

## Response to Reviews of “Co-emission of volcanic sulfur and halogens amplifies volcanic effective radiative forcing” by Staunton Sykes et al.

We are very grateful to Daniele Visoni, an anonymous reviewer and Alan Robock for their comments and efforts which have helped us improve this manuscript. Following the structure recommended by ACP, we have responded to each reviewers' comments sequentially below with italicised and underlined text showing the reviewer's comments and plain text showing our response. Text which has been added to the manuscript is coloured red. Original manuscript text is in blue and any text which has been removed from the manuscript is blue and has been struck through. The locations of changes are stated. We hope these revisions address the comments of the reviewers.

### Review 1

*In this study the authors, using a CCM model, try to understand how co-emitted halogens may alter the climate impacts of stratospheric sulfur injection by volcanic eruptions. Their results show that the inclusion of halogens dramatically changes some of the simulated impacts, and that including such emissions is crucial to properly simulate large explosive volcanic eruptions. I found the paper to be excellent: the introduction does a great job at framing the problem and the methods are properly described. The results are also explained clearly, with pretty straightforward figures (which I suggest uploading in a higher quality format, they look a bit blurred when zoomed in). Overall, the manuscript is perfectly suitable to be published in ACP.*

We thank Daniele Visoni for his comments and are pleased he assessed the work so highly. We acknowledge the suggestion to upload higher quality figures and have ensured this will be done before final submission.

### Specific Comments from Daniele Visoni:

L 69: highlight with no s

Done.

L 93: The correct name of the model is CESM1(WACCM) (Community Earth-System Model and then WACCM)

The sentence beginning on line 93 has been amended accordingly to read:

Lurton et al. (2018) simulated the 2009 Sarychev Peak eruption (0.9 Tg of SO<sub>2</sub>) in CESM1(WACCM) (Community Earth-System Model, The Whole Atmosphere Community Climate Model (~~WAGCM~~) and showed how inclusion of halogens...

L 116: from the way the phrase is written, it looks like the different halogen emissions are applied to the same amount of SO<sub>2</sub>, but in that case, the ratio would also have to

be 100 times. But Ming et al. 2020 compare also a low and high SO<sub>2</sub> injection (10 vs. 100 Tg) and in there different HCl concentrations. Just try to clarify this point.

The sentence beginning on line 116 has been amended accordingly to read:

They simulated 6 sets of experiments: a low SO<sub>2</sub> (10 Tg) and high SO<sub>2</sub> (100 Tg) eruption each paired with no HCl, low HCl (0.02 Tg) and high HCl (2 Tg), and found reported significant ozone depletion over both poles for at least four years in the high SO<sub>2</sub> and high HCl experiment. ~~that a volcanic halogen emission of 0.02 Tg (HCl:SO<sub>2</sub>=0.04) into a pre-industrial background state had little impact on column ozone but 2 Tg (HCl:SO<sub>2</sub>=0.4) showed significant and prolonged ozone depletion above both poles.~~

Fig. 1: please specify what is the quiescent period against which the anomalies are calculated in the caption.

We have amended the Figure 1 caption as follows:

Figure 1 - Global evolution of sulfur, halogens and OH for the SULF56, HAL56, SULF10 and HAL10 simulations, relative to the control climatology.

We have also amended the manuscript text in Section 2.2 Experimental Design to better explain which period the perturbations compared, as follows:

We utilise atmosphere-only, time-slice experiments whereby the initial sea surface temperature, sea ice fraction and forcing agents and depth, surface emissions and lower boundary conditions are prescribed using climatologies calculated using data from the fully coupled UKESM1.0 historical runs produced for CMIP6 (Eyring et al. 2016) and averaged over the years 1990 to 2000. By averaging over the decade the atmosphere-only simulations are forced with lower boundary conditions typical of the recent historical period but not a specific date within that decade, as desired. The fully coupled transient simulations had internally generated El Nino and La Nina cycles, however, averaging the SSTs over the 1990 to 2000 period resulted in a permanent neutral signal in the SST pattern, see figure S1. The 1990s, and thus these timeslices, were characterised by high background halogen levels due to anthropogenic emissions of CFCs throughout the preceding decade. The impacts of very short lived Bromine species are accounted for by adding a fixed contribution of 5 pptv into the CH<sub>3</sub>Br lower boundary condition.

~~A control simulation was run with a 15 year spin up followed by a further 20 years.~~ A control simulation was initialised from the January 1995 initialisation file taken from the UKESM1.0 historical scenario which was run as part of CMIP6 (Eyring et al. 2016). The model was allowed to spin up for 15 years and the control was run for a further 20 years. The effect of the different explosive volcanic eruption scenarios (SULF10, SULF56, HAL10, HAL56) was investigated by running 6 10-year volcanic perturbation simulations for each scenario. The 6 simulations were initialised from 6 different years in the control run to represent the variability in QBO states. Changes are plotted as the difference between the average of the 6 ensembles and a climatology derived from the 20-year control run, cumulative forcings are calculated as the sum of the forcing over the full 10-year simulation duration.

Fig. 1f: I'm a bit confused as to why in the SULF simulations, there is a small increase in OH that I don't think is properly explained in the text. In the SO<sub>2</sub> plume, we expect a large OH depletion. I assume that can be balanced out by the influx of water vapor in the stratosphere from the lower stratospheric heating and produce globally a slight increase. But I'd suggest checking (or consider the tropical changes in stratospheric OH, where I'm sure the change is negative – albeit less than in the HAL experiments).

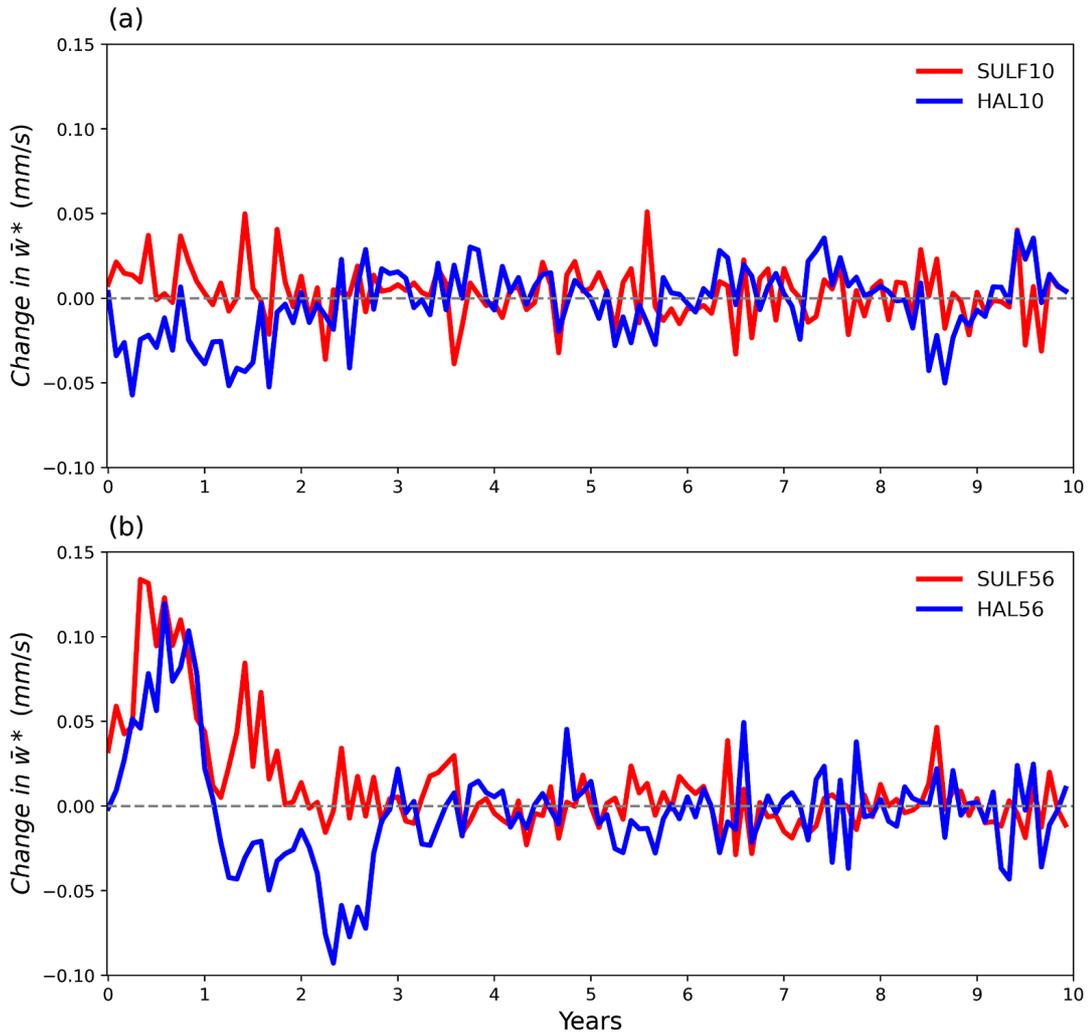
In the initial submission, Figure 1f showed the global stratospheric mean OH percentage change. In sulfur-only simulations this was shown to increase due to the influx of water vapour into the stratosphere as a result of the tropical tropopause changes explored later in the paper. We have amended Figure 1f to show the tropical (20°S-20°N) changes in stratospheric OH, which better shows the depletion of OH due to SO<sub>2</sub> oxidation.

The accompanying caption has been amended accordingly to read:

(f) Tropical (20°N-20°S ) sStratospheric OH change (%).

Line 255: it would be useful to show the changes in  $w^*$  (maybe in the supplementary next to Fig. S1) to show the difference in the transport induced by the stratospheric heating.

We have included an additional figure in the SI, showing a time series of the change in tropical (20°S-20°N)  $\overline{w^*}$  (residual mean vertical velocity) at 50 hPa (Figure S2).



**Figure S2.** Time series of 20°S - 20°N mean change in  $\bar{w}^*$  (residual mean vertical velocity) at 50 hPa.

*Fig. 3: I'd suggest switching panels a and b, as logically one might expect the lower injection scenario before.*

We agree with the reviewer that showing the smaller injection scenario first makes more logical sense. All plots and text have been amended accordingly.

*Also, I find it interesting that the relationship doesn't hold as well for the lower injection case. I suspect this might be due to different QBO phases that affect the aerosols e-folding time (see Pitari et al., 2016), and that this effect is more evident for lower injection rates while for higher injection rates the increase heating rates modify the QBO too strongly independently on the phase it's in at the moment of injection (see for instance Aquila et al., 2014) resulting in similar lifetimes. The authors could just check if that's the case verifying the QBO phase, or just mention that's a possibility for the lower correlation in panel b (unless they have a better explanation).*

SULF10, HAL10, SULF56, and HAL56 each have 6 ensemble simulations, initialised from the same 6 July initialisation files from the control simulation. Three of these initialisation files had a westerly QBO phase and the other three have an easterly QBO phase.

The correlation coefficient for regression lines is very high in both the large ( $r=0.88$ ) and very-large eruption scenario ( $r=0.95$ ). However, we acknowledge that the  $r$  value is slightly lower in the 10 Tg eruption scenarios, but suggest this is likely due to a smaller signal to noise ratio.

Fig. 8: please specify if panels a-d are global changes

The caption accompanying Figure 8 has been amended to read:

Figure 8 Evolution of **global** stratospheric mean water vapour (ppmv) in SULF56 and HAL56 (a), and SULF10 and HAL10 (b). Evolution of **global** stratospheric methane (ppmv) in SULF56 and HAL56 (c), and SULF10 and HAL10 (d).

## Review 2

### General Comments:

*This study considers hypothetical VEI 6 and VEI 7 sized eruptions and, using a coupled chemistry-aerosol model driven by scenarios including or not including halogen injection, investigate how the co-emission of volcanic sulfur and halogens alters the evolution of the volcanic aerosol plume, stratospheric ozone chemistry, and the resulting radiative forcing and UV flux. The authors investigate how volcanic halogens may interact with the sulfur aerosol life cycle and interact to modulate volcanic forcing conversely to previously reported work. I found the link between chemical and microphysical processes of particular interest. The authors reveal in their model experiments the primary importance of halogens in major volcanic emissions in the sulfur cycle in the stratosphere, a process already suspected for eruptions of much more minor amplitude. Impacts of halogen emissions on dynamics of the aerosols and the subsequent effect on aerosol microphysics are also considered here when critically missing in previous reported studies. Effects on some key stratospheric compounds like ozone, water vapour and methane are also analysed. To me, this work points at the critical need to maintain space-borne observations of stratospheric compounds which will be particularly valuable to quantify halogen injected by volcanoes and needed for model initialization. The authors finally mention open questions to be addressed in future studies reflecting the great interest to consider these events and their associated various injections (sulfur, halogens, ash, water vapour) in the future climate. This study is original and comprehensive judging by the various topics and impacts covered (microphysics, dynamics, chemistry and radiative forcing). I found the manuscript clear, well-written and nicely going straight to the obtained results. I estimate that this work deserves to be published in ACP after the following minor comments have been addressed.*

We are glad that the reviewer found the manuscript clear and concise, while remaining original and comprehensive. We are thankful for their useful and constructive comments which have helped to improve this manuscript.

### Specific Comments from Reviewer 2:

Introduction: I am not a specialist of petrological processes but could you indicate the degree of uncertainty when petrological budgets are used to derive stratospheric inputs of halogens? What are the assumptions behind the halogen injection efficiency (I would suggest to briefly recall the definition of the stratospheric halogen injection efficiency).

The stratospheric halogen injection efficiency is the fraction of the halogens degassed from the magma at the vent that are transported into the stratosphere. Textor et al. (2003b) used plume rise models to suggest that the halogen injection efficiency of explosive volcanic eruptions is 10 to 20%. We calculated halogen injection efficiencies for past eruptions using petrological data as an estimate of the amount of halogens degassed at the vent and then used ice core and satellite data to estimate the amount that made it into the stratosphere.

P5 lines 158-168: How the overall chemical species initialized? Is it based on climatological 3D fields or only surface emissions provided by CMIP6? Especially for bromine compounds, how the Bry budget initialized? Are very-short-lived species accounted for? The resulting inorganic bromine budget in the stratosphere and more generally the inorganic halogen content, computed in chemistry models is of particular importance regarding ozone chemical cycles and would have significant impact in your scenario with no volcanic emissions of halogens. Please provide a bit more information.

All chemical species in our model set up are initialised from the 3D January 1995 initialisation file from the UKESM1.0 historical scenario which was run as part of CMIP6. During the simulations, surface emissions of BC, C<sub>2</sub>H<sub>8</sub>, C<sub>5</sub>H<sub>8</sub>, CO, DMS, HCHO, Me<sub>2</sub>CHO, Monoterp, NH<sub>3</sub>, NO, NVOC, OC and anthropogenic SO<sub>2</sub> are included following the CMIP6 UKESM1.0 historical scenario and the concentrations of CO<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>, N<sub>2</sub>O, LL<sup>Cl</sup> and LL<sup>Br</sup> are specified as surface concentrations following the CMIP6 UKESM1.0 historical scenario (Eyring et al., 2016). The impacts of VSL<sup>Br</sup> are accounted for by adding a fixed contribution of 5pptv into the CH<sub>3</sub>Br surface concentration.

#### **Where:**

Long lived Chlorine (LL<sup>Cl</sup>) = CH<sub>3</sub>Cl, CH<sub>3</sub>CCl<sub>3</sub>, CCl<sub>4</sub>, CFC-11, CFC-12, CFC-113, CFC-114, CFC-115, HCFC-141b, HCFC-142b, HCFC-22

Long lived Bromine (LL<sup>Br</sup>) = CH<sub>3</sub>Br, H-1301, H-1211, H-1202, H-2402

Very Short Lived Bromine (VSL<sup>Br</sup>) = CHBr<sub>3</sub>, CH<sub>2</sub>Br<sub>2</sub>, CH<sub>2</sub>BrCl, CHBrCl<sub>2</sub>, CHBr<sub>2</sub>Cl, CH<sub>2</sub>IBr

Section 2.2 has been amended, as follows:

We utilise atmosphere-only, time-slice experiments whereby the initial sea surface temperature, sea ice fraction and forcing agents and depth, surface emissions and lower boundary conditions are prescribed using climatologies calculated using data from the fully coupled UKESM1.0 historical runs produced for CMIP6 (Eyring et al. 2016) and averaged

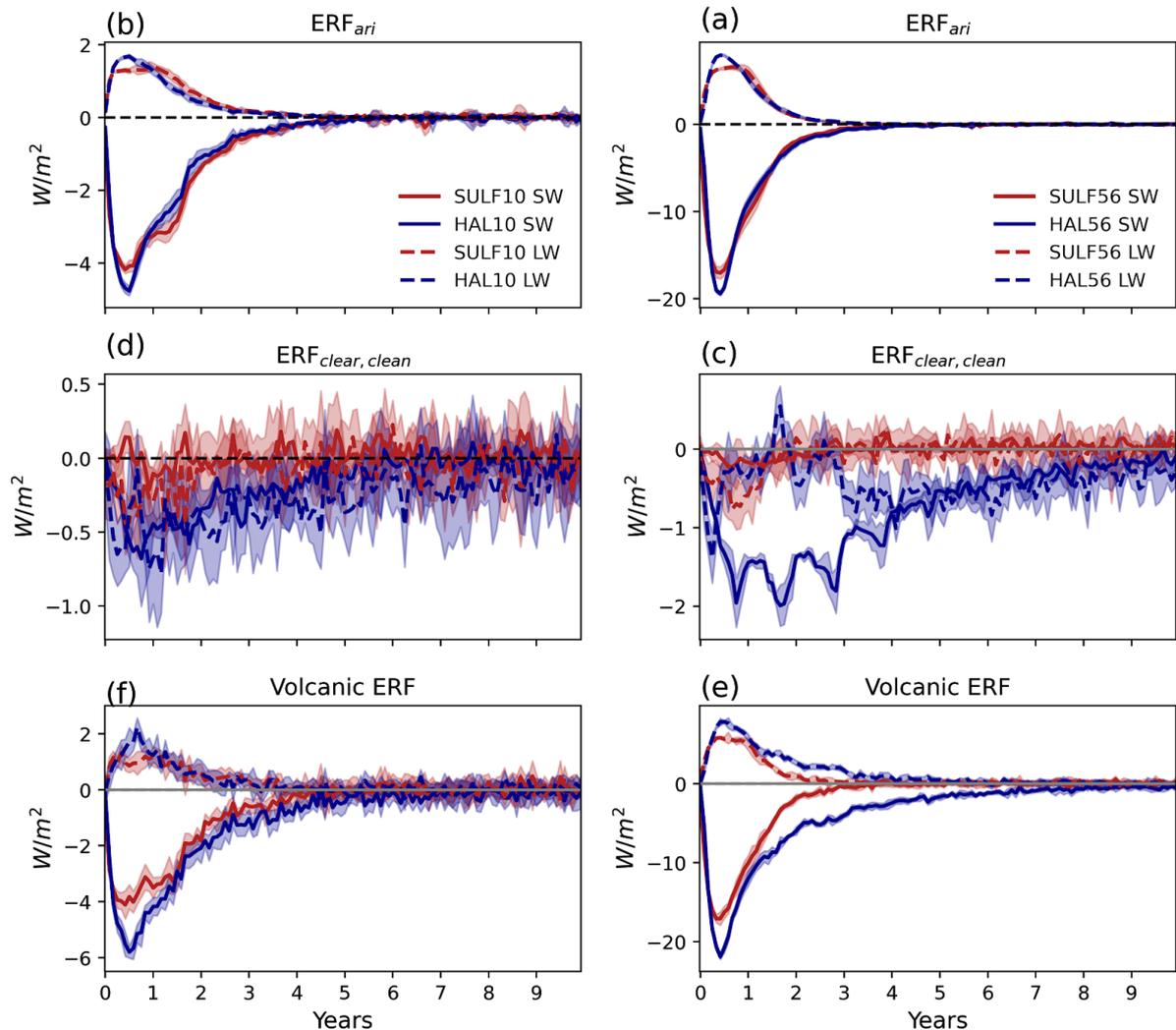
over the years 1990 to 2000. By averaging over this timeframe the atmosphere-only simulations are forced with boundary conditions typical of the recent historical period but not a specific date within that decade. The fully coupled transient simulations had internally generated El Niño and La Niña cycles, however, averaging the SSTs over the 1990 to 2000 period resulted in a permanent neutral signal in the SST pattern, see figure S1. The 1990s, and thus these timeslices, were characterised by high background halogen levels due to anthropogenic emissions of CFCs throughout the preceding decade. The impacts of very short lived Bromine species are accounted for by adding a fixed contribution of 5 pptv into the CH<sub>3</sub>Br surface concentration.

~~A control simulation was run with a 15 year spin up followed by a further 20 years. A 20-year control simulation was run after a 15 year spin up, initialised from the January 1995~~ initialisation file taken from the UKESM1.0 historical scenario which was run as part of CMIP6 (Eyring et al. 2016). The effect of explosive volcanic eruptions was investigated by running a series of 10 year volcanic perturbation simulations spun off from 6 different years in the control run to represent the variability in QBO states. Changes are plotted as the difference between the average of the 6 ensembles and a climatology derived from the 20 year control run, cumulative forcings are calculated as the sum of the forcing over the full 10 year simulation duration.

*P5 line 170 onwards: You do not differentiate between SW and LW radiation conversely to the work of Schmidt et al. (2018). What justifies this choice?*

We have included an additional figure into the SI showing a breakdown of the TOA ERF<sub>ari</sub>, ERF<sub>cc</sub>, and total volcanic ERF anomalies into the contributing shortwave and longwave changes (Figure S4), and amended the manuscript text as follows:

The radiative impact of sulfate aerosols depends on the particle size, amongst other things (Timmreck et al., 2010, Pinto et al. 1989). Using Mie scattering theory, (Lacis, (2015) found that the scattering cross section per unit mass at 550 nm is largest for sulfate aerosol with effective radius of ~0.20 μm. The smaller Reff in HAL10 and HAL56, compared to SULF10 and SULF56, is closer to 0.20 μm and results in more efficient scattering of SW of radiation per unit mass (Timmreck et al., 2010). Therefore, we simulate 11% and 22% higher peak global-mean stratospheric aerosol optical depth (SAOD) anomalies at 550 nm in HAL10 and HAL56 than their equivalent SULF simulations (Figure 4), despite having a 14% and 9% lower smaller peak aerosol burden. Correspondingly, we simulate an 8% and 6% increase in the peak global-mean ERF<sub>ari</sub> in HAL10 and HAL56 compared to SULF10 and SULF56 (Figure 4), driven by a 14% and 11% increase in peak global-mean SW forcing (Figure S4). The SAOD and ERF<sub>ari</sub> anomalies are a balance between the offsetting effects of smaller aerosol and shorter lifetime which result in a net-zero impact on cumulative ERF<sub>ari</sub> despite a significant increase in the peak global-mean ERF<sub>ari</sub> (Figure S2a,b).



**Figure S4.** Evolution of global mean SW and LW TOA  $ERF_{ari}$  (top),  $ERF_{cc}$  (middle), and  $ERF$  (bottom) flux anomalies.

P5 line 188: why variations in the surface albedo were not taken into account? As you state in the manuscript, the model can be forced by surface boundary conditions. Is it for calculation-time issues?

This work is the first paper to investigate the role co-emitted halogens may have on the ERF of volcanic eruptions. We utilised an atmosphere only model set up with fixed SSTs and sea ice to ensure the experiments were systematic, allowing us, in the first instance, to understand the atmospheric chemistry and radiative forcing response (Forster et al., 2016). Running the fully coupled transient UKESM model would make it harder to disentangle the drivers of the forcing and would involve a significant step up in computational cost. Although this was not feasible in this work, we have updated the future work section to reflect that this is an important next step.

P6 line 202: An injection altitude distributed around 21 km has been chosen. Volcanic impact depends on injection altitude especially because the residence time of aerosols is linked to this parameter. I am aware that strict choices must be done for

costly long term simulations but what justifies this value? Did the authors conduct sensitivity tests on this parameter?

In the interest of keeping the model calculation time to a reasonable level, we decided to introduce the volcanic species into the model centred at a single fixed altitude. We chose an altitude of 21km to be consistent with large historical eruptions from the satellite era. Introducing the volcanic species at 21 km results in an aerosol cloud at an altitude of ~25km after aerosol self lofting which is consistent with 1991 Pinatubo and 1982 El Chichón (Winker and Osborn, 1992, DeFoor et al., 1992).

This model set up was designed to compare the impact of co-emission and sulfur-only eruption scenarios. Although we did not conduct sensitivity tests on the injection height parameter, we utilised the exact same eruption source parameters in both the sulfur-only and co-emission scenarios allowing the simulations to be directly compared, and the differences isolated.

Section 2.2 has been amended as follows:

The volcanic emissions are prescribed by direct injection of SO<sub>2</sub>, HCl and HBr into the stratosphere with a Gaussian plume vertical distribution centred on 21 km, lasting for 24 hours on July 1st. An injection altitude of 21 km was chosen as, allowing for lofting, this results in a volcanic plume altitude consistent with recent historical eruptions from the satellite era (Guo et al., 2004). The gases were injected in the tropics (5°S latitude and 0° longitude) to represent a typical tropical explosive eruption (Newhall et al., 2017).

P6 lines 206-219: there is a lot of assumption behind the stratospheric halogen injection efficiency. The values given by Textor et al. (2013) strongly differ from other reported studies. Is this factor highly variable from one eruption to another? Why El Chichon and Mazama eruptions reassures the numbers taken for HAL56? For HAL10 it is not clear to me why the Pinatubo HCL:SO<sub>2</sub> molar ratio must match the one for Mt Mazama. I guess petrological processes somewhat differ for these events. Please clarify.

Two factors determine the amount of halogens that an eruption injects into the stratosphere: the amount of halogens degassed from the magma and the fraction of emitted halogens that reach the stratosphere (here after halogen injection efficiency). The amount released from the magma is variable and depends on the geochemistry of the volcano (beyond the scope of this paper). The halogen injection efficiency is also variable and is dependent on the specific eruption conditions, Textor et al. (2003) used complex model studies to estimate the stratospheric halogen injection efficiency of explosive eruptions to be 10 to 20%, whereas, the halogen injection efficiency was shown to be 8.1% following ~7.6 kya Mount Mazama (Bacon et al., 1992), ~2.5% following 1982 El Chicon, and ~0% following 1991 Mount Pinatubo (Mankin and Coffey, 1984; Woods et al., 1985). What is clear from these studies, however, is that significant concentrations of halogens are injected into the atmosphere following some explosive volcanic eruptions. Given this uncertainty and variance, we designed a simulation matrix that spanned a range of possible explosive volcanic emissions scenarios. In section 2.2, we compared this simulation matrix to the stratospheric halogen injection efficiency and HCl:SO<sub>2</sub> ratios of past eruptions and other studies in order to

contextualize our emissions scenarios within the prior literature. We concede that this could have been explained clearer in the manuscript and have amended the aforementioned section to read:

Since historical stratospheric volcanic SO<sub>2</sub> fluxes are variable and the volcanic flux of HCl and HBr into the stratosphere remains uncertain, we developed a simulation matrix that spans a range of possible explosive volcanic emissions. The four sets of experiments have one **large high** SO<sub>2</sub> (56 10 Tg), and one **very large low** SO<sub>2</sub> (40 56 Tg) emission scenario both with (HAL5610 and HAL4056) and without halogens (SULF5610 and SULF4056), as shown in Table 1. These eruption sizes (1056 and 5640 Tg SO<sub>2</sub>) are similar in size to a typical VEI 7 (e.g. 1257 Mt. Samalas) and VEI 6 (e.g. 1991 Mt. Pinatubo) and VEI 7 (e.g. 1257 Mt. Samalas) eruption, representing 1 in 500 -1000 year and 1 in 500-1000 year events respectively.

HAL56 utilises the 1257 Mt. Samalas HCl and HBr emission estimates from Vidal et al. (2014) and assumes a conservative ~5% stratospheric halogen injection efficiency, less than the 10-20% predicted by plume modelling in Textor et al (2013) and closer to the observed efficiency following El Chichón (>2.5%) and in the ice core record of Mt. Mazama (8%), as well as the fraction supported by Wade et al. (2020). HAL10 has a SO<sub>2</sub> injection similar to that found to reproduce the SAOD following 1991 Pinatubo (Mills et al., 2016) and a 10 times smaller flux of HCl and HBr than HAL56. This results in a HCl:SO<sub>2</sub> ratio of ~0.26 and ~0.47 in HAL10 and HAL56 respectively, similar to the estimated stratospheric injection ratio for Mt. Mazama (0.3) (Zdanowicz et al., 1999) and the ratios used in Ming et al. (2020), and Brenna et al. (2020) but smaller than the ratio used in Cadoux et al. (2015). HAL10 has a SO<sub>2</sub> injection similar to that found to reproduce 1991 Pinatubo in Mills et al. (2016) and a 10 times smaller flux of HCl and HBr than HAL56, resulting in a HCl:SO<sub>2</sub> molar ratio of ~0.26, very close to the estimated stratospheric injection ratio for Mt. Mazama (0.3) (Zdanowicz et al., 1999).

*P8 figure 1d: what is the reason for the overlapping HAL and SULF S global burden at an early stage of the simulations? This surprises me because this feature is not visible on SO<sub>4</sub> which already shows a marked difference over the first months (figure 1c).*

The overlap in the global burden in the first 6 months of the HAL and SULF simulations seen in the total sulfur burden is caused by the offsetting effects that co-emission of halogens has on SO<sub>2</sub> and SO<sub>4</sub> burdens. Co-emission of halogens results in slower oxidation of SO<sub>2</sub> to SO<sub>4</sub>, resulting in a slower decline in SO<sub>2</sub> and a corresponding slower increase in SO<sub>4</sub>. The net effect is a very similar total S trend in HAL and SULF simulations over the first few months.

*P11 lines 268-279: The investigation of Reff is interesting since it provides a (integrated) description of the impact of HAL scenario on aerosol sizes. However, it would have been also valuable to examine more comprehensively the impact on microphysics. Although GLOMAP is a modal microphysical module (as far as I understood) did the authors get information about the effects on size distributions (geometrical standard deviation, total concentration)? For instance, concentrations might be reduced if particle sizes increase but with different ratios. Concentration*

(well, the whole size distribution) is also important for the SAOD and ERF calculated in the manuscript (figure 4).

In GLOMAP-mode, the mode geometric standard deviations ( $\sigma_g$ ) are constant whilst the geometric mean diameter  $D$  can vary between the size ranges shown for each mode (see Table below). Particle number and mass are transferred between modes when  $D$  exceeds the upper limit for the mode, referred to as mode-merging (Mann et al., 2010).

**Table 3.** Standard aerosol configuration for GLOMAP-mode.

| Index | Name       | Size range                       | Composition         | Soluble? | $\sigma_g$ |
|-------|------------|----------------------------------|---------------------|----------|------------|
| 1     | Nucl-sol   | $\overline{D} < 10$ nm           | SU, POM             | Yes      | 1.59       |
| 2     | Aitken-sol | $10 < \overline{D} < 100$ nm     | SU, BC, POM         | Yes      | 1.59       |
| 3     | accum-sol  | $100 < \overline{D} < 1$ $\mu$ m | SU, BC, POM, SS, DU | Yes      | 1.59       |
| 4     | coarse-sol | $\overline{D} > 1$ $\mu$ m       | SU, BC, POM, SS, DU | Yes      | 2.0        |
| 5     | Aitken-ins | $10 < \overline{D} < 100$ nm     | BC, POM             | No       | 1.59       |
| 6     | accum-ins  | $100 < \overline{D} < 1$ $\mu$ m | DU                  | No       | 1.59       |
| 7     | coarse-ins | $\overline{D} > 1$ $\mu$ m       | DU                  | No       | 2.0        |

In theory, examining the full aerosol size distribution is possible, however, to do this comprehensively is beyond the scope of this study and would substantially increase the length of the manuscript. The differences in the aerosol  $R_{\text{eff}}$  give an integrated view across the aerosol modes and clearly explains the differences in volcanic forcing between SULF and HAL simulations.

*P11 lines 268-279: I think a short comparison with maximum Reff values reported for different past eruptions (Pinatubo in particular) would be interesting to include here.*

We have included a short comparison of observed  $R_{\text{eff}}$  values following 1991 Pinatubo from balloon and satellite observations as follows:

The maximum global mean  $R_{\text{eff}}$  was 0.38  $\mu$ m and 0.59  $\mu$ m in SULF10 and SULF56, respectively. The maximum global mean  $R_{\text{eff}}$  simulated in SULF10 is similar to that derived from measurements following 1991 Pinatubo, with an estimate of 0.4 - 0.5  $\mu$ m from balloon borne measurements (Deshler et al., 1997) and 0.45  $\mu$ m obtained from GLOSSAC satellite observations (GloSSAC, version 1.1; (Thomason et al., 2018)). The shorter total sulfur e-folding lifetime of sulfur following HAL10 and HAL56 eruptions results in reduced aerosol growth and smaller aerosol effective radii ( $R_{\text{eff}}$ ). Peak global-mean  $R_{\text{eff}}$  is ~15% and ~10% smaller in HAL10 and HAL56 compared to their equivalent SULF simulations (Figure 1e). The reduced aerosol growth is a direct result of the shorter sulfur lifetime, rapid spreading and removal of aerosol. Volcanic sulfate aerosols grow through microphysical processes of condensation and coagulation (Kremser et al., 2016). The faster removal of sulfate aerosol in HAL10 and HAL56 reduces the growth via condensation and coagulation and results in smaller peak global-mean aerosol  $R_{\text{eff}}$ . This theory is supported by Figure 3 which shows a scatter plot of 3-year global-mean aerosol effective radius as a function of the global sulfur burden e-folding time with a significant correlation within both 10 Tg ( $r=0.88$ ) and 56 Tg ( $r=0.95$ ) eruption ensembles. The positive correlation between these two variables holds only for each eruption size scenario. The larger SO<sub>2</sub> injection in HAL56 and SULF56 leads

to larger-sized sulfate aerosols, faster sedimentation and shorter removal time compared to HAL10 and SULF10, as seen by comparing Figures 3a and 3b.

P11 line 291: please add the wavelength here.

Done

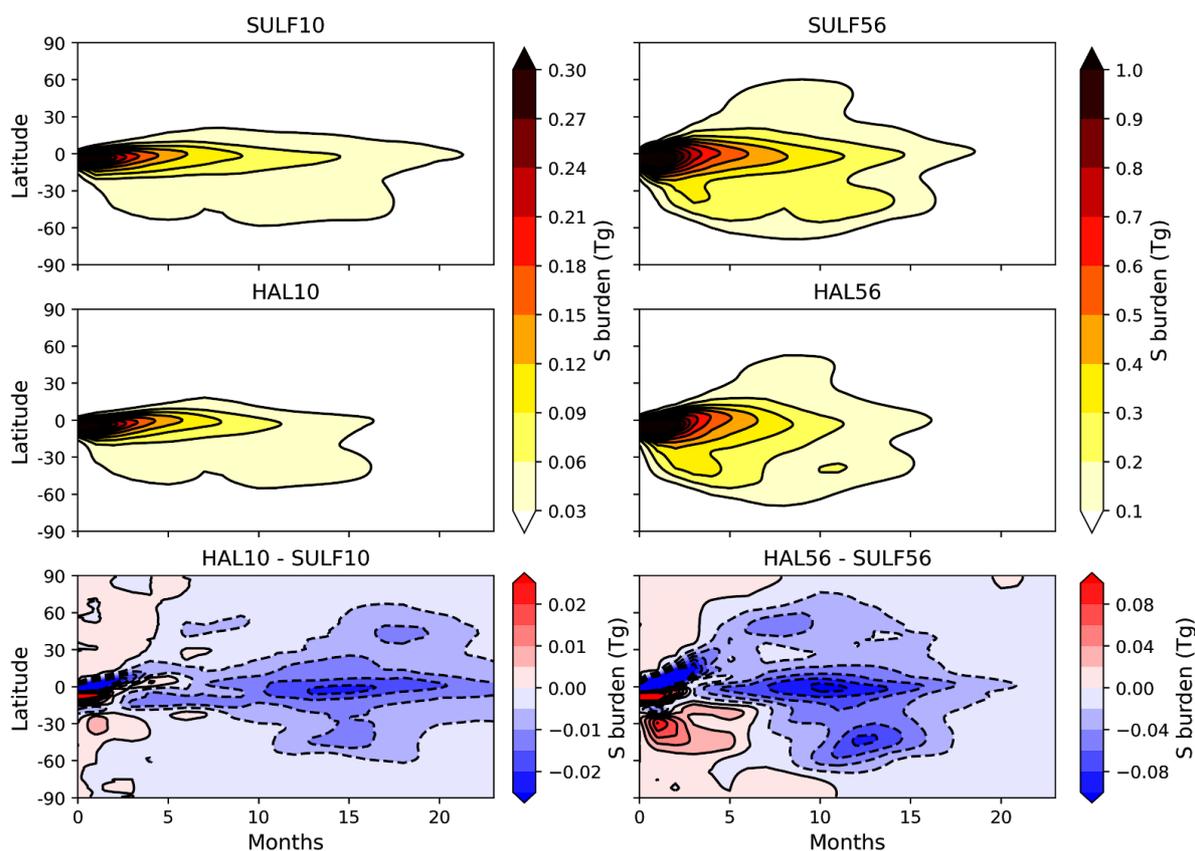
P12 figure 4: the color coding (red for SULF and blue for HAL) is the opposite than in previous figures. I think it would be preferable to homogenise this.

Done

P14 figure 6: A latitudinal transport is visible in the lower stratosphere, particularly for HAL10 simulation. What process can be related to the hemispherical difference? Since the plot is integrated over 3-years I am not sure that the dominating phase of the QBO (which has been shown to impact volcanic aerosol transport from the tropics) can be an explanation.

This hemispherical difference arises because the volcanic sulfur and halogens are injected just south of the equator, meaning that they are predominantly dispersed into the Southern Hemisphere. We have added two figures into the SI document to better show how the injected species are dispersed, a latitude-time evolution (figure S4) and an altitude-latitude profile for the first five post eruption months (S5), and we have amended the manuscript text as follows:

Volcanic halogen catalysed ozone depletion was simulated across all latitudes, but the largest magnitude changes in HAL10 (-40%) and HAL56 (-80%) were within the aerosol cloud and the polar regions, where the co-emitted halogens were activated on aerosol surfaces and PSCs respectively (Figure 5). Ozone depletion predominantly occurs in the tropics between 25 and 30 km in the first post eruption year, with depletion maxima of -3.5 ppmv and 6 ppmv in HAL10 and HAL56 respectively (Figure 6). By year 3, the ozone depletion shows a bimodal altitude distribution in the stratosphere similar to that found in Brenna *et al.* (2020), with depletion maxima both in the lower (20 km) and upper (40 km) stratosphere. As the volcanic SO<sub>2</sub> and halogens were introduced into the stratosphere just south of the equator, they were predominantly dispersed into the southern hemisphere (Figure S6), leading to larger ozone depletions compared with the northern hemisphere. In both HAL10 and HAL56 tropical ozone was found to recover first with significant depletions recurring during the winter in the polar regions for the remainder of the simulation.



**Figure S6.** Evolution of total sulfur burden in SULF10 (a), SULF56 (b), HAL10 (c) HAL56 (d). The difference between HAL10 and SULF10 (e) and HAL56 and SULF56 (f).

*P15 lines 360-366: I guess the methane increase, although limited, for the sulfur-only scenario is chemically due to the reaction of  $\text{CH}_4 + \text{OH}$ . This would mean that less OH is present under volcanically-impacted periods. There is a complex interplay between HOx, nitrogen and halogen chemistry that can result in OH reduction (and  $\text{CH}_4$  increase) unless the dominating process deals with the very high amounts of  $\text{SO}_2$  that may sequester OH through reaction  $\text{SO}_2 + \text{OH}$  (subsequently leading to the formation of sulfuric acid). In my mind, OH was rather increased for summertime midlatitude eruptions (as shown for the 2009 Sarychev eruption) reflecting a possible seasonal effect. Changes in methane amounts are also likely resulting from radiative/dynamical origin with more troposphere-to-stratosphere transport resulting from the aerosol heating in the tropopause region. A significant part of ozone changes following major eruption has been attributed to changes in transport (see e.g. Pitari, G. and Rizi, V.: An estimate of the chemical and radiative perturbation of stratospheric ozone following the eruption of Mt. Pinatubo, *J. Atmos. Sci.*, 50, 3260–3276, 1993). Similar process could apply for methane. Do the authors have an idea about the process behind the  $\text{CH}_4$  slight increase?*

The enhanced stratospheric methane burdens simulated in SULF10 and SULF56 must be attributed to one of two things:

- longer  $\text{CH}_4$  lifetime (less OH or  $\text{CH}_4$  is transported higher where there is less OH) (chemical)

- Enhanced transport from troposphere due to aerosol heating (vertical ascent)

Very recently Kilian et al. (2020) published results showing how stratospheric methane is impacted by volcanic aerosols, namely, due to changes in transport due to the heating effect. Three months after a Pinatubo sized eruption they simulated a 80ppbv (+10%) change in CH<sub>4</sub> between 40 hPa and 10 hPa. They attributed this change to upward propagating regions of elevated CH<sub>4</sub> concentrations as the result of enhanced tropical vertical ascent transporting relatively methane-rich air from the lower stratosphere into the upper levels. As the stratospheric CH<sub>4</sub> burden was also shown to increase, they suggested that the lofting of methane must also coincide with an increase in the stratospheric methane lifetime, but did not calculate this.

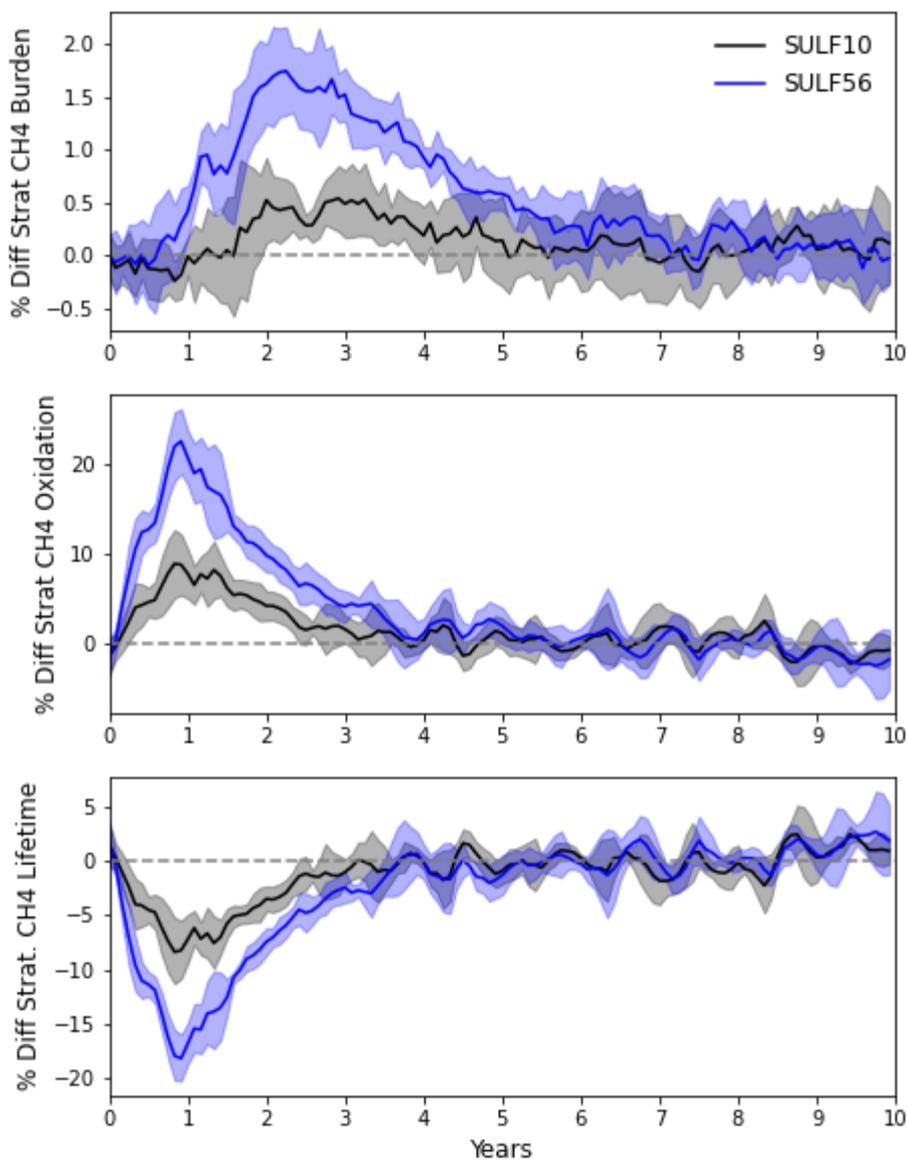
In SULF10 and SULF56, we simulated an increase in stratospheric CH<sub>4</sub> burden but also an elevated methane oxidation flux and reduction in methane lifetime (Figure S8). In other words, the flux of methane oxidation by OH, Cl and O(1D) increased. This means the increased stratospheric methane burden in SULF10 and SULF56 cannot be due to a lengthening of the lifetime and must in fact be due to increased transport from the methane rich troposphere.

We have updated the manuscript as follows:

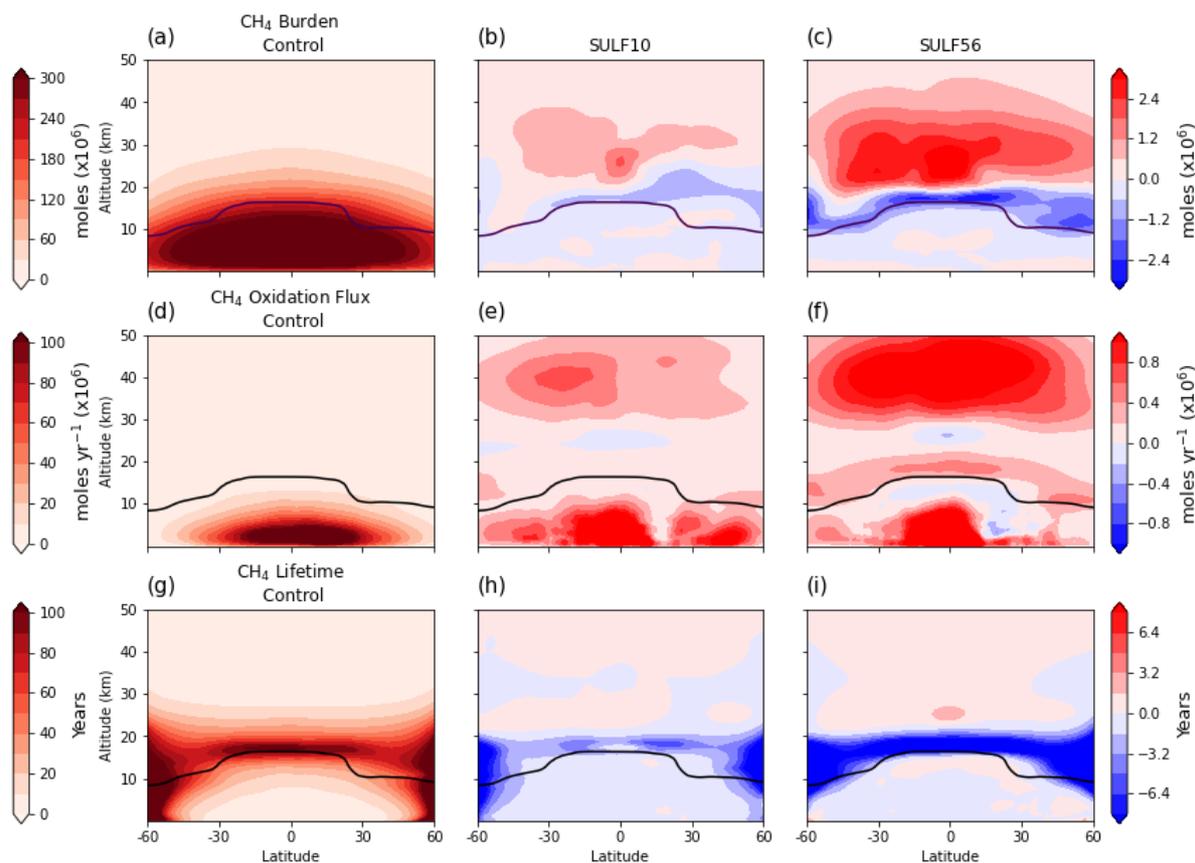
SWV and stratospheric methane are linked. SWV has two main sources: transport from the troposphere and chemical production from methane (Loffler et al., 2016). **By contrast, methane's only source is sourced from transport from the troposphere and it is destroyed by oxidation via OH, forming SWV, and reaction with halogens via equation 52.**

Following sulfur-only eruptions we simulate small enhancements in stratospheric water vapour (SWV) and stratospheric methane (Figure 8). SULF10 and SULF56 result in a peak global stratospheric mean increase in SWV of 0.4 ppmv (+7%) and 1.1 ppmv (+17%) and a 10 ppbv (0.6%) and 30 ppbv (1.8%) increase in stratospheric methane respectively. Perturbations to SWV and stratospheric methane peak 2-3 years after the eruption and recover within 7 years. The increase in stratospheric methane following sulfur-only eruptions is in broad agreement with Loffler et al. (2015), who showed stratospheric methane mixing ratios increased by ~5% following simulations of El Chichon and 15-20% following the larger Mt Pinatubo and Kilian et al. (2020) who reported a 10% increase in CH<sub>4</sub> between 40 and 10 hPa, also following simulations of Pinatubo. Killian et al (2020) suggested that this was due to enhanced vertical ascent as a result of aerosol heating, lifting relatively methane-rich air from the lower stratosphere into the upper levels. As Kilian et al. (2020) simulated an increase in stratospheric CH<sub>4</sub> burden, they suggested that the lofting of methane must also coincide with an increase in the stratospheric methane lifetime but did not calculate this. In SULF10 and SULF56 of this work, we simulate an increase in tropical vertical ascent (shown at 50 hPa in Figure S2), however, we simulate a coinciding reduction in the stratospheric methane lifetime, driven by an increase in methane oxidation by OH and Cl (Figure S8). This suggests that the increased stratospheric methane burden following sulfur-only eruptions SULF10 and SULF56 is not due to a lengthening of the stratospheric methane lifetime and, instead, is likely due to increased transport of methane across the tropopause from the methane rich troposphere as a result of increased vertical ascent in the tropical stratosphere

(Figure SIX). Due to the model set up employed in this study we were unable to diagnose this any further.



**Figure S8.** Change in CH<sub>4</sub> burden time series (top). Change in CH<sub>4</sub> loss flux (middle). Change in CH<sub>4</sub> lifetime (bottom). CH<sub>4</sub> lifetime has been calculated (as it is typically done in the troposphere) as the burden (moles)/ loss flux (moles/yr). Loss flux includes loss by OH, Cl and O1D.



**Figure Caption.** Zonal CH<sub>4</sub> burden in the control climatology (a), Change in zonal CH<sub>4</sub> burden for SULF10 (b) and SULF56 (c). Zonal CH<sub>4</sub> loss flux in the control climatology (d), change in zonal CH<sub>4</sub> loss flux for SULF10 (e) and SULF56 (f). CH<sub>4</sub> lifetime in the control climatology (g), Change in CH<sub>4</sub> lifetime for SULF10 (h) and SULF56 (1). All values are annual means over the second post eruption year. CH<sub>4</sub> lifetime has been calculated (as it is typically done in the troposphere) as the burden (moles)/ loss flux (moles/yr). Loss flux includes loss by OH, Cl and O1D.

*P18 line 404: I found the demonstration about ozone change (although largely trusted) as the dominant driver of ERFclear, clean a bit abrupt. Please provide more details here about the method used rather than only citing the Rap et al. reference.*

We have amended the manuscript to include more details regarding the ozone radiative kernel method, as follows:

To calculate the resulting radiative forcing of the simulated ozone changes we use the ozone radiative kernel (O3 RK) technique introduced in Rap et al. (2015) and updated for the whole atmosphere as outlined in Iglesias-Suarez et al. (2017) (figure SIZ). The O3 RK is constructed by calculating the change in LW and SW flux caused by a 1 ppb perturbation in ozone added to each atmospheric layer in turn. The change in SW and LW flux is diagnosed using the offline version of the SOCRATES radiative transfer model, based on Edwards and Slingo (1996). The LW component of the O3 RK (Fig. SIZb) is positive throughout the atmosphere, with a maximum in the tropical upper troposphere lower stratosphere (Fig. SIZa). The SW component is negative above ~12km altitude and positive below ~12km

altitude. This results in a net O3 RK which is positive everywhere except above ~25km between 60°S and 60N. ~~Using the whole atmosphere ozone radiative kernel from Rap et al., 2015,~~ Using the O3 RK, we are able to show that the stratospheric ozone change is the dominant driver of the ERF<sub>clear, clean</sub>, accounting for ~75% of the ERF<sub>clear, clean</sub> (Figure 9 a,b). The remainder is likely predominantly due to SWV changes with a small contribution from stratospheric methane changes. The latitudinal pattern of ozone radiative forcing reflects the locations of the ozone change, with largest forcings at the poles.

P20 line 461: specify "for two winters".

Done

P21 figure 12: this figure (as figure S4) is very interesting but I think plotting anomalies (by subtracting each simulation with the control run) would have been more meaningful especially to highlight the effect of SULF and HAL scenarios on the NH high latitudes. Such figure could be added in the supplementary material.

Done

### Short Comment - Alan Robock

1a. I don't understand how you carried out your simulations, and why you did them the way you did. Did you take an 11-year average of SSTs and sea ice, and then prescribe them, including their seasonal cycle, repeating the same average year for each entire simulation?

Yes, we used an atmosphere-only version of UKESM1.0, which utilises climatology ancillary files in place of interactive models to represent the SST and sea ice fields. These 12 month ancillary files were calculated by averaging the 1990 to 2000 SST and sea ice fields from the fully coupled transient model simulation that was run as part of the Historical CMIP6 project.

We decided to run with an atmosphere-only model set-up for a number of reasons:

- 1) To keep the computational cost of the long term simulations reasonable
- 2) In order to be systematic. This model set up allowed us to better understand the atmospheric chemistry and radiative forcing response in the first instance
- 3) Using fixed SSTs enables ERFs to be estimated (Forster et al., 2016)

1b. Why 11 years? Do you understand that this removes all interannual variation in SSTs, and removes many surface feedbacks with the climate system?

We forced our atmosphere-only simulations with 1990-2000 averaged boundary conditions taken from the Historical CMIP6 simulations because we were interested in the volcanic response in present day average conditions. It is acknowledged that the initial conditions in volcanic perturbation simulations will have an impact on the simulated response, however, this study is not about the role of initial conditions. As such, we did not seek to replicate a specific set of observed initial conditions from any point in the 1990s but rather an atmospheric state representative of the recent historical period. In forthcoming papers we

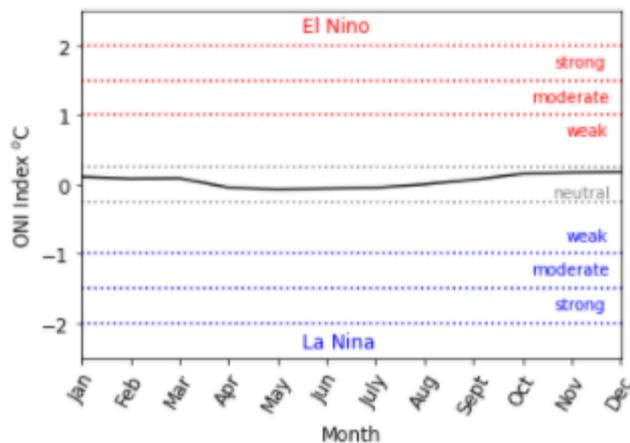
have repeated this model set up using 2050-2060 averages from ScenarioMIP simulations allowing the impacts of climate change to be isolated.

We understand this model set up means our results do not include many surface feedbacks, however, this model set up has allowed us to understand the atmospheric chemistry and radiative forcing response in the first instance. There would be a significant step up in the computational cost of running fully coupled simulations, although this was not feasible in this work, we have updated the future work section to reflect that this is an important next step.

1c. The 11-year period you are choosing, 1990-2000, includes the 1991 Pinatubo eruption. Was it included in the forcing for the coupled GCM you used? How did that affect the climate, and why are you averaging over its impacts.

The fully coupled ESM simulation, from which we spun off our simulations, did simulate 1991 Pinatubo. Seller *et al.* (2019) showed this resulted in a global mean surface temperature reduction of  $\sim 0.5$  °K. By averaging over the 1990 - 2000 time period we forced our atmosphere-only simulations with boundary conditions typical of the recent historical period but not a specific date within that decade.

1d. Furthermore there were a moderate and a huge El Niño in that decade. Did the GCM simulate them? Does your SST pattern have a permanent El Niño? How does that affect the climate response?



**Figure S1.** Oceanic Niño Index (ONI) for the SST climatology prescribed in the Atmosphere-Only simulations in this study. The ONI is calculated by comparing the SST climatology to the 1971-2000 base period from the fully coupled transient simulation run as part of CMIP6 Historical project in the Niño 3.4 region ( $5^{\circ}\text{N}$ - $5^{\circ}\text{S}$ ,  $120$ - $170^{\circ}\text{W}$ ).

The fully coupled transient simulation run as part of CMIP6 Historical project did simulate El Niño and La Niña cycles, however, as they did not themselves use prescribed (observed) SSTs they differ from the observed 1990s signals. Averaging the SSTs using the 1990 to 2000 period, by design, resulted in a permanent neutral signal in the SST pattern, see S1. This study is not about the role of initial conditions, but rather understanding the volcanic response in a model with boundary conditions typical of the recent historical period.

In response to Professor Robocks points 1a,b,d, we have amended section 2.2 as follows:

We utilise atmosphere-only, time-slice experiments whereby the initial sea surface temperature, sea ice fraction and forcing agents and depth, surface emissions and lower boundary conditions are prescribed using climatologies calculated using data from the fully coupled UKESM1.0 historical runs produced for CMIP6 (Eyring et al. 2016) and averaged over the years 1990 to 2000. By averaging over the decade the atmosphere-only simulations are forced with boundary conditions typical of the recent historical period but not a specific date within that decade. The fully coupled transient simulations had internally generated El Niño and La Niña cycles, however, averaging the SSTs over the 1990 to 2000 period resulted in a permanent neutral signal in the SST pattern, see figure S1. The 1990s, and thus these timeslices, were characterised by high background halogen levels due to anthropogenic emissions of CFCs throughout the preceding decade. The impacts of very short lived Bromine species are accounted for by adding a fixed contribution of 5 pptv into the CH<sub>3</sub>Br surface concentration.

A control simulation was run with a 15 year spin up followed by a further 20 years. A 20-year control simulation was run with after 15 year spin up, initialised from the January 1995 initialisation file taken from the UKESM1.0 historical scenario which was run as part of CMIP6 (Eyring et al. 2016). The effect of explosive volcanic eruptions was investigated by running a series of 10 year volcanic perturbation simulations spun off from 6 different years in the control run to represent the variability in QBO states. Changes are plotted as the difference between the average of the 6 ensembles and a climatology derived from the 20 year control run, cumulative forcings are calculated as the sum of the forcing over the full 10 year simulation duration.

*And why did you use the Vidal et al. emissions for Cl and Br, but not for SO<sub>2</sub>? They said Samalas emitted 158 Tg, but you several times say 56 Tg is representative of the Samalas eruption.*

This work is not simulating a 1257 Samalas sized eruption, but rather the effects of hypothetical volcanic eruptions. As a result of the variable source geochemistry and prevailing atmospheric conditions, the amount of halogen injected into the stratosphere during explosive volcanic eruptions is highly variable and differs for specific eruptions. What is clear, however, is that halogens are injected into the atmosphere following some explosive volcanic eruptions. Given this uncertainty and variance, we designed a simulation matrix that spanned a range of possible explosive volcanic emissions scenarios. The 10 and 56 Tg volcanic eruptions are hypothetical eruption scenarios similar to a VEI 6 (e.g. 1991 Pinatubo) and VEI 7 (e.g. 1257 Samalas and 1815 Tamborra) eruption. We used the halogen emission estimates from Vidal et al. (2016) with a 5% halogen injection efficiency in HAL56, and ten times smaller in HAL10. This results in a HCl:SO<sub>2</sub> ratio of ~0.26 and ~0.47 in HAL10 and HAL56 respectively, similar to the estimated stratospheric injection ratio for VEI 7 Mt. Mazama (0.3) (Zdanowicz et al., 1999) and the ratios used in Ming et al. (2020) and Brenna et al. (2020), but smaller than the ratio used in Cadoux et al. (2015). This simulation matrix spans a range of possible explosive eruption scenarios allowing us to explore how the co-emission of volcanic halogens impacts the atmospheric chemistry and radiative forcing.

To clarify this point, we have amended the manuscript as follows:

The eruption scenarios simulated in this work are hypothetical, but they are comparable to Volcanic Explosivity Index (VEI) 7 (e.g. 1257 Mt. Samalas) and VEI 6 (e.g. 1991 Mt. Pinatubo) eruptions, representing 1 in 500-1000 year and 1 in 50-100 year events respectively, with plausible amounts of co-emitted halogens based on satellite observations and volcanic plume modelling.

2. You need, in the introduction, an explanation of the ways volcanic eruptions affect stratospheric ozone, including before there were anthropogenic CFCs there and now. Since you are a chemist, chemical reactions would be useful. And dynamic processes also should be explained.

We have amended the introduction as follows:

Sulfur gases emitted into the atmosphere by volcanic eruptions have a strong direct climate impact through the formation of sulfuric acid aerosol, which reflect incoming sunlight and cool the Earth's surface. Volcanic aerosols also have the potential to alter the chemistry of the stratosphere, including ozone with significant impacts on both longwave and shortwave radiative fluxes. Ozone is impacted dynamically by stratospheric circulation changes induced by aerosol heating, and chemically by changes to ozone loss cycles. Aerosol heating in the tropics increases the vertical ascent transporting ozone to higher altitudes and latitudes, resulting in an ozone decrease in the tropics and an increase at high latitudes (Kinne et al., 1992). The addition of large amounts of volcanic aerosols increases the surface area of the stratosphere on which heterogeneous reactions can take place. Heterogeneous reactions in the stratosphere drive changes in the partitioning of NO<sub>x</sub>, ClO<sub>x</sub>, BrO<sub>x</sub> and HO<sub>x</sub> species between reservoir and active forms. Unlike stratospheric clouds (PSCs), which only occur in the extremely cold temperatures inside the winter polar vortex, volcanic aerosols provide surfaces for heterogeneous reactions at all latitudes and at all times of the year. N<sub>2</sub>O<sub>5</sub> reacts with water vapour on the surfaces of these volcanic aerosols to form HNO<sub>3</sub>, effectively sequestering reactive NO<sub>x</sub> species into a long-lived reservoir and limiting the availability of NO<sub>x</sub> radicals to take part in catalytic reactions which deplete stratospheric ozone, reducing the chemical destruction of ozone (Crutzen, 1970; Johnston, 1971). In contrast, these reactions liberate reactive ClO<sub>x</sub> and BrO<sub>x</sub> species from their long-lived reservoirs, increasing the chemical destruction of ozone (Solomon et al., 1996, Solomon, 1999, Aquila et al., 2013). The net chemical impact of stratospheric volcanic aerosol loading on stratospheric ozone is dependent on the stratospheric chlorine loading.

The net chemical impact of stratospheric volcanic aerosol loading on stratospheric ozone is dependent on the stratospheric chlorine loading (e.g., Timmreck, 2012). A large volcanic eruption in low-chlorine atmospheric conditions, such as a preindustrial atmosphere or a future atmosphere, is expected to result in net stratospheric ozone gain (Langematz et al., 2018), however, when the chlorine loading of the stratosphere is high (e.g., Tie & Brasseur, 1995), an eruption will lead to stratospheric ozone loss. High-chlorine loading may arise from anthropogenic or natural emissions. Petrological data suggest that volcanic eruptions in some geological settings may also release substantial amounts of halogen gases into the atmosphere (Krüger et al., 2015; Kutterolf et al., 2013, 2015). Petrological analysis of the 1257 Mt. Samalas eruption suggests as much as 227 Tg of hydrogen chloride (HCl) and 1.3 Tg of hydrogen bromide (HBr) could have been emitted into the atmosphere alongside 158

Tg of sulfur dioxide (SO<sub>2</sub>) (Vidal et al., 2016). The portion of the halogens erupted at the vent that reaches the stratosphere is not well constrained and has been the subject of debate in the community for decades. Halogens are soluble (especially HCl) and may be scavenged by water, ice hydrometeors and ash in the volcanic plume (Halmer et al., 2002). Despite efficient scavenging, direct stratospheric injection of volcanic halogens is predicted by theory, and sophisticated plume models suggest that between 10% and 20% of the HCl emitted at the vent of large explosive eruptions could reach the stratosphere (Textor et al., 2003).

3. I see no discussion of the impacts of volcanic eruptions on stratospheric dynamics. How do changes in stratospheric circulation affect the ozone distribution and the aerosols? You say something about a lower branch of the BDC, but do not show what you are talking about, and how the circulation changes in response to the volcanic eruptions. How does the polar vortex respond, and how does this affect the size of the Ozone Hole?

We have amended the text in the manuscript to include more information on how stratospheric circulation impacts the ozone distribution in the sulfur-only simulations, as follows:

In sulfur-only simulations, we simulate a modest reduction in global-mean ozone column, -9 DU (-3.9%) in SULF10 and -15 DU (-6.6%) in SULF56 (Figure 5a,c). This ozone depletion is catalysed by halogen radicals activated from background halogens on the surface of volcanic aerosol. We simulated a decrease in tropical ozone of <0.5 and <2 ppmv between 23 to 28 km and a symmetrical increase in zonal mean tropical ozone above in SULF10 and SULF56 respectively. This tropical ozone dipole pattern is mostly attributed to volcanic heating. Volcanic heating by the aerosol increases the vertical ascent, and brings ozone up from below enhancing the local mixing ratio.

This dynamical ozone change will not take place in HAL simulations as we simulate stratospheric cooling. Furthermore, the ozone response is dominated by chemical loss.

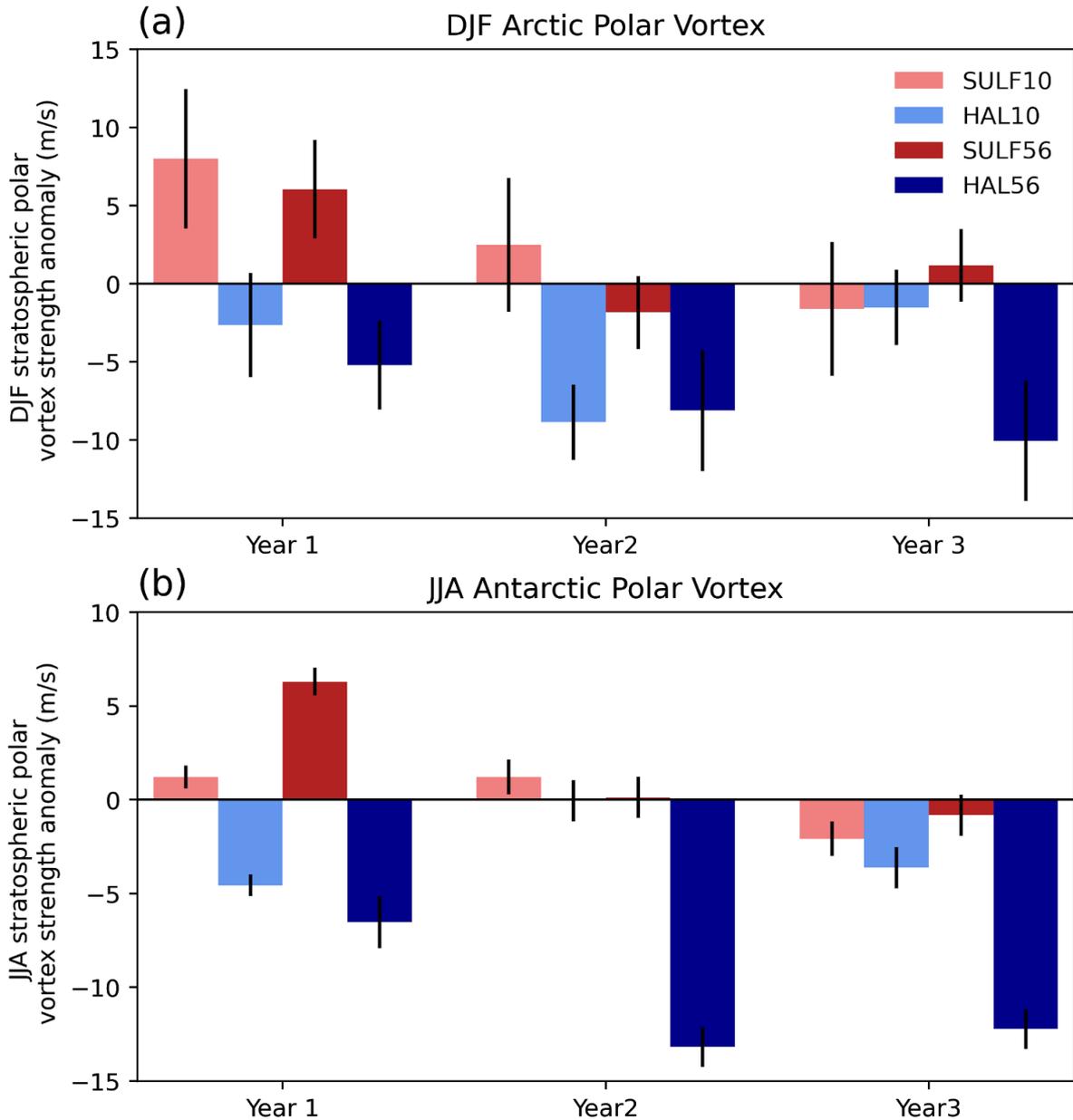
In the manuscript we discuss the reduction in rate of poleward transport with altitude and point to Figure S3 to showcase this point. This figure shows the Northward transformed eulerian mean air velocity (ms<sup>-1</sup>) in the control simulation, averaged over the full 20 year simulation length. Furthermore, we discuss the changes to tropical vertical ascent ( $\bar{w}^*$ ) and the impact this has on stratospheric methane burden.

We have included a discussion of simulated changes in polar vortex strength and the consequences for ozone in the resubmitted manuscript, as follows:

In simulations with co-emitted halogens we simulate more dramatic ozone depletions; HAL10 resulted in a peak global-mean ozone reduction of 65 DU (-22%) 1-2 years after the eruption followed by a gradual recovery over the next 3-4 years (Figure 5d). HAL56 resulted in a peak global-mean ozone reduction of 175 DU (-57%) 1-2 years after the eruption followed by a gradual recovery the remainder of the 10 year simulation, with an average reduction of 82 DU (-27%) over the 10 year simulation (Figure 5b).

Volcanic halogen catalysed ozone depletion was simulated across all latitudes, but the largest magnitude changes in HAL10 (-40%) and HAL56 (-80%) were within the aerosol cloud and the polar regions, where the co-emitted halogens were activated on aerosol surfaces and PSCs respectively (Figure 5). In both HAL10 and HAL56 tropical ozone was found to recover first with significant depletions recurring during the winter in the polar regions for the remainder of the simulation. Ozone depletion shows a similar bimodal altitude distribution in the stratosphere similar to that found in Brenna *et al.* (2020), with 3 year mean depletion maxima (-1 ppmv and -3.5 ppmv in HAL10 and HAL 56) in the lower (20 km) and upper (40km) stratosphere (Figure 6).

The simulated changes in stratospheric heating following sulfur-only and co-emission eruption scenarios affect the dynamical response of the upper atmosphere, for example, the strength of the Arctic and Antarctic polar vortices (Robock, 2000; Toohey et al., 2014). In SULF10 and SULF56, the positive stratospheric temperature anomalies in the tropics lead to an increased meridional temperature gradient. As a result, we simulated a strengthening of the polar vortex (defined as mean zonal wind speed at the vortex edge, between 55° - 65° latitude and 1 to 30 hPa) in both the Arctic and Antarctic in the first post eruption winter. In contrast, the negative stratospheric temperature anomalies in HAL10 and HAL56, lead to a decreased meridional temperature gradient and a weakening of the polar vortices. In HAL10 we simulated significant weakening of the polar vortex in the first two post eruption winters in the Arctic, and the first and third post eruption winter at the Antarctic. In HAL56, we simulated significant weakening of the polar vortex for 3-4 years at both poles. Polar vortex strength is an important driver of ozone depletion, with stronger polar vortexes leading to enhanced ozone depletion (Solomon, 1999; Zuev and Savelieva, 2019). Lawrence et al. (2020) linked an unusually strong Arctic polar vortex with the record breaking ozone loss observed in the 2019/2020 Arctic winter. As such, the strengthening of the polar vortices simulated in sulfur-only simulations may intensify ozone depletion in the first post eruption winters in both the Arctic and Antarctic. Furthermore, the weakening of the polar vortices simulated in co-emission scenarios will likely dampen the ozone response in both the Arctic and Antarctic. In addition, the simulated changes in polar vortex strength may have important consequences for the North Atlantic Oscillation and Southern Annular Mode (Driscoll et al., 2012; Jones et al, 2020; Kwon et al., 2020 ).



**Figure S7.** Change in the polar vortex strength (zonal wind anomaly 55°-65° latitude, 1 - 50 hPa) in first three post eruption years. Errors bars represent the ensemble range.

4. All the time plots need an x-axis in years and not months, so that the seasonal cycles are easy to discern. Months since July 1 are confusing and obscure what is happening. You can start all your plots on January 1 of the year you injected the gases, so we can see what the variability was before the experiment started, too. The latitude plots need labels of 0°, 30°, 60°, and 90°, not 50 only. Climate scientists are used to looking at the different regions of Earth on the natural coordinates.

Done.

5. You use many acronyms without defining them. And you define some acronyms more than once.

Done.

6. All variables need to be in italics, and chemical symbols should not be in italics.

Done.

7. You use  $r$  for both correlation and radius, even in the same figure, which is very confusing. You can only use a symbol for one thing in a paper.

We have defined ' $R_{\text{eff}}$ ' as the aerosol effective radius and ' $r$ ' as the correlation coefficient in the regression analysis. We have corrected a typographical error in the y axis label of figure 3 from ' $r_{\text{eff}}$ ' to ' $R_{\text{eff}}$ '.

8. You use a lot of global averages without showing the spatial and temporal patterns. A lot of what is happening depends on location and time of year. The specific processes are what is of interest. You can average the final result afterwards, but what is important is why things change and where and what time of year they change. This affects the climate response, as well as fluxes of UV. Global average UV does not harm anything, but local increases are important.

In the manuscript we show zonal mean stratospheric temperature changes in Figure 2, temporal and latitudinal column ozone changes in Figure 5, zonal mean ozone changes in Figure 6, tropical mean changes in stratospheric water vapour in Figure 7, and latitude-time hovmoller diagrams of column ozone and UV changes in Figure 12. In the SI document, we also include zonal total sulfur burden plots Figure S4 and S5, TOA net flux anomalies due to stratospheric ozone changes in Figure S8 and S9, and the zonal mean change in column ozone and WHO UV index in Figure S13.

To ensure the manuscript is concise and accessible, we only show spatial and temporal patterns when they feature important results and add to the understanding of the primary focus of this work.

9. Why do you average results over three years? What is special about that? I would be interested in the winter and summer seasons for each year, which is where the chemistry and dynamics responses determine the patterns. Three-year averages do not address the processes.

Done.

10. You use VEI as an index for the size of volcanic eruptions that affect climate, but that is wrong. Please see the explanation of why in Newhall et al. (2018), for example. Mount St. Helens, for example, was VEI 5, but had not sulfur and no impact on climate. I also recommend reading the original Newhall and Self VEI paper, which explains that it is an index of explosivity, and stratospheric inject is used as one criterion to assign VEI, but it should not be done in the opposite direction.

We politely disagree with Professor Robock on this point, we haven't specifically used VEI as an index representative of climate impact. Instead, we compare our hypothetical 10 Tg and 56 Tg eruption scenarios to VEI 6 1991 Pinatubo and VEI 7 1257 Samalás in order to

provide context for the reader in terms of an index understood by all volcanologists and that also gives context in terms of recurrence rates and not the likely climate effects, as is done by Newhall et al (2018).

In the avoidance of doubt we have added the following line into the manuscript:

The eruption scenarios simulated in this work are hypothetical, but they are comparable to Volcanic Explosivity Index (VEI) 7 (e.g. 1257 Mt. Samalas) and VEI 6 (e.g. 1991 Mt. Pinatubo) eruptions, representing 1 in 500-1000 year and 1 in 50-100 year events respectively (Newhall et al., 2018). VEI is used here to provide context of the recurrence rates but is not used as an index representative of climate impact.

11. You use a mixture of different styles of reference, and they all need to be in the same style.

Done.

### Specific Comments from Short Comment:

*I provided 105 comments in the attached manuscript, all of which I recommend you address.*

We thank Professor Robock's for his comment and have addressed each individually below.

*Line 17 ???*

We have amended the line in the manuscript as follows:

~~stratospheric volcanic eruption emission scenarios (56 Tg SO<sub>2</sub> ± 15 Tg HCl & 0.086 Tg HBr and 10 Tg SO<sub>2</sub> ± 1.5 Tg HCl & 0.0086 Tg HBr)~~ We simulate a large (10 Tg) and very large (56 Tg) sulfur-only eruption scenario and a corresponding large (10 Tg SO<sub>2</sub>, 1.5 Tg HCl, 0.0086 Tg HBr) and very large (56 Tg SO<sub>2</sub>, 15 Tg HCl, 0.086 Tg HBr) co-emission eruption scenario.

*Line 20 This is not the correct index to use for the impacts of volcanic eruptions on climate*

Addressed in the general comments.

*Line 20 Your first paragraph says it emitted 158 Tg of sulfur dioxide. So why did you only emit 56 Tg? I don't understand the inconsistency.*

Addressed in the general comments.

*Line 23 Acronyms have to be defined.*

Done.

*Line 26 What does this range mean? What does it depend on? & Completely counteracting? What is typical?*

We have amended the line in the manuscript as follows:

Volcanic halogens catalyse the destruction of stratospheric ozone which results in significant stratospheric cooling (1.5–3 K); counteracting the typical stratospheric radiative offsetting the aerosol heating (~1.5 °K and ~3.5 °K) simulated in sulfur-only scenarios and resulting in net stratospheric cooling (~-2 °K and ~-3.5 °K) from volcanic sulfate aerosol.

*Line 35 Which is which?*

We have amended the line in the manuscript as follows:

The dramatic global-mean ozone depletion simulated in both large (22%) and very large (57%) co-emission scenarios would result in very high levels of UV exposure ...

*Line 35 Did you account for the scattering by the volcanic aerosols?*

Covered in the manuscript discussion section.

*Line 46 158 Tg of sulfur dioxide*

Addressed in the general comments.

*Line 56 Read et al. [no comma here and in the rest of the paper]*

Done

*Line 69 highlight*

Done

*Line 73 compared to what?*

We have amended the line in the manuscript as follows:

Spectroscopic measurements of the the El Chichón stratospheric eruption plume indicated an HCl increase of 40% compared to measurements taken prior to the eruption,

*Line 74 Why don't you call it salt?*

Done

*Line 75 Chichón*

Done

*Line 85 This needs to be at the beginning of the introduction, and you need to include a discussion of how volcanic eruptions can affect ozone. Since you are a chemist, why don't you include some chemical reactions, as well as discuss the physical mechanisms? You jump into lots of past work, but do not start out describing the scientific questions, and what the various factors could be that influence the result.*

Addressed in general comments.

*Line 92 of what? & Line 96 of what?*

We have amended the line in the manuscript as follows:

Lurton et al. (2018) simulated the 2009 Sarychev Peak eruption (0.9 Tg of SO<sub>2</sub>) in CESM1(WACCM) (Community Earth System Model, The Whole Atmosphere Community Climate Model and showed how inclusion of co-emitted halogens (27 Gg of HCl) allowed better agreement with modelled and observed data. Co-emission of halogens resulted in a lengthening of the SO<sub>2</sub> lifetime, due to the further depletion of OH, and a corresponding delay in the formation of aerosols, and allowed giving better agreement between modelled and observed SO<sub>2</sub> burden, showing how co-emitted halogens could impact volcanic sulfur processing.

*Line 100 You have to explain what the mechanism was. What does background chlorine have to do with it?*

We have amended the line in the manuscript as follows:

Tie and Brasseur (1995) utilised model calculations to show how background atmospheric chlorine loadings altered the ozone response to volcanic sulfur injections. In conditions typical of the pre-1980 period, the ozone column abundance was shown to increase after a large volcanic eruption with the response being independent of the magnitude of the eruption. The increase in column abundance resulted from suppression of the NO<sub>x</sub> catalysed ozone loss cycle, driven by the sequestration of reactive nitrogen to its reservoir species via heterogeneous reactions on the surface of volcanic aerosol. The ozone response was shown to be independent of the magnitude of the eruption, as the heterogeneous conversion of active nitrogen to its reservoir was saturated. However, after 1980, higher background chlorine levels as a result of anthropogenic emissions of chlorofluorocarbons, meant that the ozone response became negative in winter at mid and high latitudes, with the magnitude of depletion increasing with eruption size. The suppression of NO<sub>x</sub> catalysed ozone loss was counterbalanced by an increase in the ClO<sub>x</sub> catalysed ozone loss, resulting in a transition in the column ozone response. Unlike in pre-industrial conditions, the ozone response was dependent on the eruption size as the heterogeneous conversion of chlorine species from reservoir to reactive is not saturated.

*Line 104 (2015) [delete comma]*

Done

*Line 105 give date*

Done

*Line 108 Explain what this is. A ratio of the number of moles of CL to S, or moles of HCL to SO<sub>2</sub>? Is that the same thing?*

Done.

*Line 110 Klobas et al. (2017) also [get rid of the commas]*

We have amended the line in the manuscript as follows:

The ~~stratospheric injection~~ molar ratio of HCl and SO<sub>2</sub> injected into the stratosphere (HCl:SO<sub>2</sub>) in this study was 0.64, considerably larger than observations from MLS (<0.1) and ice core records of Mount Manzanara (<0.3).

*Line 112 Is this the same as molar ratio? Why are you using two different ways of doing it? I find it confusing.*

We have amended the line in the manuscript as follows:

They simulated hypothetical Pinatubo sized eruptions with ~~stratospheric injection~~ a HCl:SO<sub>2</sub> mixing ratio of ~0.14...

*Line 115 Define acronyms*

We have amended the line in the manuscript as follows:

Ming et al.; (2020) simulated explosive tropical eruptions in ~~the~~ a coupled chemistry-climate model which consisted of the United Kingdom Chemistry and Aerosol (UKCA) module together with the UK Met Office Unified Model (UM) ~~UM-UKCA~~ and interactive ...

*Line 116 Describe what this is*

We have amended the line in the manuscript as follows:

... stratospheric aerosol model GLOMAP-mode. ~~and~~ They found that a volcanic halogen emission of 0.02 Tg (HCl:SO<sub>2</sub> = 0.04) into a pre-industrial background state had little impact on column ozone but 2 Tg (HCl:SO<sub>2</sub> = 0.4) showed significant and prolonged ozone depletion above both poles.

*Line 120 What is this? Is it the same as molar ratio?*

Addressed in general comments.

*Line 122 Where? Locally? Globally? It does not matter in Antarctica, as there is little UV there*

We have amended the line in the manuscript as follows:

Consequently, UV radiation increases of >80% were simulated in the tropics, averaging to >40% for 2 years.

*Line 124 Define acronyms*

Already defined above.

*Line 127 Does the halogen injection efficiency matter if you already assume a molar ratio ?*

They are different things. The halogen injection efficiency is the fraction of halogens degassed from the magma that are injected into the stratosphere. The HCl:SO<sub>2</sub> ratio is a measure of the amount of HCl injected into the stratosphere relative to that of SO<sub>2</sub>. We have amended the line as follows to minimise confusion:

They investigate the combined effect of the sulfur (523 Tg S) and halogens (120 Tg Cl, 0.2 Tg Br) emissions of the Los Chocoyos supereruption assuming a 10% halogen injection efficiency and resulting in a stratospheric HCl:SO<sub>2</sub> molar ratio ~0.4, on volcanic gases, ozone and surface UV.

*Line 133 define acronym*

We have amended the line in the manuscript as follows:

Very recently Wade et al. (2020) compared HadGEM3-ES (Earth System configuration of the Hadley Centre Global Environment Model version 3) simulations of the 1257 Mt. Samalas eruption, utilising the halogen degassing estimates from Vidal et al. (2016) and stratospheric halogen injection efficiencies of 20% and 1%, with the available surface temperature proxies.

*Line 136 of what?, records.*

We have amended the line in the manuscript as follows:

Their results suggest it is unlikely that 20% of degassed halogens reached the stratosphere, however smaller fractions gave good agreement with multi-proxy surface temperature records.

*Line 137 The aim of this study should be to answer scientific questions, not confrontation. What does this mean? How about "use?"*

We have amended the line in the manuscript as follows:

The aim of this study is to simulate hypothetical VEI 6 and VEI 7 sized eruptions, both with and without halogens, in a coupled chemistry-aerosol model with hypothetical VEI 6 and VEI 7 sized eruptions, both with and without halogens, and investigate how the co-emission of volcanic sulfur and halogens alters the evolution of volcanic aerosol, ozone, stratospheric composition, and the consequential radiative forcing and UV flux.

*Line 137 These are not indices of the size of stratospheric injection. 1980 Mt. St. Helens was VEI 5 and put no sulfur into the stratosphere.*

Addressed in the general comments.

*Line 144 needs to be defined the first time the acronym is used.*

Done.

*Line 147 Which ones?*

We have amended the line in the manuscript as follows:

The model is free-running in the atmosphere, forced by sea ice and sea surface temperatures surface boundary conditions...

*Line 148 You have to explain the differences.*

Done

*Line 160 define acronym*

Fast-Jx is the name of the scheme, not an acronym. We have provided the reference for more details.

*Line 170 Variables have to be in italics in the text as well as in equations: ERF*

Done

*Line 171 F and in the rest of the paper.*

Done

*Line 186 In the troposphere?*

Done

*Line 189 What about on land? What about snow changes or vegetation?*

We have amended the line in the manuscript as follows:

The third term,  $ERF_{\text{clear, clean}}$  accounts for changes not directly due to aerosol or cloud interactions, largely the result of changes in surface albedo and atmospheric composition. ~~In this study, we fix surface temperature and sea-ice fields meaning that surface albedo is expected to be unchanged and any  $F_{\text{clear, clean}}$  changes are the result of atmospheric compositional changes.~~

*Line 194 So you are using 11-year averages of SSTs with no interannual variation?*

*Why?*

Addressed in the general comments.

*Line 200 What does this mean physically? Forcing is instantaneous.*

The cumulative volcanic forcing is calculated as the time-integrated forcing across the entire Earth's surface area and it is an estimate of the total energy loss in Joules due to the volcanic eruption. We have amended the manuscript as follows:

Changes are plotted as the difference between the average of the 6 ensembles and a climatology derived from the 20 year control run, cumulative forcings are calculated as the ~~time-integrated forcing across the Earth's surface area and , and sum of the forcing over the full 10-year simulation duration~~ and represent the total energy loss (J) as a result of the volcanic eruption.

*Line 202 Why?*

Covered in the general comments.

*Line 208 Why this amount? Why not 60 Tg, like from Tambora?*

Covered in general comments.

*Line 210 On line 46, you say 158 Tg.*

Covered in general comments.

*Line 214 Chichón*

Done

*Line 216 Ming et al. (2020) and Brenna et al. (2020)*

Done

*Line 210 This was about 17-20 Tg, as measured by satellites. Mike Mills only got a good simulation with 10 Tg, but if your model includes fast removal processes, you need to use the amount that was observed and let the model calculate how the SO<sub>2</sub> changes. Otherwise, you have to explain why you claim it was only 10 Tg. & Line 217 not correct. Half of what was observed.*

This work is not simulating the 1991 Pinatubo eruption, but rather the effects of hypothetical volcanic eruptions of similar size. We have amended the sentence to avoid confusion as follows:

HAL10 has a SO<sub>2</sub> injection similar to that found to reproduce the SAOD following 1991 Pinatubo (Mills et al., 2016) and a 10 times smaller flux of HCl and HBr than HAL56, resulting in a HCl:SO<sub>2</sub> molar ratio of ~0.26, very close to the estimated stratospheric injection ratio for Mt. Mazama (0.3) (Zdanowicz et al., 1999).

*Table 1 SO<sub>2</sub>, This is not in the table. Why is it in the caption?*

**Table 1** - Showing the eruption masses of SO<sub>2</sub>, HCl and HBr in Tg for the four sets of experiments. ~~Equivalent effective Stratospheric Chlorine (EESC) = [Cl] added to stratosphere + 60 × [Br] added to stratosphere (Cadoux et al., 2015).~~

*Line 227 This can't be the first person to propose this equation.*

Done

*Line 230 Not italics.*

Done

*Line 230 This is not the first equation. What about the one for ERF?*

Done

*Figure 1 Units?*

Done

*Figure 1 Units?*

Done

*Figure 1 Use years, not months, for the labels.*

We thank Professor Robock for this comment, but we feel changing the x-axis labels to years would not suit this particular figure. In particular, panel (a), which only shows 6 months. We have amended the x-label of all suitable figures to use 'Years' rather than 'Months'.

*Line 240 What does this mean?  $1/e$  is 0.368.*

The caption has been amended to read:

Dashed horizontal lines in (a) (b) and (d) represent the mass remaining after one e-folding lifetime  $1/e$ .

*Line 243 e-folding lifetime?*

Done

*Line 246 where?*

The sentence has been amended to read:

Significant differences in stratospheric temperature change are simulated between the sulfur-only and halogen simulations.

*Line 250 But don't heterogeneous reactions on the volcanic aerosols result in significant stratospheric ozone depletion without additional halogens? You have to compare them. You have to diagnose all the processes that are occurring.*

*Covered in the general comments.*

*Line 252 define acronym or just write it out if you do not use the acronym again.*

Done

*Line 252 Sun [it is a proper name.]*

Done

*Line 253 define acronym or just write it out if you do not use the acronym again.*

Done

*Figure 2 This is a non-linear scale. Why? If you keep it, you have to label every change in the shade so that readers can understand the values.*

We have amended the scale to be linear and labelled every colour step.

*Line 263 units?*

The figure caption has been amended to read:

Figure 2 Zonal mean temperature difference ( $^{\circ}\text{K}$ ) averaged over the first 3 years post eruption. (a) SULF56, (b) HAL56, (c) SULF10, (d) HAL10.

*Line 264 How can there be units of  $T_g S$  plotted as contours? Are they per unit layer thickness?*

This plot is showing the difference in zonal sum of total sulfur ( $T_g$ ) between SULF10 and HAL10, and SULF56 and HAL56, respectively. We have amended the caption as follows: ~~Zonal~~ difference in the zonal sum of total sulfur ~~burden~~ averaged over the first 2 years post eruption (e) HAL56 -SULF56, (f) HAL10 - SULF10.

*Line 266 There are none in (e) and (f).*

The figure caption has been amended to read:

Differences that are not significant at the 95% confidence interval according to a Mann–Whitney U test are indicated with stipples. Zonal difference in sulfur burden averaged over the first 2 years post eruption (e) HAL56 -SULF56, (f) HAL10 - SULF10. ~~Differences that are not significant at the 95% confidence interval according to a Mann–Whitney U test are indicated with stipples.~~

*Line 268 ???*

The shorter lifetime of sulfur in the atmosphere following HAL10 and HAL56 eruptions results in ~~stunted~~ reduced aerosol growth and smaller aerosol effective radii ( $R_{\text{eff}}$ ).

*Line 276 ( $r$  [and you have to define  $r$ ] You are using  $r$  for two different things, radius and correlation coefficient, and it is confusing.*

Covered in the 'general comments' section.

*Figure 3 You use  $r_{\text{eff}}$  here and  $R_{\text{eff}}$  in the text. Which is it? And variables have to be in italics in the text and in the figures. And you are using  $r$  for two different things, radius and correlation coefficient, and it is confusing.*

Figure caption has been amended to read:

3-year mean  $R_{\text{eff}}$   $\mu\text{m}$

*Line 283 Why are there 6 different triangles of each color on each plot? What does each mean? How are these defined? Which triangle is for which year?*

Each triangle is an ensemble member simulation. The blue triangles are the 6 ensemble members for the SULF simulations, and the red triangles are the 6 for the HAL simulations.

We have updated the caption to read:

**Figure 3** Global-mean aerosol effective radius over the first 3 post eruption years as a function of the global total sulfur e-folding time. Triangles show the results from individual ensemble members (a) SULF10 (Blue) and HAL10 (red), (b) SULF56 (blue) and HAL56

(red). Both plots have regression lines fitted with correlation coefficient ( $r$ ) showing strong positive correlation.

*Line 285  $r$*

Done

*Line 285 correlation is not causation. Larger particles should fall faster and have shorter lifetimes. Why do you just show global averages, and not show the mechanisms as a function of time and space?*

Larger particles do fall out faster. The aerosol effective radius is significantly larger in SULF56 than SULF10, resulting in a significantly shorter e-folding lifetime of total sulfur. However, when we fix the mass of injected SO<sub>2</sub> and compare sulfur-only and co-emission scenarios SULF10 with HAL10 and SULF56 with HAL56, we can see the transport change having an impact as a second order effect.

We have amended the manuscript to clarify as follows:

The faster removal of sulfate aerosol in HAL10 and HAL56 reduces the growth via condensation and coagulation and results in smaller peak global-mean aerosol  $R_{\text{eff}}$ . This theory is supported by Figure 3 which shows a scatter plot of 3-year global-mean aerosol  $R_{\text{eff}}$  as a function of the global sulfur burden e-folding time for each individual ensemble member, with a significant correlation within both 10 Tg ( $r=0.88$ ) and 56 Tg ( $r=0.95$ ) eruption ensembles. The positive correlation between these two variables holds only for each eruption size scenario. To a first order, the aerosol  $R_{\text{eff}}$  is determined by the magnitude of the volcanic sulfur injection. The larger SO<sub>2</sub> injection in HAL56 and SULF56 ensemble simulations leads to larger-sized sulfate aerosols, faster sedimentation and shorter removal time compared to HAL10 and SULF10 ensemble simulations, as seen by comparing Figures 3a and 3b. However, when we fix the mass of sulfur injected and compare sulfur-only and co-emission scenarios, we can see the transport change having a second order effect. The faster removal of sulfate aerosol in HAL10 and HAL56 ensemble simulations leads to smaller-sized aerosol due to reduced opportunity for aerosol growth compared with SULF10 and SULF56 respectively.

*Line 289 0.25  $\mu\text{m}$*

Done

*Line 326 What is the gray shading?*

The grey shaded area represents the ensemble range. We have amended the figure caption as follows:

**Figure 5** Ozone percentage difference in response to the simulated volcanic eruptions (a) SULF56, (b) HAL56, (c) SULF10, (d) HAL10. Global averages of total column ozone perturbation are traced atop each panel as a function of time. Temporal average ozone anomalies are traced right, note different scales. Global-temporal averages are enumerated in the top right. Red colors indicate column ozone enhancement, and blue colors indicate column ozone depletion. Grey shaded areas represent the ensemble range.

*Line 326 ???*

Done

*Line 333 Why? What are the mechanisms?*

Covered below.

*Line 333 You already defined this.*

Done

*Line 334 How is this calculated? Global average? Compared to what?*

We have amended the text as follows:

Following sulfur-only eruptions we simulate small enhancements in SWV and methane (Figure 8). SULF10 results in a 0.4 ppmv (+7%) and a 10 ppbv (0.8%) peak increase in global stratospheric mean SWV and methane 3-4 years after the eruption, compared to the control. SULF56 results in a 1.1 ppmv (+17%) and a 30 ppbv (2.5%) peak increase in global stratospheric mean SWV and methane 3-4 years after the eruption compared to the control, perturbations recover gradually over the remainder of the simulation.

*Line 354 [delete]*

Done

*Line 356 How does this change in response to the volcanic eruptions? Is there more because the tropopause is warmer in the tropics.*

Covered in response to the general comments from the anonymous reviewer.

*Line 361 Chichón*

Done

*Line 377 Not correct. Show the circulation differences.*

Addressed in the general comments.

*Line 390 30°S-30°N, Line 390 15-20 km*

Done

*Line 395 No. Colder land, especially for the 56 Tg cases, will produce more snow.*

We have amended the text as follows:

As we ~~prescribe~~ surface temperature and sea ice were prescribed, surface albedo changes were small ~~are assumed to be unchanged~~, meaning that ERF<sub>clear, clean</sub> predominantly represents the forcing from atmospheric composition changes

*Line 407 Should show this.*

*We have included a new figure in the SI, and amended the text as follows:*

*The latitudinal pattern of ozone radiative forcing reflects the locations of the ozone change, with largest forcing at the poles, as shown in figure S8 and S9.*

*Line 434 For some reason, you already defined this in a figure caption. In any case, you have to explain what it is, how it is calculated, and why you are introducing it only here.*

*We have amended the manuscript as follows:*

*Plotting the column ozone percentage change against the magnitude of injected halogens expressed as Equivalent Effective Stratospheric Chlorine (EESC is a measure of the ozone destruction potential;  $EESC = [Cl]_{\text{added to stratosphere}} + 60 \times [Br]_{\text{added to stratosphere}}$ ; Cadoux et al., 2015) from this study and a number of previous studies, we find an exponential decay curve describes this relationship:*

*Line 451 This definition needs to be in the text.*

Done.

*Line 488 [delete] Every sentence in the paper should be worth noting, or it should not be in the paper.*

Done

*Line 468 Does this account for blocking by the volcanic aerosol cloud? & Line 490 You have to do this calculation. Use the TUV code. The claim that it is not important needs to be validated.*

In this work we utilise the simple heuristic relating column ozone to clear-sky surface UV as defined in Madronich (2007) to estimate how the surface UV would change as a result of the column ozone changes in HAL10 and HAL56. In the manuscript we caveat this estimate with the fact that volcanic aerosols will reflect UV and partially offset this change, however volcanic aerosol levels reduce rapidly after peaking in the first post eruption year.

*Line 532 But why not the SO2 estimates?*

Covered in the general comments.

*Line 543 This is the same. So it is not more than offsets.*

We have amended the wording in the manuscript to minimise confusion, as follows:

*Ozone depletion catalysed by volcanic halogens leads to stratospheric cooling (HAL10  $\approx$  -2 K, HAL56  $\approx$  -3.5 K) which more than offsets which offsets the volcanic aerosol heating (SULF10  $\approx$  1.5 K, SULF56  $\approx$  3.5 K) and results in a net stratospheric cooling (HAL10  $\approx$  -2 K, HAL56  $\approx$  -3.5 K).*

*Line 573 Why? How can you tell how much, if any, there will be?*

We have amended the text as follows:

This work shows for the first time that co-emission of plausible amounts of halogens can amplify the effective radiative forcing in simulations of explosive volcanic eruptions. This work highlights the need to include volcanic halogen ~~fluxes~~ emissions when simulating the climate impacts of past or future eruptions and the critical need to maintain space-borne observations of stratospheric compounds to better constrain the stratospheric injection estimates of volcanic eruptions.

