interval calculation. The corrected values do not alter the interpretation or conclusions.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

We have clarified the points above on Line 195-200.

Also, how big is the ensemble spread of these values? Is it adequately represented by the way the authors calculate the 95% confidence interval? I would expect even the value of standard deviation calculated based on just 3 data points will be subject to large uncertainty..

⇒ The standard deviations with the addition of new members are shown below (Table R1). The standard deviations are not large for first year JF and JJA, as anticipated above. The

standard deviations are larger for years 6-15 due to the trend towards climatology, but the confidence intervals, as stated above, still show significant results.

	DSRF	DLRF	TS
First year JF	-26.52 ± 0.78 (-14.80 ± 0.61)	$+2.59 \pm 0.30$ ($+1.72 \pm 0.33$)	-0.59 ± 0.10 (-0.38 ± 0.09)
First year JJA	-26.92 ± 0.57 (-13.00 ± 0.30)	$+6.20 \pm 0.56$ ($+6.44 \pm 0.46$)	-0.06 ± 0.12 ($+0.39 \pm 0.08$)
Years 6-15	-2.99 ± 8.00 (-0.75 ± 2.52)	-1.71 ± 2.17 (-0.32 ± 2.02)	-0.57 ± 0.38 (-0.14 ± 0.30)

Table R1. As in Table 1, but with ± 1 standard deviations instead of 95% confidence intervals.

On a related note – Section 2.2 discusses the nuclear winter experiments performed, but nowhere does it say what data are used for calculating the responses. Is it the preindustrial control simulation? If so, how many years and/or ensemble members are being averaged?

⇒ Yes, the preindustrial control simulation is used for calculating the responses. A climatology is calculated using the 100-year preindustrial control simulation. For each injection scenario, all ensemble members are being averaged.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

We have clarified this on Line 147-148.

(ii) I agree with the other reviewer that better understanding of the drivers of the reported boreal summer warming in year 1 is required before publication. In fact, I'm not convinced by the authors' claim that this is driven predominantly by the change in downward LW forcing, as the change in downward SW is at least factor of ~2 larger and opposite sign. And so considering just changes in direct downward radiative forcing, the overall change in RF should be still negative. And so the changes have to be driven by indirect processes, either changes in the upward LW/SW and/or latent and sensible heat fluxes.

⇒ It is true that the global mean surface DSRF response is about 2 times the global mean surface DLRF response in the opposite sign in first year JJA. In fact, after taking into

account latent heat flux and sensible heat flux, the global mean net flux at the surface is still negative (defined as positive downward). However, as we described in the manuscript (Line 273-276 and 290-292 on the original manuscript), the surface net energy flux imbalance needs to be considered in its entirety. This takes into account upward and downward shortwave radiative flux, upward and downward longwave radiative flux, latent heat flux, and sensible heat flux, and shows a lot of structure during the first boreal summer (Fig. 6b). The response is mostly positive downward over land surface and mostly negative downward over the ocean surface. The global mean temperature response is positive in that time period due to the land surface's lower heat capacity. With a higher heat capacity and mixing, the ocean temperature response is muted.

⇒ In the original manuscript, we described the full surface energy budget in its entirety, but we agree with the reviewer that we need to make this clearer, and we do so in the revision.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

We have changed “surface net flux” to “surface net energy flux” (Line 327-328). We have also changed “surface net energy imbalance” to “full surface energy imbalance” (Line 310-311).

In fact, to me the pattern of surface temperature changes looks very much like a dynamically induced response (I was thinking something to do with NAM maybe), or a BC-induced modulation of the model's internal modes of climate variability. In fact in figure 2d, it looks like an injection drives an oscillatory behavior in the model, that is largest for high BC injections. How realistic is this response, and could that somehow explain what is happening at the surface? Also, I see that the authors do principal component analysis of the DJF Tas response (to comment on links with AO modulation). I think it might be helpful to do the same for the JJA in year 1, my guess would be that the pattern obtained would mimic the response to BC simulated in JJA that year (this is just my guess, but maybe worth checking).

⇒ In the original manuscript, we did principal component analysis (PCA) of the DJF Tas response and attempted to attribute part of the Tas response to a positive AO modulation effect. However, since the DJF Tas response is dominated by direct radiative flux changes, we ended up not including this analysis.

⇒ For JJA Tas, while we expect the warming pattern to be modulated by advection of air warmed by the heat sources (Fig. 6b), we decided not to include a PCA as showing the temperature advection response pattern (Fig. A8) is sufficient to indicate that advection is

playing a role in modulating the Tas response pattern, e.g. for Tas response in western North America (Fig. 5c).

⇒ In Fig. 2d, the oscillatory behaviour in the global mean temperature is largest for low BC injections in the first year, for reasons mentioned in the paper. As RC2 suggests, it is largest for high BC injections in the long term. The peaks are during boreal summer, and the troughs are during boreal winter, suggesting that the cooling is more efficient during boreal winter than during boreal summer. This can be seen by comparing Fig. 5b and Fig. 5d.

On a related note – Fig. 6 shows the map of changes in DLRF and net RF response. If one shows DLRF, I think it is important to also show DSRF (since as discussed above DSRF is much larger so it's important to see whether that is also true locally over the region in question), as well as decompose the net RF budget calculation. The authors try and discuss parts of the latter in the text, but the reader can't verify (or visualize) the accuracy of the statements without seeing the plots somewhere (even if it's just in the supplementary material)

⇒ This comes back to our earlier point that Fig. 6 shows the full surface energy flux response, not just the net radiative response. Again, this takes into account upward and downward shortwave radiative flux, upward and downward longwave radiative flux, latent heat flux, and sensible heat flux, as explained in the original manuscript on Line 273-276 and 290-292.

⇒ As shown by Fig. 6b, locally on land in the Northern Hemisphere, the surface net energy flux response is positive (defined as positive downward) in first year JJA.

⇒ **Author's corresponding changes in the marked-up version of the manuscript:**

To avoid ambiguity, we have changed “surface net flux” to “surface net energy flux” on Line 327-328. We have also changed “surface net energy imbalance” to “full surface energy imbalance” (Line 310).

For reader's reference, we have included the decomposition of the surface energy budget in the Appendix (Fig. A7 or Line 443-447) and have referred to it on Line 311-312.

Minor changes:

- L. 17-18 “reduced net long wave and turbulent fluxes”. Please state direction. Throughout most of the manuscript it is argued that the downward LW increases, but in the abstract you say the net LW actually decreases - I guess that’s because net is defined as upward, but it is confusing to the reader here.

⇒ Indeed, in the abstract, the net LW is defined as upward.

⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

We have changed “linked to reduced net longwave and turbulent fluxes” to “linked to reduced cooling via net longwave and turbulent fluxes” on Line 23.

- Table 1 - the number of decimal places that are reported is too high, given all the uncertainties.

⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

We have kept the number of decimal places to 2 in Table 1.

- Table 1 – “the confidence intervals for the response in years 6-15 are calculated taking into account autocorrelation”. Does it mean you use data for years 6,7,8,9,10,11,12,13,14 and 15 separately, instead of mean over years 6-15? If so, this is not clear, and given there is a strong trend present within that period, I’d expect the 95% confidence interval to be much larger in that case (as I would expect it to reflect that trend)

⇒ For year 6-15, the aggregated timeseries that is used to compute the confidence interval is the aggregation of all monthly output in year 6-15 of all ensemble members. The number of samples is the ensemble-average effective sample size after applying a lag-1 correlation to the monthly timeseries of each ensemble member. Indeed, after fixing the aforementioned bugs, the 95% confidence intervals become larger for the year 6-15 quantities (Table 1), but the main conclusion stands. Thank you for pointing this out.

- Fig. 2 caption – please make sure you include description of all inserts, and with correct labels (some are missing, and some are mixed up I think)

485 ⇒ The caption should say “Insets of (a) and (c)” instead of “Insets of (a) and (d)”. Thank
486 you for catching this.

487 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

488 We have changed the caption of Fig. 2 (Line 205).

489
490 - *Fig. 2 – “shading indicates standard error of the mean” – but the plot shows the difference in*
491 *two means? (i.e. with and without BC injection)*

492
493 ⇒ Thank you for asking. The shading indicates ± 1 standard error of the ensemble mean.

494 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

495 We have replaced “Shading indicates the standard error of the mean” with “Shading
496 indicates ± 1 standard error of the ensemble mean” (Line 168, 207).

497
498 - *Fig. 5 – what are orange lines?*

499
500 ⇒ The solid and dashed orange lines indicate the sea ice lines for the scenario and for the
501 climatology, respectively. It is defined as the grids with 15% sea ice area.

502 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

503 We have added this to the figure caption (Line 292-293).

504
505 - *Section 2 – while the aerosol microphysics scheme used is a bulk one, we are still missing*
506 *some information of the assumed BC aerosol properties, including size.*

507
508 ⇒ The assumed hydrophilic BC effective radius and variance are 0.166 microns and 0.617,
509 respectively. The effective refractive index for internally mixed aerosol is computed using
510 the Maxwell-Garnett (MG) approximation the radiative (Wu et al., 2018). It assumes that
511 the host material (SO₄) contains all other composite materials (BC and OC).

512 ⇒ **Author’s corresponding changes in the marked-up version of the manuscript:**

513 We have added the sentence “The assumed hydrophilic BC effective radius and variance
514 are 0.166 microns and 0.617, respectively.” on Line 122.