We thank reviewer #2 thoughtful comments on our manuscript. Below are our responses to each of the question raised.

Question: A weak aspect is the exploitation of the two different models. It seems a lot of effort to run both models for the analysis, yet not very much is made of comparing their performance and discussing possible reasons for their different performance or what can be learned from this. In many sections and figures, only one of the two models is shown (different single models for different sections), which is not well justified.

Answer: In the revision, we showed both model results for all the cases discussed in the manuscript. Therefore, we added some new plots which are new Fig. 8, 10, 13, 15 (shown below).

*Question: Page 3, paragraph 2: There have in fact been multi-model assessments of Arctic sur- face ozone in global CTMs. See Monks et al., (2015), Emmons et al., (2015), and an older study by Shindell et al., (2008). Monks et al and Shindell et al both show over-prediction of surface ozone at Barrow in spring, likely as a result of missing halogen chemistry. However, Emmons et al show a general model *underprediction* over the depth of the Arctic troposphere in April compared with ozone sondes, suggesting that the halogen-induced bias may not be pervasive in the Arctic troposphere. It would be helpful to see these previous studies highlighted in the text for context.*

Answer: In the revision, we added the following sentences:

“Previous multi-model assessments of Arctic surface ozone in global chemistry transport models (CTMs) gave quite different implications on the role of halogens. For instance, Monks et al. (2015) and early modelling work by Shindell et al. (2008) both showed over-prediction of surface ozone at Barrow in spring, implying a result of missing halogen chemistry. However, Emmons et al. (2015) showed a general model under-prediction in April compared with ozone sondes, suggesting that the halogen-induced bias may not be pervasive in the Arctic troposphere. Thus, relatively little is known about model skill in reproducing polar spring boundary layer ozone, on time scales of hourly, daily and monthly, leaving a large gap in our understanding of the global ozone budget in the polar regions.”

Question: Page 5, line 3: “The retrievals were performed on a 0-4 km grid with 0.2 km resolution.” Not clear what this means. What is a 0-4km grid?

Answer: In the revision, we have the new sentence “The retrievals were performed on a 0-4 km altitude grid with 0.2 km resolution. Due to the altitude of the instrument (610 m) and the lack of low or negative elevation angles, the retrieved profiles are only sensitive to well-mixed BrO in a deep boundary layer, and to lofted BrO event.”

Question: Page 5, line7: Both models are driven by ERA-Interim data. For UKCA, please briefly explain what this means for the climate model. i.e. does this imply nudging with a certain degree of relaxation? Over what altitude range? It is important to recognise that this is different from a purely offline model (such as pTOMCAT). What else is prescribed / free-running between the models? Clouds? Surface fluxes?

Answer: In the revision, we added the following two paragraphs to highlight the model differences in atmospheric dynamics:

“A global chemistry transport model, p-TOMCAT, and a global chemistry climate model, UKCA, are used in this study. The offline p-TOMCAT used 6-hour ERA-Interim dataset to drive its winds, temperature and moisture. The ERA-Interim data were taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al., 2011). In this study, a nudged UKCA version is used to ensure a model meteorological field close to the real situation for data-model comparison. We follow the work of Telford et al. (2008) with a standard nudging relaxation parameter G=1/6 h^{-1}, which value lies within the range of relaxation parameters used by other models (Jeuken et al., 1996; Hauglustaine et al., 2004; Schmidt et al., 2006). We used the 6-hour ERA-Interim winds and temperature to constrain UKCA model’s dynamical field. However, nudging is not applied to all levels; no nudging being applied above level 50 (~48 km), or below level 12 (~2.9 km (the actual height varies depending on the orography). To avoid instability of the model, moisture is not nudged to reanalysis data, therefore it is free running.

Both models applied a non-local boundary layer mixing scheme, but p-TOMCAT based on the parameterisation of Holtslag and Boville (1993), while UKCA based on the scheme of Lock et al. (2000). In terms of convective mass flux, p-TOMCAT applied the scheme of Tiedtke (1989) – which has been updated to increase convective transport to the mid and upper troposphere (Barret et al., 2010; Feng et al., 2011), and UKCA applied the bulk convection model of Gregory and Rowntree (1990). As shown in a multi-model inter-comparison in tropics, these two models showed different behaviour in terms of deep convective transport of tropical boundary layer tracers (Hoyle et al., 2011). The clouds and precipitation schemes are also different between the two models (Russo et al., 2011), resulting in different wash-out rates for aerosols and soluble chemical compounds. The precipitation bias in the op-TOMCAT model (Giannakopoulos et al., 2004) is remedied by applying a correction to force the simulated precipitation values towards Global Precipitation Climatology Project (GPCP) observations (Adler et al., 2003), following the work in Legrand et al. (2016). This corrected precipitation scheme has been used in recent sea salt aerosol modelling works (Rhodes et al., 2017; Yang et al., 2019). However, precipitation in UKCA is free running, therefore the two models may have different wet removal rates for soluble gaseous-phase species. Details of other model configurations, mainly in chemistry scheme used are described in sections 3.1 for p-TOMCAT and 3.2 for UKCA.”

Question: Page 5, line 23: Is the Law et al., (2000) study the most up-to-date reference for the model chemistry scheme? How up-to-date is the kinetic data used? How do these data compare with that used in the UKCA model for the same tropospheric reactions? Does p-TOMCAT include non-halogen related heterogeneous chemistry (it seems that UKCA does)-e.g. N2O5 hydrolysis on aerosol, which is likely important for winter / early spring ozone and NOy in the Arctic. Given the focus on comparing ozone performance between the models, it is important to acknowledge any important differences in the chemical schemes of the models.

Answer: We added the following two paragraphs Section 3.1 to highlight he chemistry differences between the two model.

“The ozone photochemistry scheme applied to the model has been detailed in previous studies (Law et al., 1998, 2000) and Savage et al. (2004), with updates including an isoprene chemistry scheme, same as the one implemented to the UKCA model by Young et al. (2009) according to the method of Poschl et al. (2000), a hydrolysis reaction of N_2O_5 on aerosols and cloud droplets (Yang et al., 2005), a tropospheric bromine scheme involving both gaseous-phase reactions (Yang et al., 2005) and heterogeneous reactions (Yang et al., 2010), and a

Fast-J photolysis scheme developed by Voulgarakis et al. (2009b), which is not used in this study. They found that N_2O_5 hydrolysis can cause net NO_X loss at high latitudes by up to 60% in the northern hemisphere and ~80% in the southern hemisphere (Yang et al., 2005). They found that including halogen-related heterogeneous reactions on aerosols and cloud droplet can significantly increase polar BrO partitioning by a factor of ~3 (Yang et al., 2010). This heterogeneous reaction scheme for halogen reactivation was also implemented to the UKCA model (Yang et al., 2014; Dennison et al., 2019; Ming et al., 2020)."

Ozone is dry-deposited in the bottom model layer with dry deposition velocity inferred from the study of Ganzeveld and Lelieveld (1995) by Giannakopoulos (1998). The original dry deposition velocity over ocean and snow (=0.05 cm s^{-1}) is reduced to 0.01 cm s^{-1} in this study following recent modelling work by Hardacre et al. (2015) and Luhar et al. (2018) as well as the Helmig et al. (2007). Since p-TOMCAT only covers part of the stratosphere with a top layer height of ~31 km, a simplified stratospheric chemical scheme has to be used, including a pre-prescribed top boundary condition for ozone. Therefore, p-TOMCAT model is quite different from the UKCA model in upper troposphere and lower stratosphere. However, it is unlikely that the downwards transport of air mass in the polar region may significant influence near surface bromine."

And in section 3.2, we added one sentence: "In p-TOMCAT model, heterogeneous reactions occur also on cloud droplets, but UKCA does not include such reactions. Therefore, in free troposphere, the BrO partitioning in UKCA may be lower than that in p-TOMCAT, which may result in more soluble inorganic bromine species being washed-out by precipitation in UKCA, as discussed in section 4." In section 4 we added one sentence: "In addition, p-TOMCAT considers heterogeneous reactions on cloud droplets while UKCA does not, this difference may explain why BrO partitioning in p-TOMCAT is higher than that in UKCA, especially in free troposphere, where BrO partitioning can be as large as 50% (Fig. S2). In addition, the higher BrO partitioning in p-TOMCAT also attribute less BrY removal by dry and wet depositions."

Question: Page 7, line 15: Care needs to be taken in over-interpreting the reason for differences between the models and assigning this to mainly physical parameters (and I agree that such differences between a climate model and CTM would be expected - although see also my comment on nudging), as it may be that there are important differences between the model chemical schemes (see point above). This is why it would be useful to point out more about these potential differences. I am not sure it is the case that only or a dominance of physical factors can be assumed.

Answer: Point taken. We remove the relevant sentence (below) in the revision: "Given that both models used in this study share a very similar tropospheric bromine-chemistry scheme, it is likely that the major factor causing model-to-model differences is in the physical set-up, rather than the chemical set-up. This is reasonable given that p-TOMCAT is a CTM while UKCA is a GCM."

Question: Page 7, line 32: Effect of dry deposition on ozone abundances during long-range transport into the Arctic. A key references here showing suppressed high latitude ozone due to deposition loss to vegetation in Siberia is Stjernberg et al., (2012).

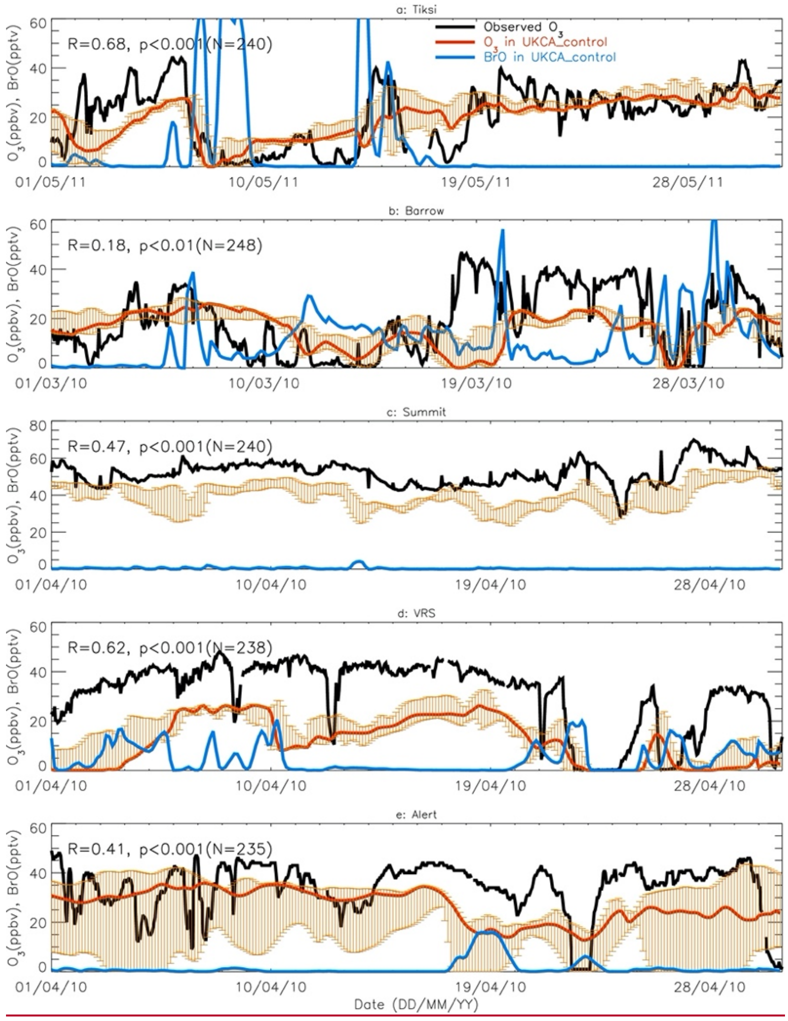
Answer: In the revision we added the following sentence: “For example, the suppressed high latitude summer ozone in Siberia is related to deposition loss to vegetation during long-range transport into the Arctic (Engvall Stjernberg et al., 2012).”

Question: Section 4.1: Discussion of model and observed seasonal cycles. I think it is worth explicitly pointing out that inclusion of the halogen chemistry the control pTOMCAT model leads to severe underestimation of spring ozone at Summit and Pallas.

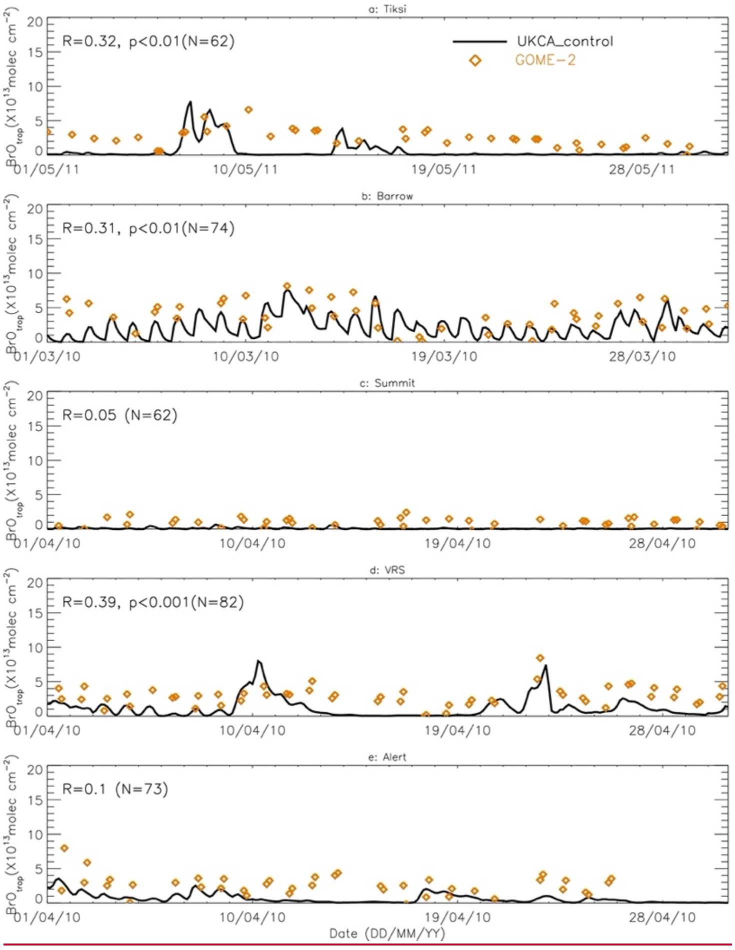
Answer: Point taken. We clearly mentioned it in the revision.

Question: It does not seem obvious to me why in presentation of the results in Section 4.3.1 model output switches from using p-TOMCAT to UKCA. Perhaps I have missed something in the applicability of the simulations to different periods. In general, only one model is shown for each part of the results comparing with observations. Would it be more informative to show both models where possible?

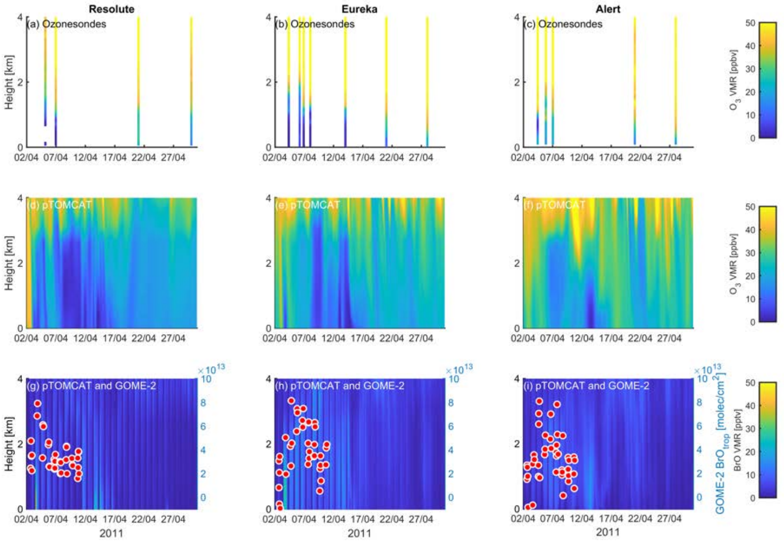
Answer: In the revision, we showed both models outputs for all the cases that were discussed in the manuscript (see below).



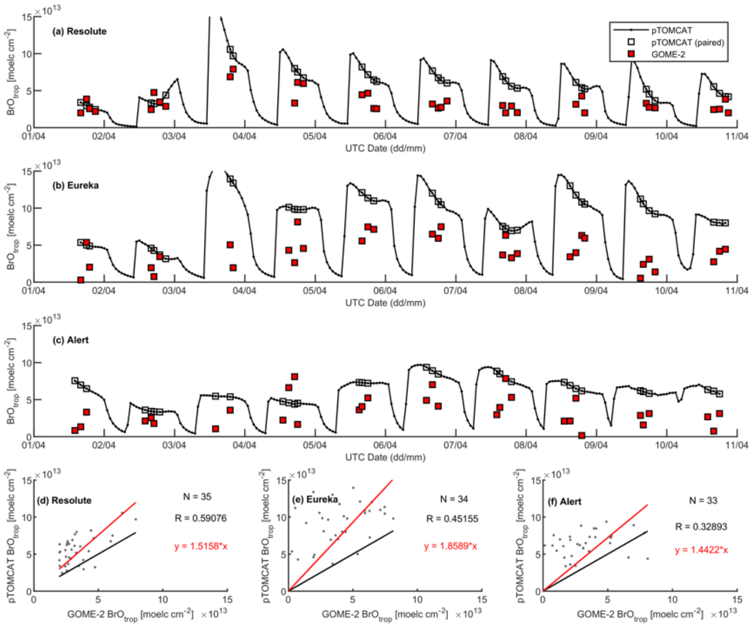
New figure 8: Same as Fig.7 but for UKCA_control run result.



New figure 10: Same as Fig. 9 but for UKCA_control run result.



New Fig. 13: Same as Fig.12 but for pTOMCAT_control run result.



New Fig.15: Same as Fig. 14 but for pTOMCAT_control run result.

Question: Throughout: I find the use of the labels "pTOMCAT_SI_OO_VSLS" and "UKCA_SI_OO_VSLS" to name the two control runs overly complicated and distracting when reading. It is better just to call these "pTOMCAT_control" and "UKCA_control" in the text. The names of the other experiments are then enough to highlight what is missing/included for the other runs.

Answer: Point accepted, and we have used these new labels in the revision.

Question: I would recommend a more explicit short "Summary" or "Conclusions" section to definitively set out the key findings of the study and their context in the wider picture. At the moment, the final paragraph is a bit brief and needs to be separated from the main discussion.

Answer: Thank you for the constructive suggestion, in the revision we have re-written the 5: Discussion section and the 6: Summary section, shown below:

"5 Discussions

Regarding tropospheric total BrY in the Arctic, as mentioned previously, the model-to-model difference can be as large as 100% at near surface layer (under the same bromine loading). As a consequence, the ozone loss due to bromine chemistry can be different by a factor of two. The relatively high BrY in pTOMCAT_control run is partly due to the higher BrO partitioning in p-TOMCAT (attributed to the inclusion of heterogeneous reactions on cloud droplets) and thus less wet removal of soluble bromine species, and partly due to stronger vertical mixing of air masses in lower the troposphere and thus less dry deposition removal of reactive bromine species from the surface layer.
On a global scale, the uncertainty of the sea spray (from open ocean) source can be a factor of four (Lewis and Schwartz, 2004). On sea ice, the blowing snow related SSA production is

sensitive to both snow salinity and bulk sublimation flux calculated (as a complex function of near surface wind speed, temperature and relative humidity, etc) (Yang et al., 2019).

Although we lack snow data on Arctic sea ice to strictly constrain the3.5 times Antarctic Weddell Sea snow salinity used for the Arctic (Section 3.1), the likelihood of higher snow salinity in the Arctic implies that there is more SSA generated from same amount of blowing snow sublimation flux (also with slightly larger SSA in size). As a consequence, there is more reactive bromine released from blowing snow in the Arctic than in the Antarctic. Given that the snow salinity effect on SSA mass production is almost linear, the uncertainty caused by this factor can be linearly estimated when more snow data in the Arctic are available. However, in terms of relative bromine release from SSA, the actual emission flux varies and depends on the salinity, this is due to the cut-off threshold size applied (i.e., a dry NaCl radius of 10 μm in the control run). Therefore, the reactive bromine release from SSA is a function of snow salinity and SSA spectrum.

Another factor that may directly affect reactive bromine emission is the depletion factor. Fig. S6 shows simulated ozone from the pTOMCAT_Fixed_DF run, in which a fixed bromine DF scheme (Table S1) is used. For comparison, the pTOMCAT_control run result is shown, in which the seasonal DF scheme (Table 1) is applied. As can be seen, the timing of the spring ozone minimum shifts slightly from April in pTOMCAT_control towards March in pTOMCAT_Fixed_DF, which makes the model agreement poorer, as the observed ozone minimum is in May at the four coastal sites. To achieve better agreement with the observations, the model needs either an even larger seasonal amplitude of bromine DF than that in Table 1 or a further shift of the DF phase by at least one month, e.g., to allow the annual maximum DF (=0.53) to shift from May to June. However, due to lack of year-round SSA bromide data in the Arctic, we could not validate the DF patterns used in this study as this requires systematic measurements of the SSA bromide depletion strength in the Arctic. This is critical as local SSA is a large source of bromine and the seasonal DF not only affects the timing but also affects the total bromine flux to the atmosphere. Model bias also comes from applying the same depletion factor scheme (i.e., Table 1) to both open-ocean-sourced sea spray and sea-ice-sourced SSA. As we know that freshly released sea spray is alkaline with pH>8, and therefore, the anions in sea spray may buffer the absorbed nitrate and sulphate before getting acidified to allow bromide to be released through heterogeneous reaction, e.g., $\mathrm{HOBr+Br^- \rightarrow Br_2}$ (e.g., Sander et al., 2003; Breider et al., 2009). On sea ice, the situation could be different as surface snow may have been pre-acidified before grains are lifted in to the air to form SSA. Unfortunately, this difference in the process of bromide liberation from SSA particles is beyond the scope of this study, but we note that it could result in bias, e.g., in bromide releasing from air-borne SSA in both strength, timing and locations.

To investigate model sensitivity to the above key parameters used in describing sea-ice sourced SSA and reactive bromine release from SSA, we performed additional model experiments (in Table 3) by altering one or a few parameters in each experiment and comparing the output with the pTOMCAT_control output (for year 2007). For most key parameters, we designed a pair run with one applying a higher value and the other a lower value than in the control run. Model results are shown in figure 16 with derived sea-ice sourced $\mathrm{Br_Y}$ (April) and ozone (as well as change with respect to the control run) shown in Table 3.

Since the control run applied a 3.5 times Weddell Sea salinity, the 10 times salinity run in pTOMCAT_high_salinity and 1 times salinity in pTOMCAT_low_salinity is roughly ~3 times and ~1/3 of the control run salinity, respectively. Comparing to the pTOMCAT_control run, the sea-ice sourced $\mathrm{Br_Y}$ (April) in pTOMCAT_high_salinity increases by +94.8%, corresponding to additional ozone loss by -37.5%. Sea-ice sourced $\mathrm{Br_Y}$ (April) in pTOMCAT_low_salinity decreases by -60%, corresponding to ozone increase by +61%. It is interesting to note that ozone and Br_Y percentage change in pTOMCAT_low_salinity is at a ratio of 1:1, but in pTOMCAT_high_salinity run the ozone percentage change is only < 1/2 of the Br_Y percentage change. Sea-ice sourced Br_Y in pTOMCAT_SSA20μm (with a large cut-off radius size of 20 μm) increases by +42.3%, corresponding to additional ozone loss by -21.1%, which is almost half of the Br_Y percentage change. On the contrary, sea-ice sourced Br_Y in pTOMCAT_SSA5μm (with a small cut-off radius size of 5 μm) decreases by -55.1%, corresponding to ozone increase by +45.4%. Sea-ice sourced Br_Y in pTOMCAT_2×DF (a doubled DF) increases by 84.5%, corresponding to additional ozone loss by -35.7% (less than 1/2 of the Br_Y change). Sea-ice sourced Br_Y in pTOMCAT_0.5×DF (a halved DF) decreases by -47.3%, corresponding to additional increase by +45.8% (almost same amount of the Br_Y change). Sea-ice sourced Br_Y in pTOMCAT_spectrum_1 (with a small N=10) reduces by -20.1%, corresponding to ozone increase by +16.5%. Sea-ice sourced Br_Y in pTOMCAT_spectrum_2 reduces by -42.2%, corresponding to ozone gain by +38.6%. In all model experiments with reduced Br_Y from sea ice, the percentage change in ozone is almost in the same amount of Br_Y change. However, in the Br_Y increasing cases, the ozone percentage (loss) change is only half or less than that of the Br_Y percentage change indicating that ozone consumption efficiency is getting lower at higher reactive bromine loading, therefore, introducing extra reactive bromine to the environment will not necessary result in equivalent amount of ozone loss as at low reactive bromine loading.

The above model experiments clearly show the possible range of modelled ozone and Br_Y in the Arctic caused by uncertainty of each key parameter involved in the parameterisations. From these runs we can derive the likely maximum effect from the sea-ice sourced SSA from blowing snow. For example, the mean DF values in spring (March, April and May, see Table 1) are ~0.5, a doubling DF indicates all bromide in SSA is released to the air, thus the pTOMCAT_2×DF run represents an extreme scenario with the maximum effect from blowing snow (with other conditions unchanged), so is the pTOMCAT_SSA20μm run, as under this cut-off threshold, almost all SSA formed from blowing snow releases bromide as a source of reactive bromine. pTOMCAT_high_salinity represents another extreme case that shows the large effect from blowing snow. Their combination effect can be multiplied and result in even larger effect. Equivalently, under extremely low snow salinity (such as in pTOMCAT_low_salinty) or small DF (such as in pTOMCAT_0.5×DF), the blowing snow sourced SSA effect on Arctic surface ozone and reactive bromine will be less important than the control run. Further field measurements will be required to collect data to constrain these key model parameters.

6 summary

For the first time, using two global chemistry models, we have examined the three tropospheric bromine sources (bromocarbons, open ocean sea spray and sea-ice-sourced SSA) and their impacts on Arctic boundary layer bromine and ozone loss. Our modelling experiments show that inclusion of bromine chemistry can greatly improve Arctic surface ozone seasonality reproduction, in particular the spring ozone depletion observed at most Arctic coastal sites, such as Tiksi, Barrow, VRS and Alert. However, inclusion halogen-chemistry leads to severe underestimation of spring ozone at inland sites such as Summit and Pallas. Our model results shows that very short-lived bromocarbons contribution to Arctic tropospheric Br_Y is less than half pptv in the near surface layer, corresponding to small ozone loss of < 1 ppbv. Multi-year simulations show that inclusion of bromine chemistry can cause Arctic surface ozone loss by 10~20 ppbv in spring, with almost half of the ozone loss attributed to open-ocean sourced SSA and the other half from sea-ice sourced SSA. However, without SI-sourced bromine, models cannot reproduce Arctic ozone depletion events, and OO-sourced bromine only affects background atmospheric ozone and cannot by itself produce any polar surface ODEs.

Although a very similar tropospheric halogen scheme applied in the two models, the model-to-model differences are relatively large. For example, boundary layer Br_Y in p-TOMCAT control run is higher than in UKCA control run, which is likely related to the different wet and dry depositions of reactive bromine species. Comparing the GOME-2 satellite data, p-TOMCAT BrO_{trop} overestimates the observations by a factor of ~2 during BEEs, but agrees well with the observations during non-BEEs. On the contrary, UKCA BrO_{trop} generally underestimates the observations by ~50% during BEEs, but severely underestimates the observation during non-BEEs (e.g., more than an order of magnitude at Summit). Despite the model differences, both model’s outputs of time series of surface ozone and tropospheric column BrO (in spring) show significant correlation to the observations at most selected periods, which strongly supports the physical and chemical mechanisms implemented.

Due to the relatively coarse model resolution (e.g., 2~3 degree in horizontal direction), our models cannot resolve small scale ODEs, e.g., with a spatial scale < ~500 km (or with a temporal scale of < ~1 day). Thus, to allow a better reproduction of small-scale ozone events, a fine resolution model is needed. Ozone sonde data from three adjacent high Arctic Canadian sites (Resolute, Eureka and Alert), satellite BrO_{trop} and back-trajectory model output clearly indicate a large ODE (and BEE) in association with a stormy system, which event is successfully captured by the two models, further confirming that ODEs and BEEs can be long-distance transported. Although our global models cannot be able to reproduce small-scale ODEs, the success of the models in capturing large scale ODEs (and BEEs) gives additional evidence from a chemistry considerations to the proposed mechanism of SSA production and reactive bromine release from blowing snow on sea ice (Yang et al., 2008; 2019; Frey et al., 2019). Note that the success of the blowing snow mechanism does not eliminate possible contributions from other candidate processes, e.g. reactive bromine from the snowpack, open leads, frost flowers, sea ice surface, etc. Change in sea ice extent and type in a warming climate will influence Arctic boundary layer chemistry and Arctic climate, including the deposition of atmospheric mercury to the surface (Wang et al., 2019).”

Question: Figure 2 caption: “in various experiments” – please clarify in the caption from which model.

Answer: Done.

Question: Figure 6 caption, please correct date “(May 201)”

Answer: Done.