We thank the reviewer for their overall positive review and their constructive feedback. Below we provide point-by-point responses (in red) to the reviewer's comments (reproduced in black italics), with manuscript modifications indicated in **bold**.

This manuscript presents original and valuable earth system model results based on sensitivity experiments to investigate the tropospheric VOC chemistry effect on climate. It is well structured, written and presented. I suggest acceptance of the manuscript for publication but I have a few minor comments to be considered.

Comments

Something that I missed in the structure of the manuscript was a discussion of the induced radiative forcing due to the perturbation with explicit VOC chemistry (WACtl) versus the case with a simplified SOAG scheme and no explicit VOC chemistry (MACtl). Would you think that an estimate of effective radiative forcing from these two experiments would be feasible following the regression method of Gregory et al. (2004). For this method, it is the time development before the steady state is reached which is of interest and in your simulations the first 39 years of output (which were discarded because they showed significant trends in global mean surface temperature, indicating that the model had not fully equilibrated) may offer this opportunity.

Thank you for this suggestion, as an analysis of the ERF would facilitate comparison with other climate forcers. Unfortunately, however, the regression approach of Gregory et al. (2004) is not well-suited for our model experiments. This is because there is initial warming of global mean surface temperature (GMST) over the first ~20 years of WACtl relative to MACtl (the reasons for which are unclear), even though the steady-state GMST response is negative. That is, the transient and equilibrium radiative forcings appear to be opposite in sign, and fixed SST simulations would be more suitable for calculating the ERF (e.g., Hansen et al., 2005; Forster et al., 2016), which is not a simple undertaking. We think that the complexity in the transient vs. equilibrium radiative forcing and the reasons for it are worth investigating in a separate study, and we will add text to this effect within the discussion of Fig. 1.

line 130: I would suggest to add a few references for SOAG scheme implemented?

Thank you for the suggestion. The following text will be added here: "**The simplified SOA scheme in MACtl is known as a SOAG scheme, and it is commonly implemented in GCMs. Under this scheme, anthropogenic and biogenic precursor VOCs are assumed to have fixed mass yields and are sorted into five primary VOC bins. Within MAM, these gas-phase precursors are lumped into a single semi-volatile organic gas-phase species called SOAG, which then condense and are emitted into the model as primary aerosol particles (Liu et al., 2012; Emmons et al., 2010).**"

lines 198-199: The fact that the two cases were not initialized from the same ocean and sea ice states is a limitation. Although you mention that in the future, you plan to examine possible sensitivity to different initialization approaches, could you speculate based on other studies on the degree that this could influence your results?

Thank you for raising this point. The figure below shows the SST difference (WACtl – MACtl) averaged over just the first simulation day.

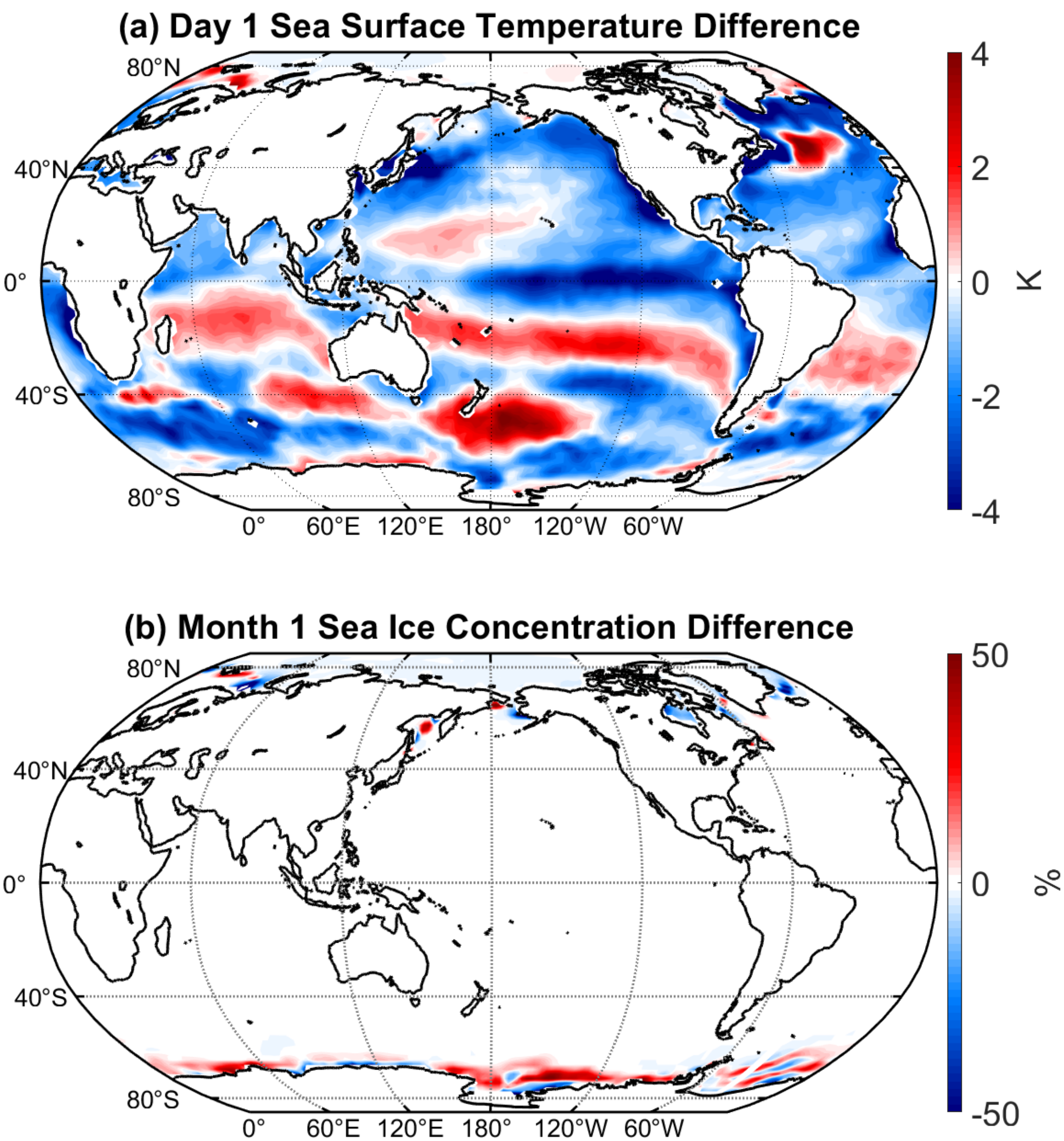

Figure R1-1: First day mean difference (WACtl – MACtl) in (a) SST (K) and (b) sea ice concentration (%).

The SST pattern is similar to the difference between a La Nina state and an El Nino state. Furthermore, the sea ice contraction differences appear to be mostly dictated by the temperature differences, with lower (higher) sea ice concentration where temperature is warmer (colder). Overall, this analysis suggest that the difference between the WACtl and MACtl initial states is a difference in the phase of internal variability, rather than a fundamentally different climate state. Thus, MACtl and WACtl appear to originate from essentially indistinguishable climate states, even though their initial conditions are not precisely the same. Therefore, we expect that any long-term difference between MACtl and WACtl would be due to the effects of climate forcing during the simulations rather than climate differences in the initial states. If we were to perform another run, "WAmatched", with exactly the same initial conditions as MACtl, then we expect that the difference between WAmatched and WACtl would just be internal variability (similar to the difference between ensemble members in an initial condition ensemble), and the long-term average of WAmatched – MACtl would look essentially the same as WACtl – MACtl. Also note that the difference (WACtl – MACtl) between the long-term average of SSTs (Fig. 1 in the manuscript) does not resemble figure R1-1 above, further indicating that the difference between WACtl and MACtl is not influenced by the difference in the initial states. We will add text to Section 2.4 to make these points.

lines 297-299: The authors mention that " This makes sense because, if low cloud cover is fixed (which is approximately true around Antarctica) and SIC decreases (increases), then the radiative effect of low clouds gets stronger (weaker), implying a negative (positive) CRE change." Maybe the authors can clarify that this is related to albedo changes and not cloud fractions changes.

Thank you for this point. We have modified the text here as follows: "This makes sense because, if low cloud cover is fixed (which is approximately true around Antarctica) and SIC decreases (increases), then **the surface albedo decreases (increases), and** the radiative effect of low clouds gets stronger (weaker), implying a negative (positive) CRE change."

Figure 5: I am puzzling by the fact that the statistical significant near surface warming over the North polar regions seen in Figure 1 is not shown in Figure 5. Do you have any explanation for this inconsistency among the two figures?

Below we show the same plot as in Fig. 5 but using a linear pressure scale, which reveals more detail in the lower troposphere.

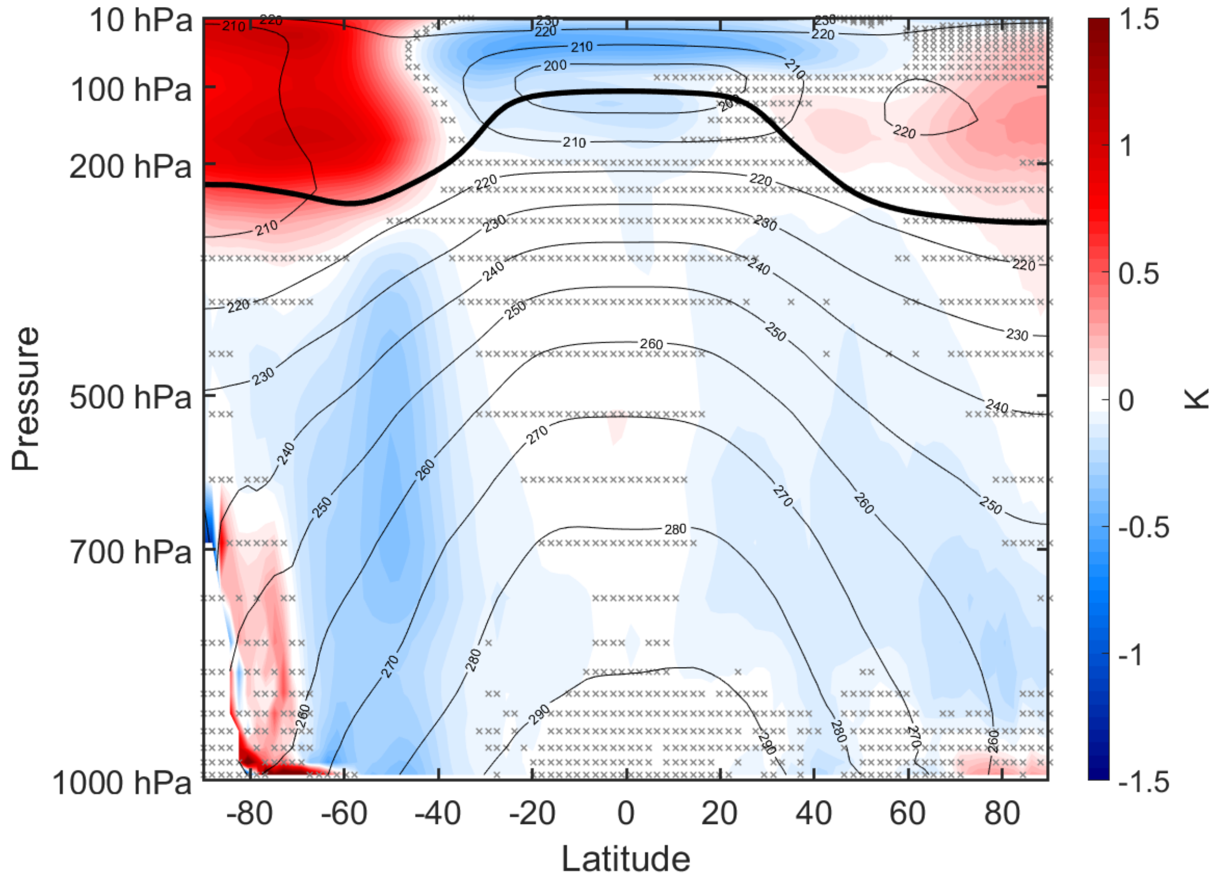

Figure R1-2: Annual mean difference (WACtl-MACtl) in atmospheric temperature (K). The black contours indicate the MACtl climatology of air temperature (K), with a contour interval of 10 K. Gray crosses indicate regions that are statistically insignificant, with uncrossed regions being statistically significant at the 95% level. The thick black line marks the MACtl climatology of the tropopause. The pressure scale is linear.

This shows that there is indeed warming in the Arctic troposphere, but it is confined to very low altitudes, as we stated in the manuscript when discussing Fig. 5. So there is no inconsistency between Figs. 1 and 5.

line 325: There is a weakening of the NH stratospheric polar jet but it is not clear the equatorward shift of the tropospheric mid-latitude jet.

Thank you, we will modify the text here as follows: "In NH, there is a qualitatively similar weakening of the stratospheric polar vortex and **zonal wind deceleration on the poleward flank of the midlatitude jet. However, the zonal wind response is quantitatively weaker in NH than in SH, and there is no clear indication of an equatorward shift of the midlatitude jet in NH.**"

line 328: This is confusing as in the next sentence you clarify that the responses are opposite to the poleward shift due to global warming. Are you referring to stratosphere warming (as shown in Figure 5) or the thermal forcings that qualitatively mimic three key aspects of anthropogenic *climate change: (warming in the tropical troposphere, cooling in the polar stratosphere, and warming at the polar surface) discussed in Butler et al (2010) referenced here?*

Thank you for pointing out this possible point of confusion. The key thing to bear in mind from earlier studies is that the circulation responds differently depending on whether warming is confined to particular latitudes or spread over all latitudes. To avoid possible confusion, we will modify the text here as follows: "**Care is needed when comparing our zonal wind response to the responses in earlier studies, since the atmospheric circulation response is highly sensitive to the regional structure of temperature changes. For example, idealized simulations have shown that a narrow band of warming around the equator produces equatorward shifts of the midlatitude jets, whereas warming spread over all latitudes produces poleward shifts (Tandon et al., 2013). In our simulations,** the zonal wind responses in both hemispheres qualitatively resemble the circulation responses to high latitude warming found in earlier studies (e.g., Son et al., 2008; Butler et al., 2010), and we will explore the mechanisms generating this warming further below. Moreover, these responses are qualitatively opposite to the poleward midlatitude jet shifts produced in simulations of global warming **(with warming over all latitudes)**, including simulations in which SST is uniformly increased (e.g., Chen et al., 2013).

lines 330-332: The authors mention that "based on such past work, we would expect that, aside from any regional temperature changes, widespread tropospheric cooling would also shift the midlatitude jets equatorward". Do you refer to past work related to the impact of aerosol cooling? Please prove some references.

Thanks for this suggestion. We will add the following text here: "**Indeed, simulations of the response to increased natural aerosols have shown widespread tropospheric cooling and equatorward shifts of the midlatitude jets (Allen and Sherwood, 2011).**"

line 342: The use of arrows on Fig.6b could help the reader to identify the counterclockwise anomaly extending poleward of the SH HC edge and a clockwise anomaly extending poleward of the NH HC edge. This is only a suggestion.

Thank you for this suggestion. We will look into implementing this suggestion in the revised manuscript.

lines 360-361: Pacific. The statement that "in addition to these HC changes, Fig. 6b also shows weakening of the Ferrel cells in both hemispheres and weakening of the polar cell in SH" needs more elaboration as it is not clear.

We will clarify this point by adding the following text: "**Weakening of the Ferrel cell in the NH is indicated by a positive anomaly where there is negative climatology, while in the SH it is indicated by a negative anomaly where there is positive climatology. The weakening of the polar cell in the SH is indicated by a positive anomaly where there is negative climatology.**"

line 383: Please define the acronym SPS.

Thank you for catching this. We will modify the text here as follows: “**Southern Polar Stratosphere (SPS)**”

Line 405: The authors mention here that the warming in the Antarctica is likely explained by shortwave heating, which does increase. A link to Figure 8 that shows clearly this would be helpful for the reader.

Thanks, we will add the following text here: “**(This increased shortwave heating is more clearly visible in Fig. 8, which plots the shortwave heating on a linear pressure scale with an adjusted shading scale.)**”

line 448 : Are PANs included in NOy as part of organic nitrates?

PANs are indeed included in the WACtl experiment as part of organic nitrates (Emmons et al., 2020). We will add text here to clarify this.

Line 547: The authors mention that in the midlatitudes of both hemispheres, the downward shifts of the clouds result in negative CRE as expected. This is clear for the SH but it is not clearly evident in the case for NH midlatitudes according to Figure 13.

Thank you for pointing this out. We will modify the text here as follows: “In the **SH** midlatitudes **and the NH high latitudes**, the downward shifts of the clouds result in negative CRE as expected.”

References:

Allen, R.J., and Sherwood, S.C.: The impact of natural versus anthropogenic aerosols on atmospheric circulation in the Community Atmosphere Model, Climate Dynamics, 36, 1959-1978, https://doi.org/10.1007/s00382-010-0898-8, 2011.

Emmons, L.K., Walters, S., Hess, P.G., Lamarque J.-F., Pfister, G.G., Filmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T, et al.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), Geoscientific Model Development, 3(1), 43-67, https://doi.org/10.5194/gmd-3-42-2010, 2010.

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J.-F., et al.: The Chemistry Mechanism in the Community Earth System Model version 2 (CESM2). Journal of Advances in Modeling Earth Systems, 12, e2019MS001882. https://doi.org/10.1029/2019MS001882, 2020.

Forster, P.M., Richardson, T., Maycock, A.C., Smith, C.J., Samset, B.H., Myhre, G., Andrews, T., Pincus, R., and Schulz, M.: Recommendations for diagnosing effective radiative forcing from climate models for CMIP6, Journal of Geophysical Research: Atmospheres, 121(20), 12460-12475, https://doi.org/10.1002/2016JD025320, 2016.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G.A., Russel, G., Aleinov, I., Bauer, M., Bell, N., et al.: Efficacy of climate forcings, Journal of Geophysical Research: Atmospheres, 110, D18104, 2005.