We thank Reviewer 1 for their positive and helpful comments towards improving this manuscript. Our responses are given in the blue font. Where necessary, we refer to the revised version of the manuscript.

General comments:

Byrom et al. have created a clear manuscript that conveys scientifically relevant findings. They do this by expanding upon work that has highlighted a stratospheric base state dependence on CMIP6 model 4xCO2 ERF, investigating the role of O3 that plays a significant role in determining stratospheric temperatures.

They construct stratospheric ozone change experiments in NorESM2, prescribing ozone from the higher stratospheric-resolution CESM2-WACCM, to demonstrate how stratospheric temperatures change with O3 concentrations. They then calculate IRF (through offline radiative code) and adjustments (through radiative kernels) that result from a quadrupling of CO2 from these different stratospheric states.

Through this, and by comparing methods with previous work, they show the importance of calculating IRF through offline code, and using appropriate kernels that avoid extrapolating to above the top of the model. They further show that IRF depends on the base state of the stratosphere, due to overlapping O3 and CO2 spectral bands and stratospheric temperature. However, stratospheric temperature adjustments do not change significantly between increases and decreases of O3. Cloud adjustments result in invariant ERF, which highlights that stratospheric O3 concentrations are unlikely to explain the inter-model spread in 4xCO2 ERF.

I would like to see this published, though I have some minor concerns that I would like to see addressed.

Specific comments:

L112: on calculating the cloud adjustment as a residual, I'm not confident that A_c is being properly estimated. It seems like the same criticism levied against calculating IRF as the residual could apply to calculating A_c as a residual. Non-zero ϵ may explain the differences in cloud feedbacks shown in Figure 3. I'd like to see further justification for assuming approximately zero ϵ.

We have now calculated the cloud adjustment following the adjusted cloud radiative effect method of Soden et al. (2008; see https://journals.ametsoc.org/view/journals/clim/21/14/2007jcli2110.1.xml), whereby the cloud radiative effect is calculated as the difference between clear-sky and all-sky ERF and adjusted to correct for cloud masking of the clear-sky forcing and adjustments. For Figure 1 we use the NorESM2-MM all-sky and clear-sky 4xCO_2 shortwave and longwave ERFs and correct for the cloud masking by using output from the corresponding PORT 4xCO2 simulation and clear-sky and all-sky radiative adjustments from the CESM-CAM5 kernels:

dLW_cloud = -d_cre_lw + cloud_masking_of_forcing_lw + (dLW_q_cs - dLW_q) + (dLW_ta_cs - dLW_ta) + (dLW_ts_cs - dLW_ts)

dSW_cloud = -d_cre_sw + cloud_masking_of_forcing_sw + (dSW_q_cs - dSW_q) + (dSW_alb_cs - dSW_alb)

This calculation is repeated for the cloud adjustment in Figure 3 using the data from the respective experiments presented.

Note that Øivind Hodnebrog provided calculations for this cloud adjustment, hence his addition as a co-author on the manuscript.

L117: It's not clear to me how you calculate the adjustment to surface temperature change (A_T_s). I think a brief explanation is warranted, or just clarification that you used the same method as in the papers that follow the same approach.

We use radiative kernels to calculate this adjustment in the same way that all of the other adjustments are calculated (apart from the cloud adjustment). With regard to the sentence: "Several studies follow this approach (e.g., Hansen et al., 2005; Forster et al., 2016; Smith et al., 2018)." – this was intended to highlight that several other studies also calculate fixed SST ERF in the same way (whereby land-surface temperatures are allowed to respond to the forcing given the difficulty in prescribing fixed surface temperatures) and not necessarily that we calculate (A_T_s) in the same way as these papers (although we do with respect to Smith et al. 2018). This section of the text has now been updated to be clearer (lines 132-137):

"Additionally, we also use kernels to calculate the adjustment due to surface temperature change (A_{T_S}) as in Smith et al. (2018), since land surface temperatures are allowed to respond to the forcing in our simulations given the difficulty in prescribing fixed surface temperatures (Forster et al., 2016). Several studies also follow this approach whereby the calculation of ERF includes the radiative response of land surface warming or cooling (e.g., Hansen et al., 2005; Forster et al., 2016; Smith et al., 2018; Smith et al., 2020a)."

Figure 3: especially given the small number of years (9, instead of the ~20 you might use from 30 year runs) used in calculating values in the figure, this figure could benefit from error bars to show standard errors. If error bars are very small, I think the results would still benefit from acknowledging this.

We agree and we have now added standard error bars to this figure (see revised manuscript).

Fig. S3: CESM2 and NorESM2 are very similar, so isn't comparing what they do here quite limiting? It might be worth clarifying this.

Yes, it is quite limiting and the use of an altogether different model with a different radiation scheme would definitely be interesting to use. NorESM2 uses a different module for aerosol physics and chemistry, including cloud and radiation interactions and has further differences in the land component and surface albedo calculation. However, we have now deleted this comparison from the text and supplementary seeing as the comparison was somewhat limited and further calculations would now need to be performed (for the CESM2 cloud adjustment).

Technical corrections:

L17-19: Unclear sentence "However… O3". This could read as though the base-state stratospheric temperatures and the spectral overlap of CO2 and O3 are counteracting the spread in CO2 ERF. Given what you write in your conclusions, it seems like this should be something along the lines of "The spread in CO2 IRF is explained by the impact of base-state stratospheric temperature on the emission of outgoing longwave radiation and the spectral overlap of CO2 and O3, but these do not explain the spread in CO2 ERF".

Thanks for pointing this out, we have now amended the text to be clearer as suggested (lines 16-19):

"These experiments impact the IRF due to the influence of base-state stratospheric temperature on the emission of outgoing longwave radiation and the spectral overlap of CO_2 and O_3. However, the impact on IRF does not result in a correspondingly large spread in CO_2 ERF. We conclude that inter-model differences in stratospheric O_3 concentration are therefore not predominantly responsible for inter-model spread in CO_2 ERF."

L233-244: I think the use of brackets here leads to less clarity. This seems like a good example of what has been criticised in previous literature (https://doi.org/10.1029/2010EO450004), where parentheses are not used for clarification. I would recommend writing separate sentences e.g. for increase vs. decrease of O3 instead of trying to save space with parentheses.

This has been noted and the sentence has now been re-written as two separate sentences (see lines 277-289):

"As stratospheric O_3 concentrations increase, TIR absorption at these wavelengths also increases to an extent that depends on the level of band saturation and the abundance of other gases absorbing at these wavelengths. The opposite occurs if stratospheric O_3 concentrations decrease. For CO_2, the main TIR bands lie in the window regions of the H_2O spectrum, with absorption centered at 4.3 µm and 15 µm (the latter of which is highly significant due to its proximity to the peak of blackbody distribution for the Earth's effective emitting temperature). Weaker bands also occur near 10 µm. Regions of spectral overlap between O_3 and CO_2 therefore arise at several wavelengths; at 15 µm the strength of CO_2 absorption largely masks the radiative effect of O_3 at 14.27 µm and absorption by both gases at 4.75 µm and 4.3 µm has little impact given their location further away from the peak of Earth's blackbody distribution. However, decreased stratospheric O_3 concentration leads to weakened absorption at 9.6 µm that can enhance absorption by CO_2 at 10 µm. Likewise, increasing stratospheric O_3 concentration results in strengthened 9.6 µm absorption that mutes CO_2. Combining the effect of base-state stratospheric temperature and spectral overlap, a 4xCO_2 perturbation therefore results in an enhancement of the IRF in the 'Strat O_3x0.5' case relative to the 'standard' experiment. Correspondingly, an increase in stratospheric O_3 has the opposite (albeit evidently weaker) effect."

Figures (generally): the figures are a bit blurry, which makes it hard to read some elements especially in Figs. 1, 3, and S3. Ideally these would be higher resolution, or the font size might be increased.

The figures have now been updated using a higher resolution, please see the revised manuscript. Apologies for the blurriness.

Fig. S3: why the change of colours from Fig. 3? I found it a bit harder to distinguish the shades of blue than the shades of orange.

Originally the results from NorESM2 and CESM2 were plotted together, so the orange and blue shades were used to distinguish the results from each model. But the CESM2 comparison has now been deleted from the manuscript.