Review #1 of "Montreal Protocol's impact on the ozone layer and climate" by Egorova1 et al., 2022, for Atmospheric Chemistry and Physics

GENERAL COMMENTS

This paper describes the climate and atmospheric benefits of the Montreal Protocol and its Amendments and Adjustments simulated using the SOCOLv4.0 Earth System Model. The analysis is relatively straightforward and adds to a body of evidence accumulated through several similar previously published papers. The work will be of interest to readers of ACP. I believe that the paper will be suitable for publication once my concerns, listed below, have been addressed.

SPECIFIC COMMENTS

Line 16: Replace 'ozone-depleting substances' with 'hODSs'. You've defined the acronym so you might as well use it.

We changed it.

Lines 20-21: I think that the abstract should concisely summarise what was learned from this study over and above what is known from similar previous studies. What is new here?

We confirm the main conclusions of previous studies about dramatic ozone loss and substantial climate warming, which support the robustness of the results. We have found some differences in the climate response compared to the model with prescribed ozone changes, which should be further addressed. We demonstrated the influence of MPA limitations on main chemical cycles, cloud amount, precipitation, and sea ice for the first time with the interactive chemistry-climate model. We showed that the climate effects of MPA are comparable with extreme climate warming following the SSP5-85 scenario.

We have added some of our new findings to the abstract.

Line 24: SOCOLv4.0 is a very good Earth System Model (ESM). But it is not state-of-the-art. State-of-the-art suggests that it is the best ESM, defining a standard that all other ESMs should aspire to. This is not the case. The phrase is anyway ridiculously overused. Just delete it.

We have deleted it. However, we note that the reviewer's definition is not commonly accepted. For example, Cambridge Dictionary says: that state-of-the-art means "very modern and using the most recent ideas and methods."

Lines 44-46: I find this sentence confusing. I have not seen any analysis, by anyone, including the Ball et al. papers of 2018 and 2019, that suggest that the ozone layer is not recovering from the effects of ODSs. That's not to say that ozone is everywhere increasing. It is not. But "ozone increasing" is not equivalent to "ozone recovering" otherwise we would be talking about tropospheric ozone increases during the 1980s and 1990s as "ozone recovery". Ozone may be decreasing in some regions of the atmosphere, but that certainly does not mean that

ozone is not recovering from the effects of hODSs. So I have no idea at all why the "recovery and the efficiency of the MPA are now being questioned" as you state. Sure models and observations of, e.g., declining ozone in some regions of the stratosphere may disagree, but that says something about the deficiency of the models, nothing about the deficiency of the MPA.

We have reformulated this sentence. We agree that the MPA efficiency relative to halogen amount is clearly visible, but for the ozone behavior it is not so obvious.

Line 50: Just one point that is worth noting and commenting on regarding "observed meteorological fields". If you run the counter-factual simulation (i.e. the world in which there was no MPA) with observed meteorological fields, then this is not a self-consistent simulation since we know that the meteorological fields would have been different had there been no MPA right? So doing an analysis this way is not a very 'clean' attribution. Using CCMs (or ESMs) for this purpose makes much more sense and provides a further motivation for your work. I think it is worth making that clear here*. Ah yes, I see you make some of these points below.*

The reviewer found the answer in the below text.

Line 83: I think that papers that explore 'the world avoided' because of the Montreal Protocol are interesting. But you haven't made much of a case for why they are required. What is the value in knowing what would have happened if the MPA had not been implemented? As an analogy, I have no desire to know what would have happened if I hadn't decided to come in to the office today. The fact is that I did. The fact is that the MPA was enacted. What is the value of knowing what would have happened had it not been enacted? I think that greater justification is needed for the significant expenditure of compute resources for such studies.

This is rather a philosophical question. One simple answer could be: we need to enlarge our knowledge of what exactly we managed to avoid in order to strengthen our confidence that the measures taken were necessary, effective, and sufficient. Such knowledge will help to prevent similar situations in the future as well as it will contribute to the ongoing dialogue concerning other environmental problems (the greatest being global climate change), by focusing society's attention on how catastrophic anthropogenic consequences could be just in several decades, but also presenting a successful example of the global joint decisions to avoid them. Besides this, the avoided noMPA case presents an interesting example of the atmospheric and climatic state that allows us to learn more about specific processes, i.e. CFC radiative impacts, the behavior of the stratosphere under such an extreme state, and its connection to the troposphere, etc. Such extreme conditions are also a good case to intercompare the models outside their "comfort zone", namely the historical period to which they are tuned. To this date, there are only a couple of other studies that employ the fully coupled chemistry-climate-ocean models to address this topic. In addition, several basic reasons why to perform extreme noMPA cases and spend resources on that are very nicely described in Newman et al., 2009, and we do not want to repeat it but added the sentence with the reference.

Line 124: Nice to see the noMPA_noRadCFC simulation which is very useful for attribution.

Thanks a lot for your compliment.

Figure 2: If this is showing the TCO difference between MPA-ssp245 and noMPA, as in MPA-ssp245 minus noMPA, surely these values should all be positive? It is also stated in the figure caption that "The colored areas mark statistically significant results at the better than 95% level" but *everywhere* is colored. In that case wouldn't it be simpler to say "All differences are statistically significantly different from zero at the 95% level"?

We agree and made changes in Fig.2 caption accordingly.

Line 154: I disagree with this statement "...learn more and verify our knowledge about...". Running models does not add to our knowledge for the simple reason that models contain only what we already know and nothing that we don't know. Running a model can add to our understanding of how some system works and could provide new insights, but no new knowledge.

We rephrased the sentence according to your suggestion. However, we do not completely agree about the model's value. It is true that any model contains what we already know, but what would be the results of nonlinear interaction between processes is frequently surprising.

Figure 3: The figure caption provides no indication as to whether these results are withMPA minus withoutMPA or vice versa. I can certainly guess and I am 99% sure that I would guess correctly, but readers should not have to guess. There appear to be dots in some regions of some panels of Figure 3 but no explanation is given as to what these represent. Finally, it is stated that "colored areas mark statistically significant results at the better than 95% level" but as far as I can see, *everywhere* in every panel of Figure 3 is colored.

We change the Fig.3 caption according to your recommendations. Black dots mean that results are not statistically significant, we added this into the caption.

Line 157: I don't understand this. What is the "earlier periods" that you refer to? and when you say "while the magnitude of the signal is larger", larger than what?

We deleted this sentence.

Line 173: I would have thought that it is due more to the Clausius-Clapeyron relationship rather than enhanced evaporation. For example, climate change could, conceivably, reduce surface wind speeds and evaporation but the warmer troposphere would still have higher water vapour loading because of the Clausius-Clapeyron relationship. If you are certain that it is the result of enhanced evaporation rather, then you need to provide evidential support for that assertion.

We added this suggestion to the text.

We have changed: Tropospheric warming is responsible for the increase of tropospheric water vapor due to enhanced evaporation in a warmer climate to: Tropospheric warming is responsible for the increase of tropospheric water vapor mostly because warmer air can hold more water vapor according to Clausius–Clapeyron law and partially due to enhanced evaporation from the warmer surface.

Line 176: But, presumably, a much clearer troposphere with far less pollution as a result of an increase in the self-cleaning capacity of the troposphere. Given the number of people that die from pollution-induced respiratory diseases every year, this would have a big health benefit right? OK so you die from skin cancer rather, but nonetheless…

Definitely, there will be some health consequences of the changed chemistry, climate, vegetation, and aerosol formation. A holistic estimate of health issues is very complicated and out of our manuscript's scope.

Line 203: I am surprised you have not cited Velders, G.J.M.; Andersen, S.O.; Daniel, J.S.; Fahey, D.W. and McFarland, M., The importance of the Montreal Protocol in protecting climate, Proc. Natl. Acad. Sci., 4814-4819, 104, 2007 here.

We added the reference.

Figure 4: I find the caption very confusing. It says that Young et al. is light blue. I don't see any light blue line in the figure. The legend says that Young et al. is green and I do see a green line. But then it says that Goyal et al. is red but the legend says it is orange. The caption however says that orange is noMPA which the caption says is red. Anyway, as a result, I could make no sense of Figure 4.

Thank you for noticing the inconsistency with colours in Fig.4, we corrected it.

Line 209: With regard to "This disagreement is not easy to understand", how do you know that these difference are statistically significant given the uncertainties and inherent unforced variability in the model simulations?

We added uncertainties, but it did not help to explain the disagreements.

Lines 237-238: I don't see any indication in the rightmost panel of Figure 5 this "except marginally significant cooling in equatorial Africa" that you refer to.

We have deleted this part of the sentence.

Line 239: You need to cite some references here to "Arctic amplification" no matter how well known.

We added a reference to Prevedi et al (2021).

Line 245: Can you are also not providing statistical significance on your results right?

Yes, we can, but the results without statistical significance have no value because the reader cannot distinguish between real signal and noise related to internal variability.

Figure 6, rightmost panel: The figure caption does not make clear whether this is noMPA minus MPA_ssp245 or vice versa. Likewise for Figure 7.

This is not noMPA minus MPA_ssp245 or vice versa for the line plot, these are the differences relative to the first 10 years of the simulation(1980-1990 mean value) for noMPA and MPA_ssp245 experiments. And this is exactly what is written in the caption of Fig.6 and 7.

Line 271: The rate of "1-4% per 1 K warming" that you quote is considerably lower than I have see in other publications, e.g. Sun, Q., Zhang, X., Zwiers, F., Westra, S. & Alexander, L. V. A global, continental, and regional analysis of changes in extreme precipitation. Journal of Climate 34 (1), 243–258 (2021), find values of 6.6% (5.1% to 8.2%; 5%–95% confidence interval) while Westra, S., Alexander, L. V. & Zwiers, F. W. Global increasing trends in annual maximum daily precipitation. Journal of Climate 26 (11), 3904–3918 (2013), found values of 5.9% to 7.7%. You need to provide strong justification for your quoted 1-4% sensitivity.

*We believe that this is a misunderstanding. The reviewer cites numbers that are larger than those presented in our paper. But while the reviewer refers to numbers for extreme precipitation changes, we present in our paper mean precipitation change. Our values agree well with mean precipitation changes from GCMs as for example discussed in Held IM and BJ Soden, Robust responses of the hydrological cycle to global warming. J Clim **19**, 5686–5699 (2006) or Jeevanjee A and DM Romps, Mean precipitation change from a deepening troposphere. PNAS, 115, 45, 11465-11470 (2018).*

Line 302: Again, I would suggest not referring to SOCOLv4.0 as the state-of-the-art ESM.

We have deleted the state-of-the-art.

GRAMMAR AND TYPOGRAPHICAL ISSUES

I understand that the author's first language is not English. The paper would benefit substantially from proof-reading by someone whose first language is English. I apologise for not taking the time to correct the many grammar and typographical errors in the paper. I just don't have the energy today (insufficient chocolate). Some suggested corrections are included below.

Line 47: By 'efficiency' do you mean 'efficacy'? I suspect you do.

We have changed it accordingly.

Line 60: No need to defined hODS again here. You have already defined the acronym twice above.

We have changed it accordingly.

Line 132: You defined the MPA acronym above but now seem to have decided to stop using it.

We have changed it accordingly.

Review #2

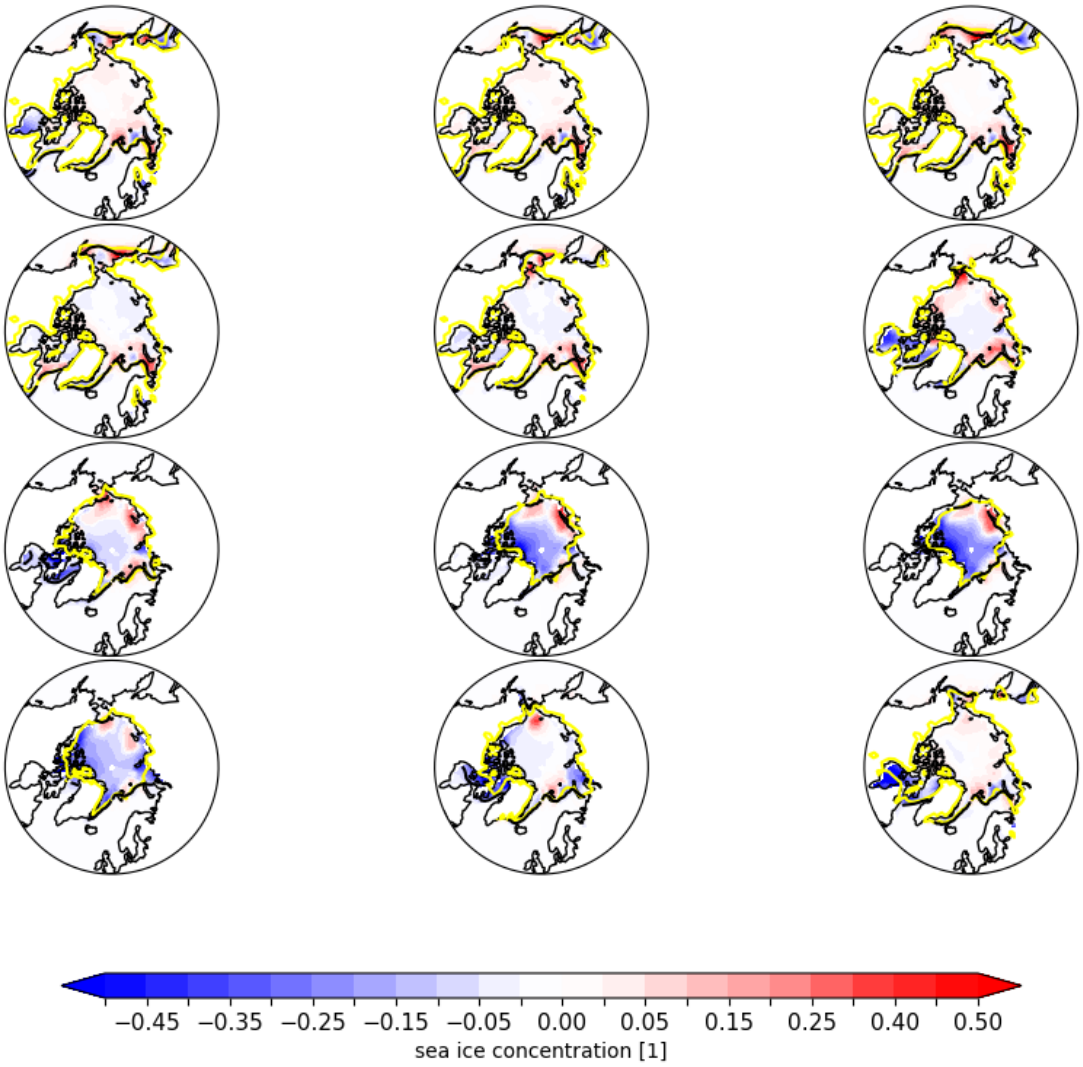
This paper uses the Earth system model SOCOL4 to examine the impact of the Montreal Protocol on ozone, atmospheric chemistry, temperature and surface climate variables. It is straightforward in design, and a useful addition to the literature (e.g., it was noted in the 2018 WMO Ozone Assessment that there had been very few 'world avoided' studies performed since the previous assessment in 2014). I enjoyed reading this paper and recommend publication after the following comments are addressed:

Major

No model validation is done in the paper, instead, the reader is referred to Sukhodolov et al. (2021). While Sukhodolov et al. (2021) evaluated ozone, temperature, water vapour and various other chemical species, they did not evaluate precipitation, cloud cover or sea ice, which are shown in this study. This should be addressed.

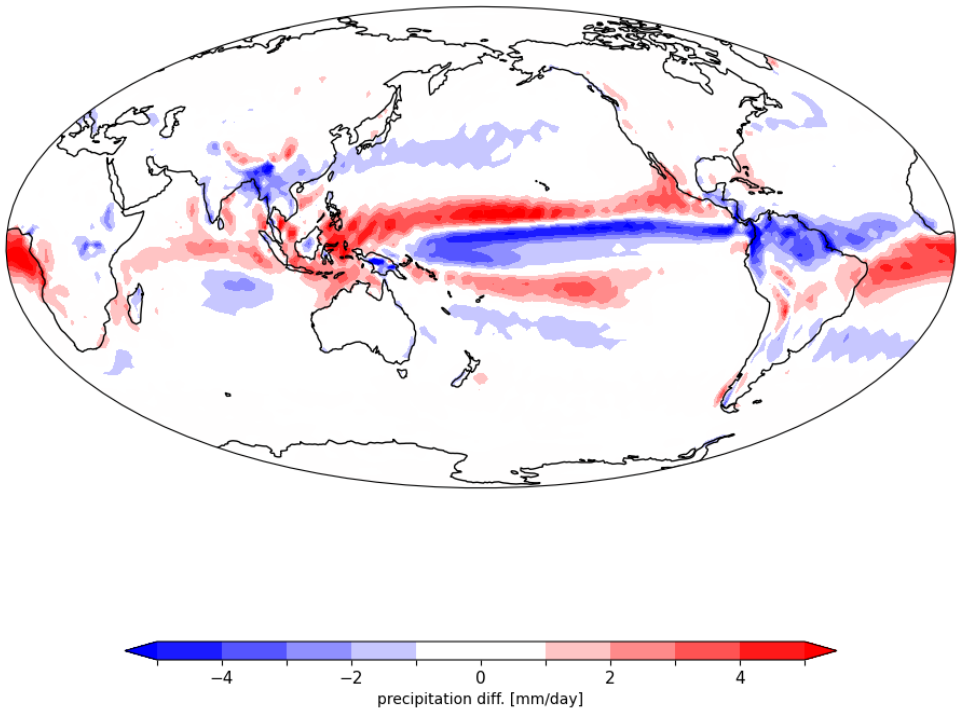
As described in Sukhodolov et al. (2021) the SOCOL version 4 model is based on the MPI-ESM model with ECHAM6 for the atmospheric dynamics and MPIOM as the ocean component. The MPI-ESM model is described in Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2. Journal of Advances in Modeling Earth Systems, 11(4), 998–1038, doi: 10.1029/2018MS001400.

The sea ice component is based on the viscous-plastic rheology and the zero-layer thermodynamic scheme. Both are well-established computational components for sea ice and are used in many climate models. Below we inserted a figure comparing the sea ice extent with ERA-5 data over the period 1990 and 2015. The simulated sea ice edge agrees very well with ERA-5 data. Especially also in the Barents Sea, a place where sea ice models usually have difficulties in modeling the observed sea ice edge. In the summer month, the sea ice concentration is a bit on the lower side in the central Arctic. The SOCOLv4 sea ice agrees well with the MPI-ESM sea ice cover reported in Mauritsen et al. (2019).



Caption: Sea ice cover between 1990 and 2015. The colors indicate the concentration. The yellow line is the ERA-5 sea ice edge and the black line is the simulation sea ice edge.

The annual precipitation simulation vs. observation bias pattern from SOCOLv4 and ERA-5 agrees well with the results from ECHAM6-LR shown in Figure B5 of Mauritsen et al. (2019). In the equatorial region near Australia and in the Indian Ocean, the model shows too much precipitation compared to observation. In the equatorial region close to the Americas, the model is too dry. These regions are known to have some differences compared to observations in many climate models and the SOCOLv4 is no exception here.



Caption: Annually averaged mean precipitation. SOCOLv4 minus ERA-5 precipitation.

The cloud cover is not shown. As the clouds are usually very noisy, and the precipitation pattern is compared to the observation, we don't think that showing the cloud cover brings additional insights into the model results.

In summary, the SOCOLv4 results of sea ice cover and precipitation are very similar to the MPI-ESM results and seem to originate from the dynamical core of the MPI-ESM model.

I was surprised to see almost no mention of polar stratospheric clouds (PSCs) in the analysis. Newman et al. (2009) discuss that the ozone layer collapses as a result of heterogeneous chemical processes. When are PSC formation temperature thresholds reached in the model, and how widespread are PSCs by the end of the simulations?

We mentioned PSC here.

Lines 163-165: In the lower stratosphere, ozone depletion is related to the enhanced liquid aerosol concentration in the Junge layer (Junge and Mandson, 1961), a colder environment, and the presence of polar stratospheric clouds.

Following reviewer's suggestion, we have changed lines 163-165 to:

In the lower stratosphere, almost complete ozone depletion similar to the finding of Newman et al. (2009) is explained by the acceleration of heterogeneous chlorine activation, which is caused by the cooling in this area. The cooling accelerates heterogeneous reactions and enhances surface area density (SAD) due to the additional generation of the liquid sulfate

aerosol in the Junge layer (Junge and Mandson, 1961) and the appearance of the polar stratospheric clouds (PSCs). The obtained increase of annual mean SAD in the noMPA run relative to the reference exceeds 100% in the entire lower stratosphere, except southern high latitudes (not shown). Lower SAD in this area is explained by a substantial decrease of PSC1 due to HNO_3 depletion caused by enhanced chlorine loading.

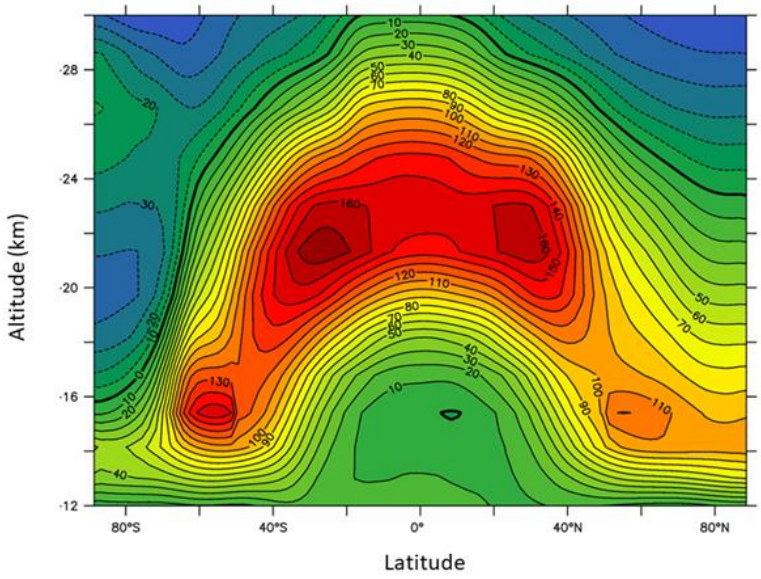


Figure. Annual and zonal mean SAD changes (%) in the noMPA run relative to the reference.

Section 3.3: discussion of studies by Garcia et al. (2012), Goyal et al. (2019) and Young et al. (2021). It would help the reader to give some context as to the models, assumptions and GHG emissions scenarios used. I see that Goyal et al. (2019) used RCP 8.5, so it is not surprising that the temperature response is smaller in SOCOL4 ssp245 simulations (line 217-218).

We have added more information about the above-mentioned models.

The authors seem surprised that warming in the model is weak until 2060 with a sharp decrease thereafter. I see from Figure 6 that cloud fraction drops away sharply after the 2060s and it could be that positive cloud feedbacks have come into play (for more explanation see FAQ7.2 in Chapter 7 of IPCC 2021).

These processes are definitely connected.

The paper would benefit from proofing by a native English speaker; I have recommended a few suggestions under 'specific' comments below.

Minor:

It would be pertinent to cite Young et al. (2021) in the introduction when previous studies that looked at ‘world avoided’ scenarios are reviewed.

Done

Please give more details of how tropospheric chemistry was included (or not) – was it as in Egorova 2003, or as in Figure 2 of Sukhodolov et al. 2021?

As in Sukhodolov et al. 2021

For how long was the model spin-up time?

We added the sentence on model spin-up.

Please give details of how statistical significance testing was done (i.e. which test was used).

We added the method that we used for the calculation of statistical significance.

It would be useful to include a plot showing how hODSs change over time in the MPA_ssp245 and noMPA scenarios. Similarly, a table summarising the different simulations performed would be useful.

hODS behavior is taken from the WMO2018 scenario (World Meteorological Organization (WMO)/United Nations Environment Program (UNEP). Scientific Assessment of Ozone Depletion: 2018 Global Ozone Research and Monitoring Project Report No 58. Geneva, Switzerland: World Meteorological Organization. 588) extended with new data (https://bpb-eu-w2.wpmucdn.com/blogs.reading.ac.uk/dist/7/201/files/2020/09/WMO2018_ESRL_extended.pdf) for MPA_ssp245. For noMPA we applied 3% annual increase of the species under MPA limitations (CFC-11, CFC-12, CFC-113, CFC-114, CFC-115, CCl4, CH3CCl3, HCFC-22, HCFC-141b, HCFC-142b, Halon-1211, Halon-1202, Halon-1301, Halon-2402, CH3Br, and CH3Cl). We do not see the necessity to repeat so many already available figures in our paper.

The table of the simulations has been added.

How do VSLSs get into the atmosphere in SOCOL4? Are they all prescribed as surface mixing ratios, or does the ESM component simulate some of the natural VSLS emissions? (eg sea spray inorganic bromine)

They are prescribed as surface mixing ratio.

lines 76-77: “the fixed tropospheric ozone approach is difficult to justify for the no MPA case due to the expected dramatic increase of UV radiation in the troposphere caused by stratospheric ozone depletion and strong tropospheric ozone radiative forcing.” Given that tropospheric ozone is not fixed in SOCOL4, why do we not see increases in tropospher ozone in Figure 3? Is it because of increases in tropospheric OH?

Yes, we have an up to a 100% increase in tropospheric OH (also shown in Figure 3), which enhances the ozone destruction, plus the free-tropospheric ozone levels are strongly affected by the reduced transport from the stratosphere. We discuss it in section 3.2.

Specific

Line 12-13: suggest rewording as 'Previous research has demonstrated the success of the Montreal Protocol' or similar (as it stands the sentence reads a little oddly).

We change it accordingly.

Line 32: should read 'The ozone layer' not 'the ozone'

We corrected it.

Line 61-62: "in the case of hODS uncontrolled increase their radiative forcing…" this statement needs to be referenced.

We added the reference.

Line 66: "comparable to the effects of carbon dioxide increase" this is vague – how much of a carbon dioxide increase? Over what time frame/under which scenario?

We rephrased the sentence.

Figure 1: please clarify that the 'saved total ozone' is MPA_ssp245 minus noMPA (if that is the case).

We add an explanation to the figure caption of what we meant by saved ozone.

Figure 2: there is no upper or lower part; please add labels (a) and (b). Since all areas are statistically significant (95% level of confidence; ? test) you may as well state that in the caption.

We changed the caption with left and right panels before it was a mistake, thanks for noticing that. Yes, the results are statistically significant everywhere, we added that in the caption.

Line 168: should read 'the secondary maximum in cooling' so that readers do not think you mean a secondary maximum in temperature.

We have added it.

Throughout, please ensure that chemistry is typeset properly, ie $O(^1D)$ not O(1D), the Greek letter Nu (v) not the English letter v when representing photolysis (*hv*), arrows instead of equals signs in chemical reactions.

We checked it and corrected it accordingly.

Figure 3: I don't understand the colorscale used. On the CH_4 plot it appears (from the dark blue shading) as though there are large decreases in tropospheric methane, but the contour

lines say it is only ~-3%, while larger decreases (shaded white) are visible in the upper stratosphere. In contrast temperature, for example, uses dark blue to show large cooling and light blue to show weaker cooling. Also, the text colour on the contours is very hard to read over dark red and blue. Please ensure you state (here or elsewhere) how Cl_y, NO_x, NO_y are defined.

We have replotted this figure.

The caption should state explicitly that results from the noMPA minus MPA_ssp245 simulation are shown. What do the black dots show? (I assume where the difference is not statistically significant, but this needs to be defined as it is inconsistent with the caption).

We have corrected the caption according to the reviewer's suggestion.

Finally, what does 'st.' height refer to on the y axis?

'st.' height refers to the standard height, it is used sometimes also, but we change it to altitude in km.

Figure 4: How were the anomalies calculated?

They are calculated as differences between the results of the experiment and reference runs.

The coloured lines stated in the caption are inconsistent with what is shown on the legend (e.g MPA-ssp585).

Thank you for noticing that. We have replotted the figure with more distinguished colours and corrected the caption, now it is consistent with the legend.

Figure 6: do you show cloud volume fraction or cloud area fraction? (I think the latter?) In the plot on the right, why not be consistent with the plot on the left and report the difference in %?

We changed the figure and now show everything in percent.