

Dear Referee,

Thank you for your comments on the paper and constructive recommendations. We have tried to follow your suggestions and have taken into account most of them. Following we mention how the manuscript has been changed according to your comments.

General comments

1. In my opinion, the **CTM results (Section 4)** are extremely problematic. I provide some more detailed discussion in my specific comments below. Here, I will just highlight the main points. The total ozone maps from the CTM shown in Figure 10 are in stark disagreement with the OMI total ozone maps from Figure 1. That alone puts the utility of the CTM experiments into question. Furthermore, the morphology of the ozone loss frequency maps (Figure 11) bears only a vague resemblance to the actual geometry of the polar vortex. The latter is not shown in the paper but it's easy to plot using the same reanalysis that the paper uses, MERRA-2, as I show below. It appears that areas of high loss frequencies from the CTM often fall outside of the vortex boundaries. This doesn't seem right. The way polar ozone loss works is that the most significant depletion occurs within the chemically processed air mass rich in active chlorine, i.e., within the polar vortex, not outside of it. In addition, or perhaps related to the above, the CTM experiments in Section 4 suggest that much, even most of the ozone loss occurred via gas-phase reactions involving NO_x. This goes against our established understanding of polar ozone chemistry. That doesn't automatically make it wrong and, yes, if true it would be a major finding – but then it would require a lot stronger evidence than a low-resolution CTM experiment that fails to reproduce the observed evolution of total ozone distributions! As it is, this result only indicates likely problems with the CTM. See my specific comments for details.

Reply:

We removed the total ozone maps from the article, as well as other maps, and transferred them to the Supplement. Instead of these figures, we have added to the article figures of the vertical distribution of ozone and associated gases, as well as the destruction of ozone in chlorine and nitrogen catalytic cycles. It can be seen from these figures that at altitudes of 15-25 km, the main role in the destruction of ozone is played by its destruction in halogen cycles, and the role of nitrogen cycles is negligible. However, at altitudes of 25-40 km, significant ozone depletion is noted, in which nitrogen catalytic cycles play the main role.

2. **Section 3.1.** Almost everything in this section has already been discussed in detail elsewhere: Evolution of total ozone in Dameris et al., 2020; minimum temperatures as compared to other cold winters and climatology in Innes et al., 2020; heat fluxes, wave activity and geopotential height and related metrics as well as surface impacts in Lawrence et al., 2020. You are clearly aware of that as you cite those other studies in the paper. In principle, it's OK to have a study that confirms previously published results, especially if it uses different methods or data sets. However, I'm not convinced that this is the case here as both papers use reanalysis data and similar diagnostics. If there are any novel or otherwise valuable aspects here, please clearly state what they are and explain how they are distinct from the findings of the existing papers. If not, then I think most of this section should be eliminated. One element that may not have been discussed before (unless I missed something) is the analysis of 3-D Plumb fluxes. Currently, the discussion of Plumb fluxes is somewhat limited in the manuscript. Perhaps one way to salvage this section would be to expand this part while significantly shortening much of the preceding material.

Reply:

Section 3.1 was reduced: former Fig.2a (Tmin mean 70-90N at 70 hPa) was moved to the Supplement, former Fig.4a (NAM index) was removed, former Figures 4b-4c (with estimation of significance of difference between Fz (2020) and Fz(climate mean)) were moved to the Supplement. Also several plots with Plumb fluxes for periods March 18-20, March 20-22, and March 22-24 2020 were added to Supplement.

3. **The different parts of the manuscript are quite disconnected from each other.** For example, chemical ozone loss is calculated using three different methods (trajectory analysis, ozonesondes, and CTM) but no attempt is made to cross-check and reconcile the results.

Reply:

We made the backward trajectory calculations for other vertical levels (400 K - 500 K) in winter 2019-2020 and obtained the results showing the maximum of ozone loss at levels 435 K-460 K that correspond to estimates based on ozonesondes. The probable reason of quantitative difference is also discussed in revised text.

4. **The description of methods and data sets used in this study is insufficient.** It's not always clear which reanalysis is used for what. It's not even clear if the NCEP reanalysis mentioned in line 138 is used at all as it's not talked about anywhere else in the manuscript. I couldn't find any information about which reanalysis was used to generate Figures 2-6. Section 2.2 uses ERA5 to initialize the ozone content of air parcels used in the trajectory analysis, but no evaluation of ERA5 ozone is provided or cited (although I don't think that ERA5 ozone has been thoroughly validated yet). The

descriptions of the trajectory model and the CTM also lack detail. For example, how is vertical advection done in the CTM?

Why should we think that the very course resolution of the CTM ($5^\circ \times 4^\circ$) is adequate? It is not clear why different reanalyses are chosen to drive the trajectory model (JRA) and the CTM (MERRA-2). This is not necessarily wrong, but it does require some explanation that is not provided. No justification is given for the choices made in the trajectory calculations, e.g., why the parcels were initialized at those specific locations and times. This list is not exhaustive. See my specific comments for details.

Reply:

In the updated version of the article, we described the methods used in more detail. In particular, we have described in detail the chemical scheme of the model and the structure of the performed numerical experiments. With regard to the rather coarse resolution of the model, we were interested in whether it was possible, using such a model, to describe the morphology of the formation of ozone mini-holes in winter-spring 2019-2020. The results of a comparison of the calculated and measured variability of the total ozone content at ozonometric stations demonstrated that this is qualitatively possible, although the quantitative differences can be quite significant.

	Specific comments	Reply
1.	L106. Note that there's a considerable debate over whether the Arctic depletion events should be called "ozone holes". If you do use that term, please drop the <i>a</i> ; just " <i>appearance of large ozone holes</i> "	Modified
2.	L138. What was the NCEP reanalysis used for? I couldn't find any mention of it in the rest of the paper.	We include following in Section 2.1: The propagation of wave activity was analyzed by using the zonal mean meridional heat flux and three-dimensional Plumb flux (Plumb 1985) calculated using NCEP reanalysis data.
3.	Section 2.1. Are these diagnostics calculated from all three reanalyses mentioned above or just one? Which one?	NCEP-R
4.	L138. The canonical reference for ERA5 is Hersbach et al (2020).	Modified
5.	L142. Please explain why this (somewhat narrow) latitude range was chosen.	This plot was moved to Supplement. The latitudinal range 70-90°N is rather appropriate for winter 2019-20 because vortex most of the winter was relatively undisturbed with center near the Pole. However for some winters with sever ozone reduction the polar vortex was more disturbed and shifted from the Pole.
6.	L144. Are the climate averages from reanalyses? Which one, specifically?	NCEP-R
7.	Section 2.2. It's hard for me to understand from this brief description how ozone loss is calculated. How were the initial parcel locations selected? Were they initialized with ERA5 ozone and then retained their ozone content (a variation of the passive tracer method)? What does "ensemble averaged" mean in this case? Please expand this section significantly. Some of this is explained in Section 3.2, but I think it belongs here. Please, also include or cite validation results of ERA5 ozone during Arctic winter/spring. It's a relatively new reanalysis and it cannot be assumed that its ozone is suitable for science, at least not without some solid evaluation.	Section 2.2 was changed

8.	L203. Please provide a citation for OMI, e.g., Levelt et al. (2018)	Done
9.	LL209-210. This is incorrect: while ClOOCl photolysis requires sunlight, air parcels depleted in ozone can get advected out of the illuminated area. Even Figure 1 suggests that this is the case: the ozone minima occur near the terminator, as you say in the next paragraph.	Thus statement has been deleted from the manuscript.
10.	L226 and below. What data are used here? One of the three reanalyses mentioned above? Which one?	NCEP-R
11.	LL226-232. Lower-stratospheric minimum temperatures in different extreme winters were also compared by Innes et al., 2020. What does the present analysis bring to the table that is new?	This plot was moved to Supplement.
12.	LL230-247. This analysis and most of Fig. 3 repeats the results of Lawrence et al. (2020). For example, see Figs 7 and 9 therein. If there is anything in the present analysis that isn't already in that study (I may have missed something), please indicate clearly what it is. Otherwise, I suggest eliminating this text and Fig 3a and b.	This part of the text was reduced, Fig.2a (with Tmin) was moved to the Supplement. Figure 2a displays the temporal evolution of heat flux (upward wave activity propagation) in the winter season 2019-2020. This figure is necessary to highlight the period of strongly reduced wave activity propagation from early February till early March 2020. The aim of the Figure 2b is to show the daily integrated values of zonal mean meridional heat flux at 70 hPa averaged over 45-75° N for three periods selected periods (normalized by the number of days in each period) of 2020 and the other five winters with strong polar vortex and ozone loss.
13.	LL288-290. I don't think you really mean <i>assume</i> . Something like this shouldn't be simply assumed, it needs to be demonstrated.	Fig2c (former Fig.3c) illustrates strong positive temperature anomalies at 925 hPa.
14.	Fig. 4 b and c. It would be better to use the same contour colors in both panels. I also suggest showing a difference plot.	These plots with estimation of difference significance were moved to Supplement
15.	LL315-317. I think you mean Fig. 6a, not c and March 10-13 not 11-13.	Modified.
16.	L321. The figure doesn't have panels e or f. I think it should be c and d.	Modified.
17.	Section 3.2. The trajectory analysis. I don't find these results very convincing. Why were these isentropic surfaces and these particular dates chosen? This choice of initial points is very restrictive. To make sure that this calculation represents the average vortex ozone loss, you would have to demonstrate that the trajectories sample the entire lower portion of the vortex (e.g. $\theta < 550$ K) uniformly throughout the chemically processed vortex air and uniformly in time. Without that it's hard to say what Figure 7 shows other than chemical loss along some arbitrarily chosen trajectories, that could be different if the trajectories were initialized differently. I would imagine that a good strategy	Section 3.2 and figure 5 were changed. In revised version the backward trajectories were initiated at the 400K, 435K, 460K and 500K levels at the latitude circles between 70° N and 85° N with 0.5° resolution in latitude and 0.5°/cos(latitude) spacing in longitude. The dates for the

	<p>might be to select initial points randomly within the vortex and at different times, but I can be convinced otherwise if you can show that your method does provide sufficiently uniform sampling and that the choice to start all the trajectories at the same time doesn't lead to most of them missing layers and times with particularly strong or particularly weak ozone depletion. I'm thinking of a situation when all the parcels initialized at 475 K in December were below 400 K by the end of February before serious depletion (maximized above 400 K, see Manney et al. 2020) even started. Another point: by selecting the initial parcel locations in a more robust way you could estimate ozone loss as a function of altitude / potential temperature and compare the result to that obtained from the ozonesondes analysis, thus providing some cross validation.</p>	<p>trajectory initiation were chosen when that domain was mainly located inside the polar vortex. Also we recalculated both forward and backward trajectories. Results are presented in the supplement. The ozone behavior in both directions is almost the same for all winters, proving that the vortex air is well mixed.</p>
18.	<p>L338. This is area-weighted average, correct? Please state that clearly if true. Also, I think the few trajectories that did venture out of the vortex should be excluded from the average.</p>	<p>In revised version the vortex interior was determined as the region with PV>14 PVU at 400 K, PV>26 PVU at 435 K, PV>36 PVU at 460 K and PV>46 PVU at 500 K, this criterion was used to filter out the initial locations outside the vortex and the trajectories leaving the vortex later on.</p>
19.	<p>LL 341-342. The fact that these rather large oscillations are there indicates that a good number of trajectories left the offspring vortices after the final SSW. It's hard to imagine what chemical mechanism would produce ~0.8 ppmv up and down changes over the course of a few days in the lower stratosphere!</p>	<p>In revised version we initiated the backward trajectories before SSW. The figure was changed.</p>
20.	<p>Figure 7. Please replace the commas in the y-axis tick labels with decimal points.</p>	<p>Done</p>
21.	<p>LL364-365. (1) This logic seems backward: One should select those observations that are inside the vortex, not handcraft the definition so that it accommodates the sonde locations. (2) how is the vortex edge defined at levels other than 475 K? Polar vortices exhibit complex 3-D geometries with edges at different levels often not lining up. Some of those stations were definitely outside of the polar vortex for some period of time (e.g., Ny-Ålesund and Sodankylä in mid-March). Were these measurements excluded from the regression analysis?</p>	<p>We excluded from analysis ozone data for the dates when this or that station was outside the vortex. We used only ozone observations inside the vortex. The vortex edge was defined as 42 PVU at 475 K isentropic surface. The PV limits for other levels were taken from source code <i>pypick</i> (available in /nadir/scr/nongraph/meteorol.</p>
22.	<p>L376. The maximum loss of 3 ppmv shown in Fig 8 occurs at 450 K. In Wohltmann et al. (2020) the vortex average (mentally subtracting the lines in their Fig. 4b) is about 2 ppmv at that level. Is it possible that this discrepancy results from different definitions of the vortex edge?</p>	<p>Certainly most of ozone data we used were obtained in the most depleted parts of the vortex. Unfortunately many sonde flights near the vortex edge terminated at quite low height. But on the other hand our results are consistent well with results of (Manney et al., 2020) -2.8 ppmv cumulative chemical ozone loss on 460 K level. Manney, G.; Livesey, N.; Santee,</p>

		M.; Lawrence, Z.; Lambert A.; Millan, L.; Fuller, R. Record low Arctic stratospheric ozone in 2020: MLS polar processing observations compared with 2016 and 2011. <i>Geop. Res. Lett.</i> 2020
23.	L383. If I understand correctly the CTM used water vapor from MERRA-2? If that is true, then it could seriously skew the results. Note that MERRA-2 stratospheric humidity is not very much informed by observations. It is, instead, relaxed to a zonally symmetric climatology (3-day relaxation time). It is therefore, especially suspect in extreme situations such as the 2020 winter/spring. Stratospheric water vapor in reanalyses is generally not recommended for scientific use with some exceptions (see Davis et al., 2017).	The water vapor content was specified according to the reanalysis data only for the troposphere, while in the stratosphere it was calculated taking into account chemical reactions and transport. In the updated version of this article, this is explained in the Methods section.
24.	LL392-404 and Fig. 9. I'm sorry but I don't think the model compares well to observations at all. Below I juxtaposed the OMI total ozone map from Fig 1 (left) and the CTM ozone from Fig 9 (right) for March 15. This is a particularly striking example, but things don't look much better on the other days considered, except perhaps on 3 March. One would expect at least the dynamical features to line up as the CTM is driven by reanalysis winds – but they don't (compare the shape of the total ozone contours and their gradients). Then looking at the ozone values the two figures have almost nothing in common: OMI shows a large complex patch of deeply low values extending between the Hudson Bay and northern Siberia while the CTM has a single weak minimum over the coast of Alaska, where OMI does not show anything noteworthy. Much weaker gradients in the CTM suggest that the model may produce far too much mixing across the vortex edge. The overall positive bias at high latitudes suggests insufficient depletion or too much resupply through descent, or, again, too much horizontal mixing. Overall, based on this comparison against OMI, I see no reason to trust the results from the CTM in this case.	Unfortunately, these maps used an unsuccessful interpolation scheme. In the updated version of the article, the maps are corrected and moved to the Supplement.
25.	Figure 10. What is shown there? The text talks about “PSC surface area”. I take it to mean PSC surface area density, but I'm confused about the units, mkm^2/cm^3 . This would be a dimensionless quantity. Maybe I misunderstood something.	To estimate the rates of heterogeneous reactions, the surface area of aerosol particles and particles of polar stratospheric clouds per unit volume is used. The dimension of this parameter is $\text{mkm}^2 / \text{cm}^3$. In the Methods section of the updated version of this article, we have clarified this.
26.	L425. What range of isentropic levels or altitudes?	The vertical range is indicated in the updated version of the article.
27.	L428. “nitrogen and hydrogen gases”. I think you mean nitrogen oxides and OH.	This refers to all catalytic cycles of ozone destruction, including nitrogen and hydrogen cycles.
28.	L429-432. This seems to suggest that the location of the vortex is an additional constraint on top of chemical depletion. I don't think this is correct. Rapid depletion is tightly confined to the interior of the polar vortex because that's where all the chemically processed air is. In fact, it is essential that the vortex air mass be isolated from the mid latitudes.	These map have been removed from the article.

	<p>Numerous studies demonstrated that chemical composition, particularly the ClOx family in the lower stratosphere exhibits a sharp discontinuity coincident with the vortex edge. Therefore, the position, extent and shape of the polar vortex are already imprinted in the spatial distribution of ozone loss rates. As a side note, that is not to say that dynamical factors don't play a role (resupply through descent, mixing).</p> <p>This brings me to Fig. 11. Since these are snapshots not time-averages, I would expect to see sharp gradients in the loss coefficient maps that would align with the edge of the polar vortex. Instead, the fields vary gradually. Below I plotted 12Z maps of the vortex edge defined as in this paper (black) and via scaled PV contours (see e.g. Manney et al., 2020) at several isentropic levels. I used MERRA-2 so this should be consistent with the CTM. On the right-hand side, I copy/pasted the depletion coefficient fields from Fig. 11. On both days the CTM produces significant depletion outside of the vortex (however defined). For example, on 15 March the CTM shows elevated depletion over Alaska and over Eurasia where it extends almost as far south as Lake Baikal, in both cases far outside of the polar vortex. On 1 April, again, the depletion coefficient map does not bear much resemblance to the vortex. Also, note the similarity between the shape of the vortex on 15 March and the OMI ozone map from the same day.</p> <p>I see two possibilities: either I grossly misunderstood what is plotted in Fig. 11 or the CTM's chemistry doesn't represent ozone depletion correctly. It may be instructive to look at ClO or ClOx maps in the CTM run. Active chlorine should be confined to the vortex with very limited mixing across the edge. You could compare ClO from the CTM with MLS. Another possible test would be to use the CTM and noCHEMall runs to calculate chemical ozone loss and compare that with the results of section 3.2. In fact, this should be done in order to establish self-consistency of the paper and get a sense of uncertainties.</p>	
29.	L432. I don't understand this sentence. What does " <i>destruction of the base</i> " mean?	Corrected.
30.	Figures 12-15 and the accompanying discussion add more doubts about the correctness of the CTM chemistry. Comparing the blue and green lines, it looks like heterogeneous chemistry plays a relatively minor role compared to gas-phase chemistry. At Pechora you estimate it to be responsible for only 25 DU of the total 70 DU of chemical loss! This raises a lot of red flags. According to our well-established understanding of polar ozone, heterogeneous chemistry is responsible for most of ozone destruction, particularly during cold winters with no major SSWs, although NOx can be important during weak-vortex winters (Sagi et al., 2017), which is not the case here. If the results presented here are correct, they require much more rigorous analysis and justification than that presented.	We added vertical analysis into updated version of the manuscript. It can be seen from these figures that at altitudes of 15-25 km, the main role in the destruction of ozone is played by its destruction in halogen cycles, and the role of nitrogen cycles is negligible. However, at altitudes of 25-40 km, significant ozone depletion is noted, in which nitrogen catalytic cycles play the main role.
31.	LL521-526. Again, at best this is a very surprising result that needs to be substantiated and supported by additional comparisons with observations.	Additional discussion has been added.
32.	L536. <i>Would have been observed or would have occurred.</i> But is this really demonstrated in the paper? If indeed NOx chemistry played the main role (which I find doubtful) then a disruption of the vortex could cause more depletion by bringing more NOx lower down leading to more loss. I think that this sentence is actually correct (while not substantiated), but it appears at odds with the results of this study (the predominant role of NOx).	We are talking about altitudes 25-40 km related to NOx catalytic cycles.
33.	LL548-L553. This first conclusion repeats one of the results of Lawrence et al., 2020 almost verbatim. There is really nothing new here	Three sentences (first conclusion) were moved upward in the beginning of our short discussion (chapter 5).

Technical corrections

There are a large number of grammatical and style issues. Below I list a few. I feel that it's not necessary at this point to mention all of them as the paper is likely to change very substantially if it gets resubmitted in the future.

L31 *Sudden stratospheric warming a sudden stratospheric warmings*

Modified

L33. Here and in several other places I think *main SSW* is supposed to be *major SSW*

Modified

L49. *statically a statistically*

Modified

L66. I suggest replacing *certain* with *selected*.

Modified

L75. *typical to Antarctica a typical for Antarctic conditions*

Modified

L87-89. Please, revise this sentence. It doesn't read well.

This sentence with comparison of winter 2019-2020 with previous one was removed according to Review 1.

L99. *supposed.* I think you mean *suggested*.

Modified

L118. *have been formed.* I think *have occurred* would sound better.

We remove several sentences on ozone loss in Antarctica according Review #2

LL120-124. This is a very long sentence. Please consider breaking it up into two.

This sentence was removed.

L129. *methodology of applied diagnoses* does not read well. How about something like *diagnostic methods*?

Modified

L136. This sentence would read better if the word *features* were dropped

Modified

LL135-139. Please, expand all the acronyms that were not previously defined (even if they are already expanded in the abstract). Also, please provide references for the models here, where they are first introduced.

Modified

L153. *Following to (Runde et al., 2016) a Following Runde et al. (2016)*

Modified

L172. *radiation transfer model a radiative transfer model*

Modified

L203. *on the board of a onboard.*

Modified

L204. *at the early March a in early March*

Modified

L315 *by Plumb fluxes a using Plumb fluxes*

Modified

L316. *dominated* I think you mean *dominant*

According to Review 1 "display dominated" was changed to "show pronounced"

LL456-458. I think something like *by the end of March (...) total ozone content at Pechora drops by almost 50 percent* would be more clear and read better.

Modified

L528. Please rephrase this sentence. It doesn't read well.

This sentence was removed

L532. I suggest changing *values* to *magnitude*.

We remove this sentence.

Thank you again for taking the time to review our manuscript.

With respect,

Sergei P. Smyshlyaev,

Pavel N. Vargin,

Alexander N. Lukyanov,

Natalia D. Tsvetkova,

Maxim A. Motsakov