We would like to thank both reviewers for their constructive comment on our manuscript. In the following, we address the respective proposals for improvement. Changes are explained in detail, answering each referee point by point. Reviewer comments are in normal font. Our answers are in italic and changes to the manuscript in blue.

Response to Referee #1

Main topic areas to address in revision

- **Source of tropopause data**, lines 155-165. Each profile is analyzed with respect to its height above the tropopause yet nowhere is it stated where this tropopause information comes from. What is the source of the 'local tropopause'? Is the aircraft doing frequent profiling to identify a tropopause (through a temperature minimum)? This must be explained. Later in the paper (line 322) there is talk of using a thermal tropopause, but this doesn't note the data source. Surely the climatological mean tropopause is not used for the constituent analyses. This issue is very important as the results depend on what information is used to determine this coordinate.

  The unknown source of your tropopause data leads right into 3 related issues: using the thermal tropopause for your analysis, the decision not to use potential vorticity from a reanalysis, and the use (and source of) of equivalent latitude.

  We addressed this issue by including a new subsection "Meteorological data" in the section "The SouthTRAC campaign". In this subsection, we shortly explain what kind of local tropopause we used in this analysis. Regarding the following comment, we switch to the more appropriate dynamical tropopause. We go into more detail on the next comment.

  "[…]
  2.1 Meteorological data
  […]
  The local tropopause information along the flight tracks of HALO was created using the Chemical Lagrangian Model of the Stratosphere (CLaMS) (e.g., Grooß et al., 2014). The underlying meteorological data are taken from ECMWF ERA-5 data (Hersbach et al., 2020). In this work, the potential vorticity (PV) based dynamic tropopause is used (e.g., Gettelman et al.,2011), taking the 2 PVU (potential vorticity unit) for the dynamical tropopause. Since the PV tropopause is not physically meaningful in the tropics, potential temperature level of 380 K was taken as tropopause if the 2 PVU level lies above. […]"

- **Use of thermal tropopause.** The thermal tropopause is inappropriate in polar winter because the temperature profile is often isothermal – the dynamical (PV) definition is needed. See the analysis of dynamical v. thermal tropopause in Zaengl and Hoinka (J. Climate 2001). This means you need PV from reanalysis data (ERA5, MERRA2...whatever). These fields are available in fairly high resolution (0.5 degree or

better) and even with interpolation they may more accurately identify the tropopause than does temperature in an atmosphere with weak vertical temperature gradients. Whatever your final analysis method is, you will need to justify it based on 1) showing that reanalysis PV doesn't give sensible results, or 2) proof that the temperature tropopause actually makes sense at high latitudes in winter.

We fully agree that the thermal tropopause is inappropriate in polar winter. Therefore, the analysis was carried out again using the dynamical tropopause, based on PV from ERA-5 data. Where we have listed the necessary information for dynamic tropopause in the manuscript can be found in the previous comment.

- **Source of equivalent latitude**. Around line 156 equivalent latitude is said to be used to sort the flight data: where does your equivalent latitude come from? Just 20 lines earlier it is stated why use of reanalysis data and its coarse resolution is a drawback to the analysis, but where to you think the equivalent latitude information comes from? It is calculated based on global PV fields which, by necessity, come from a reanalysis. So although you haven't explained the source of either the tropopause or equivalent latitude data used, it seems clear that you are using reanalysis info. This should be acknowledged. It's fine if you want to use the Greenblatt method for identifying profiles, but I'm not sure it's accurate to say that the reanalysis PV isn't good enough for your analysis. (Have you tested this?)

  We are aware that the equivalent latitude information is based on global PV fields, which come from the reanalysis, mentioned in the answer to the previous comment. At this point, however, we only us the equivalent latitude to pre-filter the measurements which helps the filter procedure to find the lower and upper envelope. The vortex and mid-latitude profiles are still based on the N_2O measurements.
  We added "(Butchart and Remsberg , 1986)" as reference to the equivalent latitude. Earlier in this section, we replaced the statement about the resolution of PV with the following:

  "[…] PV is a model-derived quantity. Although the underlying meteorological reanalysis have a fairly high resolution these days (e.g., Hersbach et al., 2020), small-scale features like vortex filaments with different chemical compositions may not be well resolved […]"

- **Antarctic and Arctic vortex size differences**. These play a role in whether Figure 10 is meaningful. The Antarctic vortex mean edge is at 60S equivalent latitude – it's a large vortex. (Sep avg ~35 million km2). Even in 2019 the Antarctic vortex at 360K had an average size until the last third of September. I'm not certain what the Arctic vortex mean edge is but it's probably closer to 70N equivalent latitude (avg March vortex <20 million km2). Because of this the hemispheric difference plot using equivalent latitude coordinate doesn't make physical sense in the 60-70 degree range. In Fig. 10 the difference at 65 degrees will be a comparison of the Antarctic vortex with the northern midlatitudes. Since the hemispheric vortex profiles are already compared in Fig. 9, perhaps add panels to that figure showing the NH/SH midlatitude differences on the 2 vertical coordinates. I don't think Fig. 10 is very useful and could be eliminated.

We agree that the different sizes of the Artic and Antarctic vortices limit the hemispheric comparison. In the text we have already indicated that this comparison is limited in the polar region and vortex edge region due to different sizes, strength, and transport barrier of the vortices. However, the figure also shows a comparison of lowermost stratosphere of the mid-latitudes of the Northern Hemisphere and Southern Hemisphere. We extended the discussion about the differences regarding mid-latitudes Cly values. Thus, we would like to keep figure 10 in this study.
Section 4.4 was extended by:

"[…] It must be noted that the polar vortices of the two hemispheres are different in size, stability and strength of the transport barrier. The comparison on equivalent latitude is therefore only possible to a limited extent. Nevertheless, possible reasons for the observed differences can be derived from the hemispheric difference of the Brewer-Dobson circulation, using the age of air as a common metric for transport. Konopka et al. (2015) showed, that north of 60 °N, age of air is always younger than south of 60 °S in the same season, implying a stronger residual circulation in the Northern Hemisphere. Analysis of Haenel et al. (2015) revealed differences in age of air trends in the lowermost stratosphere of the mid-latitudes of Northern and Southern Hemisphere with a positive trend in the Northern Hemisphere and a negative trend in the Southern Hemisphere. In addition, Mahieu et al. (2014) reported long-term total column data for HCl and $ClONO_2$ (representing Cl_y) in the stratosphere, at Jungfraujoch (46.5 °N) and at Lauder(45 °S), though the end of 2016. A negative trend of Cl_y is observed at both stations but with a non-significant trend for the Jungfraujoch data over the last decade and a slightly larger negative trend from the Lauder data. Furthermore, lower-stratosphere HCl from the Global Ozone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS)shows larger decreases at southern latitudes and increases at northern mid-latitudes (Froidevaux et al., 2015). Thus, higher values of Cl_y in the mid-latitudes during PGS seems to be plausible. […]"

- **Figure 7 discussion (l. 270).** What data are used to calculate the mean age shown in Figure 7? There is discussion just prior to this about mean age and the 'arrival time' – is this what's plotted? Maybe I'm missing something but I cannot see what observations or information are used to produce mean ages. But a bigger problem is that mean age values of 5 years are shown for Cly of 1500 ppt. This can't be right. The best estimate is closer to 3000 ppt at 5 years. See for example Newman et al (ACP 2007) or Strahan et al (2014, JGR) or compare the N2O values you observed with the ACE N2O/mean age mapping in Strahan et al (JGR 2011). No data have been presented that demonstrate that SouthTrac data, which are entirely from 390K and below, have such old age. It's more likely the maximum age there is near 3 years.

  We used SF_6 measurements of the GhOST-ECD channel in a time resolution of around 1 minute and a precision of 0.64%. The usage of SF_6 to calculate mean age was already mentioned in the manuscript in section 4.2 in the paragraph after equation 1: "[…] The concept of the age spectrum can be used to determine mean age values based on observations of chemically inert tracers in the stratosphere. For this purpose, in addition to the age spectrum, tropospheric time series of the inert tracers

are required (Engel et al., 2002). This was done for the SouthTRAC campaign by using SF_6 measurements of the GhOST-ECD and tropospheric time trends taken from the AGAGE (Advanced Global Atmospheric Gases Experiment) Network (Prinn et al., 2018). […]".
SF_6 is a commonly used tracer for age of air, but recent research suggests, that its lifetime has been overestimated and thus it may be giving higher mean ages. (e.g., Ray et al., 2017 and Leedham Elvidge et al., 2018). Leedham Elvidge et al., 2018 also provides correction functions of SF_6 derived mean age. We compared N_20/mean age mapping during SouthTRAC with ACE N_2O/mean age mapping in Strahan et al., 2011. Without SF_6-based mean age correction, toward older air, the difference reaches roughly one year, whereas with correction, difference becomes roughly half a year. It must be mentioned that we compare ACE annual mean N_2O from mid-latitude and tropical observations with mostly mid-latitude and polar observations from September to November 2019 during the SouthTRAC campaign. In addition, Konopka et al. 2015 showed, that signatures of old air within the Antarctic vortex propagate down to 340K of potential temperature and in the polar regions around 380K, oldest air anywhere can be found in September south of 60°S. The maximum Cl_y value during SouthTRAC was 1668 ppt at 4.9 years of mean age, 4.2 years when corrected by a correction function from Leedham Elvidge et al., 2018. The corrected mean age seems to fit better for the given Cl_y value, although not fully comparable with Strahan et. Al., 2014. Furthermore, Cl_y calculation using N_2O instead of CFC-12 as the reference substance leads to similar values (see new Figure S7 in the supporting information). To alert the reader that the mean age shown here may be somewhat overestimated, we have added a second x-axis in Figure 7 showing the profiles at corrected mean age. However, this is only intended as a guideline, as it represents only one possible correction of the mean age. in the paragraph to figure 7, we have thus included the following:

"[…] Recent research suggests that SF_6-based mean age is biased because the suggested lifetime has been overestimated (e.g., Ray et al., 2017). As a guideline, Figure 7 additionally shows a corrected mean age of air using one of the linear fit functions from Leedham Elvidge et al. (2018), based on a comparison of SF_6-based mean age with a combined mean age based on five alternative age tracers. In this study, however, the uncorrected mean age of air is used. […]"

- The vortex Cly profile differences (Fig. 9) imply interhemispheric (IH) differences in mean age (and age spectrum) in the lower branch of the Brewer Dobson Circulation (BDC). These are presumably driven by transport and indicate that the NH lowermost stratosphere is younger than the SH. I believe such differences are expected – see for example Birner and Boenisch, ACP 2011. Simulations driven by reanalyses may reproduce these differences (as well as the midlatitude differences), but what about chemistry climate models (CCMs)? It would strengthen this paper to put your measurements in the context of what they tell us about IH differences in the lower BDC. These measurements help confirm our thinking about the stratospheric circulation. You might comment on whether and why it's important for CCMs to reproduce similar hemispheric differences.

We agree, that including a discussion with chemistry climate models would additionally strengthen this manuscript. However, it would go beyond the scope of this evaluation and cannot be dealt within a few sentences in section 4.4.
As an outlook, we included a statement regarding chemical climate models in the summary and conclusion to pinpoint, that such interhemispheric difference should be captured therein:

"[…] These hemispheric differences can also be found in simulations based on reanalysis, e.g., Konopka et al. (2015). A comparison of the available data with chemical transport models should be subject to further studies. Furthermore, such interhemispheric differences should also be captured by chemistry climate models, which are not only used to understand past changes but also predict future changes in chemical composition […]"

- LInes 220-240. Isn't the semi-direct Cly calculation nearly identical to Schauffler's method (JGR 2003)? While that was referenced much earlier it seems far more relevant here. If this is true then you can reference Schauffler here and shorten the description, only describing any way your method differs.

  The methods are quite similar but with the difference, that Schauffler et al., 2003 did not use relevant age.

- This paper uses measurements to calculate Cly from only the long-lived Cl-containing species, but there are contributions from short lived (VSL) Cl species too (e.g., CH2Cl2, C2Cl4). It should be explicitly stated that such species are excluded from this study. The estimated size of this neglected contribution could be noted. See Hossaini et al., JGR 2019 for an observational and modeling study that estimates the VSL Cly impacts.

  You are correct. The contribution from the short-lived chlorinated substances is small but significant, as revealed in Hossaini et al., 2019. We included the information, by focusing only on the long-lived substance in this analysis, in the last paragraph of the introduction. In addition, we listed the contribution from the very-short lived chlorinated substances to the total stratospheric chlorine as well as the contribution in inorganic form from Hossaini et al., 2019 and from the WMO Report 2018.

  "[…] Calculations of Cl_y are based only on long-lived chlorinated substances. There is an additional contribution to total stratospheric chlorine from the very short-lived chlorinated substances. Engel and Rigby (2018) estimated a contribution of 115 (75 – 160) ppt from very short-lived chlorinated substances for 2016. Hossaini et al. (2019) estimated a contribution of about 111 ± 22 ppt, of which 13 ± 4.6 ppt are already in inorganic form, which is not considered in this analysis. […]"

Minor Comments

- Title: It doesn't make sense to say 'comparison' without saying what you're comparing with. The abstract reveals it is Cly in the Arctic LMS a few years earlier.

Perhaps ‘Comparison of Cly in the Arctic and Antarctic lowermost stratospheric vortices’?

We changed the title for more clarity.

“[…] Comparison of Inorganic Chlorine in the Antarctic and Arctic lowermost stratospheres by separate Late Winter aircraft measurements […]”

- First sentence of the abstract. You’re really talking about stratospheric inorganic chlorine so please say so. And the strat inorganic chlorine comes from all chlorine containing source gases with a lifetime of more than 5 months (see Hossaini et al JGR 2019), so that’s the long-lived and many of the VSL species.

  We extended this sentence to clarify that we were considering stratospheric inorganic chlorine. We also add to the sentence the information that chlorinated very-short lived substances contribute to the inorganic chlorine.

  “[…] Stratospheric inorganic chlorine (Cl_y) is predominantly released from long-lived chlorinated source gases and, to a small extend, very short-lived chlorinated substances. […]”

- line 16. “Based on the results of these two campaigns, the difference of Cly inside the respective vortex is significant and larger than reported inter annual variations.” Each campaign was a single winter – there is no information on interannual variability. I realize you are citing another paper on Cly variability in Antarctic lower stratosphere, but what about Arctic variability? Unknown? As written this statement is misleading and not supported.

  Our results are based on one winter each in the Artic and Antarctic. Therefore, interannual variations are not examined in this study, as the referee mentioned correctly. Referee #2 also marked the word “significant” as not applicable in this context. We have taken up the suggestion of referee #2 for improvement and adjusted the sentence accordingly.

  “[…], the differences in Cl_y inside the two vortices are substantial and larger than the inter-annual variations previously reported for the Antarctic. […]”

- line 20: ‘1980-ies’ is 1980s

  Done.

- line 23 OClO isn’t involved in depletion. Null cycle.

  OClO has been removed from the list of substances.

  “[…] Chlorine substances involved in rapid ozone depletion are Cl, Cl_2, ClO, and ClOOCl, and can be summarized as ClO_x. […]”

- line 45. In both hemispheres, polar winter temperatures are above radiative equilibrium because of dynamical (wave-driven) heating. It's not just the absence of insolation.

  You are fully correct. Although temperatures within the polar vortex are basically driven by radiative processes, they are also determined by dynamics and the transport of heat by atmospheric motions. But we don't want to go into too much detail here about how the temperatures inside the vortex come about. For this reason, we will shorten the statement in the manuscript by excluding the statement about the ultraviolet heating. The new sentence is:

  "[…] Due to a temperature difference and consequently to a latitudinal pressure gradient between the polar and mid-latitude stratosphere (e.g., Schoeberl and Hartmann, 1991), a state with a strong westerly wind in the stratosphere is established (polar night jet). This jet acts as a transport barrier, leading to strong latitudinal gradients of potential vorticity and long-lived substances like N_2O results (e.g., Hartmann et al., 1989) […]"

- Since the paper is comparing Cly in the 2 vortices, do you have any comments/conclusions about differences in maximum potential O3 depletion in each LMS vortex?

  Comments and conclusions about differences in maximum potential O_3 depletion in the respective LWS vortex is beyond the scope of this manuscript.
  Beside the comparing Cl_y in the two vortices, mostly important is the chlorine activation e.g. the production of active chlorine from the reservoir species on polar stratospheric clouds.

- line 85. You don't need to define payload

  The definition of payload has been removed.

- line142. I would emphasize that you mean mixing within the vortex. E.g., "...benefits mixing on isentropic surfaces inside the vortex..."

  We have rewritten this sentence for more clarity.

  "[…] In addition, N_2O has a small variability inside the vortex on constant isentropic surfaces (variability of about 6 ppb (Greenblatt et al., 2002a)). This is an indication of well mixed air inside the polar vortex due to the long isolation in polar winter. […]"

- line 143. 'descent' not 'descend'

  The sentence containing this typo was removed due to repetition.

- line 146. What you're describing is that as you approach the tropopause the vortex ceases to exist so there is no longer a barrier to mixing. There is nothing to distinguish.

  We have rewritten this sentence for more clarity, that the vortex barrier vanishes towards the tropopause.

  "[…] Towards tropopause altitudes, the transport barrier of the polar vortex disappears, and a classification is not possible. […]"

- line 152. By "Stratospheric transport and mixing is related to the isentropic surfaces" do you mean that transport and mixing occur on isentropic surfaces? This is unclear, please rephrase.

  We do mean quasi-isentropic mixing. As this is not the only effect on the composition of air in the UTLS, we further extended this sentence.

  "[…] The composition of the lowermost stratosphere is affected by the diabatic decent inside and outside the polar vortex and quasi-isentropic mixing with air from lower latitudes. […]"

- line 155 'had contact to the vortex core'? This is awkward and unclear. Is the intended meaning that the reference profile was entirely inside the vortex, away from the edge and mixing at the edge?

  This sentence was rewritten for more clarity.

  "[…] The vortex reference profile (see Fig. 2) was generated from all flights that are assumed to contain measurements within vortex air. […]"

- line 189, 'extensively'

  By rewording this sentence, this word is omitted.

- lines 194-7. "The metric describing the combined effect of all ozone depleting substances (ODS) as an equivalent amount of inorganic chlorine in the stratosphere, related to tropospheric source gases in a simple, is the equivalent effective stratospheric chlorine (EESC)". Awkward sentence. I suggest: "Equivalent effective stratospheric chlorine (EESC) is a simple metric that sums the effect of all ozone depleting substances (ODS) as an equivalent amount of inorganic chlorine in the stratosphere..."

  We have rewritten this sentence according to the suggestion.
  In addition, this sentence, and the following regarding the EESC was included in the introduction and removed at this point.

- General comment on 'pre-filtered', 'pre-required'. Drop the 'pre', it's not needed.

The data is filtered only because the subsequent procedure for the upper and lower envelope works well programmatically. We would like to keep "pre-filtered".

- To clarify the meaning, I'd suggest a slight rewording (line 220): "Cly can be calculated as the difference between total chlorine entering the stratosphere and the organic chlorine that remains bound in chlorinated halocarbons"

  To make this part of the section clearer, we have rewritten it.

  "[…] Organic chlorine (CCl_y) can be calculated directly from the up-sampled GhOST-MS measurements. Thus, Cl_y can be calculated from Eq. 1 if the mixing ratios of the major chlorine-containing substances at the stratospheric entry point (Cl_{total}) are known. […]"

- line 256: 'rations'...you meant 'ratios'

  Done.

- line 239: "in the following..." Move this statement to the beginning of the next paragraph where you actually describe the semi-direct method and then reword. For example you can begin the next paragraph with: "The semi-direct Cly calculation is used in the case where no measurements of chlorine containing substances are available. This method is based on [trace gas?] correlations found in previous measurement campaigns."

  The following paragraph does not describe the semi-direct Cly calculation. Instead, it describes the indirect Cly calculation based on scaled correlation from previous measurement campaigns, as stated at the beginning of the paragraph. The semi-direct Cly calculation is described in the first paragraph of this section. The sentence mentioned here serves to introduce the term "semi-direct Cl_y" as a term for Cl_y from in-situ measurements, which reappears later in the paper. We would leave this sentence at this point in the text.

- Figure 9. It would be more useful to give titles to each panel other than 'a' and 'b'. Those labels normally go inside the panel.

  A figure earlier in the paper also uses (a) and (b). For the sake of consistency, we would like to leave it at that, with a detailed description in the figure caption.

- Fig. 9 shows that the NH data reaches 405K while in the SH 385K is the maximum. Do these represent the same maximum altitude for flights, and does this difference indicate that the SH LMS vortex is much colder than the NH vortex?

  Figure 9 shows only the maximum potential temperature captured inside the respective vortex. However, this also depends on the flight patterns. Flights during the PGS campaign were operated from Kiruna (68° N). In contrast, flights during the SouthTRAC campaign were operated from Rio Grande (53° S). The longer distance to and from the vortex had to be considered for the flight planning and has an influence

on the maximum potential temperature that could be reached inside the respective vortex. The maximum height of the flights cannot be taken from this.

- Fig. 10 is saying that the NH midlatitudes are older than the SH. Anything to say about that?

  We extended the discussion on the differences of NH and SH mid-latitudes. The extended passage can be found in the comment "Antarctic and Arctic vortex size difference".

- Lines 298-302: Since you are identifying profiles as vortex, midlatitude, or edge already, I imagine the effect of the SSW is that you measured more edge and midlatitude profiles in November than you might have in another year. But you've pointed out that vortex descent has essentially ceased by September, so as long as you are sampling vortex air that hasn't mixed with midlatitudes, I would expect that the Cly profiles you measure aren't affected by the SSW. In other words, the mean age profile for air masses that are truly vortex air masses might well look similar to other years. Thus, the statement implying that the fraction of CCly found as Cly being affected by the SSW right may not be right. On the other hand, there aren't data from other years and maybe this point should be made. There is no information on interannual variability.

  Thank you for mentioning this point. We do not want to create a link between the amount of Cl_y and the size of the ozone hole or the minor SSW event. Instead, we wanted to note that the minor SSW event led to an early chlorine deactivation.
  We further use a more appropriate reference for the size of the Antarctic ozone hole.

  "[…] Inorganic chlorine within the vortex could be obtained from Θ between 330 to 385 K. Cl_y inside the vortex increases significantly up to a value of 1687±19 ppt. Thus, in late winter and early spring at this altitude about half of the recorded chlorine is found in inorganic form. Despite this amount of inorganic chlorine in the lower stratosphere, the total polar ozone column was higher than usual in September 2019. As a result of the minor SSW event, chlorine deactivation began earlier in 2019 and the ozone hole was about 10 x 10^6 km^2 in size, thus only 20 % of that in 2018 mid-September (Wargan et al., 2020). […]"

- In general, 'data' is a plural noun, thus, 'data are...' not 'data is'.

  Done.

Response to Referee #2

Title

- The comparison being made in this paper is not clear from the title. It would be better to craft a different title capturing the idea that in situ measurements from separate aircraft campaigns are being used to compare Cly abundances in the Antarctic and the Arctic LMS.

  We changed the title for more clarity.

  "[…] Comparison of Inorganic Chlorine in the Antarctic and Arctic lowermost stratospheres by separate Late Winter aircraft measurements […]"

Abstract

- L1: The wording "... long-lived chlorinated source gases. These include the reservoir species" seems to imply that HCl and ClONO2 are chlorinated source gases. Thus, These include --> Cly includes

  Done.

- L5: in late winter --> in austral late winter

  Done.

- L8: The sentence "Cly from a scaled correlation was compared to directly determined Cly ..." is confusing, since the previous sentence states that not all source gases were measured during PGS. It needs to be made clear that this "validation" was performed for SouthTRAC.

  We have rewritten this sentence to create less confusion and to show that SouthTRAC data were used to validate Cly from a scaled correlation.

  "[…] Using SouthTRAC data, Cly from a scaled correlation was compared to directly determined Cly and agreed well. […]"

- L12-13: The values (40%, 20%) here appear only in the abstract and conclusions, not in the main text. In my opinion, it is not appropriate to include "new information" in the abstract and summary sections that has not been thoroughly discussed in the body of the paper. Please add some corresponding statements in section 4.

  We include this information into section 4.4 after the discussion of Figure 9.

  "[…] The fractions of total chlorine which are in the form of Cl_y inside the vortex and in the mid-latitudes during PGS at the same distance from the local tropopause as for the highest values within the vortex during SouthTRAC, are about 40% within the vortex and about 20% in the mid-latitudes. […]"

- L13-14: Differences inside the respective vortex reaches up to 565 ppt more --> Differences inside the two vortices reach as much as 565 ppt, with more

  We have rewritten this sentence according to the proposal.

- L15-16: As far as is known --> To our knowledge; within the respective polar vortex --> within the Antarctic and Arctic polar vortices

  Done; Done.

- L16-17: "the difference of Cly inside the respective vortex is significant and larger than reported inter annual variations". I have a number of comments on this sentence:
  - The authors have not done a statistical analysis, so I do not think that "significant" is an appropriate word here.
  - This statement could be erroneously interpreted as implying that their study examines interannual variations. Moreover, the Strahan et al. [2014] paper on which this statement is based looked only at the Antarctic, not the Arctic.
  - The word "respectively" is used many times throughout the manuscript, sometimes (as here) incorrectly. I have recommended alternative wording in a few places.
  - Thus, I suggest instead: "the differences in Cly inside the two vortices are substantial and larger than the interannual variations previously reported for the Antarctic".

  We agree that "significant" should not be used in this context. Furthermore, our results are based on one winter each in the Artic and Antarctic. Therefore, interannual variations are not examined in this study. We are happy to use the appropriate suggestion and change the sentence accordingly.

  "[…], the differences in Cly inside the two vortices are substantial and larger than the inter-annual variations previously reported for the Antarctic. […]"

Section 1

- L20: 1980-ies --> 1980s; also, pre-dominantly --> predominantly

  Done.

- L23: substances, which are involved --> substances involved

  Done.

- L23-24: OClO is a consequence of, and thus a good qualitative indicator of, halogen activation, but it does not itself participate in ozone destruction, as this sentence implies. Thus it is not normally considered part of ClOx.

  OClO has been removed from the list of substances.

"[…] Chlorine substances involved in rapid ozone depletion are Cl, Cl_2, ClO, and ClOOCl, and can be summarized as ClO_x. […]"

- L25: within --> through

  Done.

- L26: While I applaud the recognition of some of the original papers, I think it would be good to also include some review articles (e.g., Solomon, Rev Geophys, 1999) and/or some more up-to-date citations (e.g., the most recent WMO Ozone Assessment).

  We added Solomon et al. 1999 as a reference because this paper clearly illustrated the activation of chlorine form the reservoir species.

- L30: Citing only the Newman et al. [2004] paper here gives the erroneous impression that it is the only relevant reference. At the very least, "e.g.," needs to be added to this citation. This is another instance where it might be appropriate to cite the WMO Report.

  "e.g." was added for this citation.

- L35: used again --> again used

  Done.

- L39-40: Citations for the long-term trends in Cly and N2O should be given.

  As a reference for negative Cly trends we included Newman et al. 2007. As a reference for a positive trend of N2O, we included chapter 1 of the 2018 Ozone Assessment Report.

- L45: between polar --> between the polar

  Done.

- L46: Here and throughout the manuscript, when "mid-latitudes" is used as an adjective to modify a noun (e.g., stratosphere, profile, reference, etc.), it should be singular: "mid-latitude". When it is used as a noun itself (as in "at mid-latitudes"), then it is plural.

  "Mid-latitudes" was changed to "mid-latitude". In the following, this term is adapted, according to its use.

- L46: add "e.g.," in front of "Schoeberl"

  Done.

- L48: add “e.g.,” in front of “Hartmann”

  Done.

- L62: SouhTRAC --> SouthTRAC

  Done.

- L65: delete duplicate period after “4”

  Done.

Section 2

- L70: capable to reach --> capable of reaching

  Done.

- L72: Rio Grande (RGA), Argentina (53°S, 67°W) --> Rio Grande, Argentina (RGA, 53°S, 67°W); also, regions for --> regions of

  Done; Done.

- L74-75: The actual dates (not just “September” and “November”) for the two phases of the campaign should be given.

  The two sentences about the campaign phases were supplemented with the start and end dates of the respective phase.

  “[…] The first phase took place from September 6th to October 9th, 2019 to target the dynamical objectives (e.g., Rapp et al. 2020). The second phase took place from November 2nd to 15th, 2019 to, among others, sample polar vortex remnants. […]”

- L77-80: 9 transfer flights + 10 Phase I flights + 3 Phase II flights = 22 total flights, not 23 as stated in L77

  There was a short local flight on Sal during the first transfer from Oberpfaffenhofen to Rio Grande, which I did not list in the text. I apologize for the inconsistency and add the following sentence:

  “[…] Within the first transfer from EDMO to RGA, there was an additional local flight operated from SID. […]”

- L78: Rio Grande (RGA), Argentina --> RGA

  As I mention the three-letter codes of the other airports at this point, I would keep the current wording for Rio Grande.

- L84: Beside --> Besides (or, “in addition to”); also, delete comma after “instruments”

  Replaced “Beside” with “In addition to”; Done.

- L108-109: Given the in-flight conditions as described in this paragraph, it makes sense that the mean precisions during the campaign are poorer than those measured in the lab. So it is 2 puzzling that the mean precisions during the flights improved over those measured beforehand for methyl chloroform. Do the authors have an explanation?

  Unfortunately, no explanation could be found. The chromatographic peaks of methyl chloroform were examined closely without finding any problem or abnormality.

- L108-112: It is stated that for CFC-113 “the amount of water in the analytical system should be kept as low as possible”, but it is not clear whether that was actually done during the campaign. The implication is that in fact this was not done properly and that is why the in- flight precision of CFC-113 is so much worse than that determined in the lab, but this point needs to be clarified. I also think it is questionable whether the measurement of CFC-113 really stands out as an “exception” (L108) for its degraded performance. In fact, the precision estimated during the flights is even worse relative to the pre-campaign value (a factor of 5 difference) for CFC-11.

  We have rewritten the section about the poorer precisions of CFC-113 and included CFC-11 to this section.
  We also mention in the text, that we dry the air before pre-concentration. More detailed information about drying can be found in the already listed publication about the GhOST-MS.

  “[…] The exceptions are CFC-11, CFC-113, and methyl chloroform. […] It is difficult to determine exactly what the poorer precisions of these substances can be attributed to. The chromatographic peak of CFC-11 is very narrow and variable environmental conditions (due to changes in altitude, pressure, and temperature in the cabin) have an influence on the peak shape. The amount of water in the analysis system is also important and is kept as low as possible by drying before pre-concentration. As the chromatographic peak of CFC-113 is close to the chromatographic peak of water, small changes in water can affect the chromatographic peak of CFC-113. […]”

- L109: precision CFC-113 --> precision of CFC-113

  Done.

- L109: It seems odd to me that the authors make the effort (L85) to define “payload”, which is a widely known and not particularly technical term, but not “elutes”, which many of their readers (including me) may not know. Also, should it be “CFC-113 is eluted by water” rather than “CFC-113 elutes near water”?

  The definition of payload was removed from the manuscript.

To avoid having to introduce and explain the term "elute" unnecessarily, the sentence was rewritten, and "chromatographic peak" was used for the description. Changes can be seen in comment to L108-112.

- L110: add a comma between "water" and "the amount of water"

  Done.

- L111-113: Precision values are not given for GhOST-ECD SF6 either in the text or in Table 1, yet SF6 measurements are mentioned later in the paper.

  We included the precision of SF_6 in the text in section 2.2 in addition to the precision of CFC-12 with the GhOST-ECD.

  "[…] CFC-12 and SF_6 with the GhOST-ECD channel were measured with a precision of 0.2 % and 0.64 %. […]"

- Table 1 caption: have been determined shortly before the SouthTRAC (ST) campaign and mean precisions during the flights --> have been determined in the laboratory shortly before the SouthTRAC (ST) campaign, and mean precisions were calculated during the flights

  We have rewritten this part of the sentence according to the proposal.

- L120: prior --> prior to

  Done.

- L122: was of --> was

  Done.

- L123: post-flight corrected --> corrected post-flight

  Done.

Section 3

- L125: The occurrence of chlorine activation also depends on factors other than temperature (e.g., humidity, the availability of suitable aerosol particles) and has been observed outside the polar regions, so it would be better to say "tends to occur" rather than "occurs" here.

  We have taken the suggestion from the reviewer.

- L129: conclusion --> conclusions

  Done.

- L134-135: Modern meteorological reanalyses have fairly high resolution these days [e.g., Fujiwara et al., ACP, 2017]; although they still do not resolve very small-scale features, it is not entirely fair to characterize them as having “rather coarse resolution”.

  We agree that the wording “rather coarse resolution” is not appropriate. Therefore, we have rephrased the sentences regarding PV for air mass classification:

  “[…] PV is a model-derived quantity. Although the underlying meteorological reanalysis have a fairly high resolution these days (e.g., Hersbach et al., 2020), small scale features like vortex filaments with different chemical compositions may not be well resolved. […]”

- L138-145: This part of the paragraph is poorly written, repetitive, and hard to follow. It should be reorganized to improve the flow.
  - The sentence “It can be measured ... atmosphere.” is out of place, as it comes in between two sentences that say essentially the same thing. It should be moved and the other two sentences combined to reduce repetition.
  - Not only is the statement “the isolation inside the vortex benefits mixing on isentropic surfaces and therefore a small variability on isentropes (variability of about 6 ppb)” grammatically incorrect, but also it makes no sense. I’m not sure what “benefits mixing” means? Perhaps “inhibits” was meant? In any case, this sentence needs to be rewritten.
  - The sentence “The low mixing ratios inside the vortex ... N2O” again repeats the same information already stated twice above. (Also, descend --> descent)

  Regarding these three comments, we have rewritten this part of the paragraph.
  The sentence “It can be measured ... atmosphere.” was moved to the top, followed by a small description of the polar vortex. Subsequently, N_2O inside and outside the vortex is described. The sentence about the variability of N_2O inside the vortex is rewritten for more clarity:

  “[…] A tracer like N_2O can be measured in situ with a sufficiently high time resolution to reveal small structures in the atmosphere. During the dark polar winter, stratospheric temperatures are below those of the mid-latitude stratosphere. The associated pressure gradient between the pole and mid-latitudes, as well as the Earth’s rotation, leads to enhanced circumpolar winds, also known as polar night jet or polar vortex (e.g., Schoeberl and Hartmann, 1991). Furthermore, the decreasing polar temperatures lead to a subsidence of polar air, also known as diabatic decent (e.g., Schoeberl and Hartmann, 1991; Bauer et al., 1994). A tracer like N_2O exhibits a horizontal gradient across the vortex edge in the stratosphere with lower mixing ratios inside the vortex and higher mixing ratios outside the vortex. In addition, N_2O has a small variability inside the vortex on constant isentropic surfaces (variability of about 6 ppb (Greenblatt et al., 2002a)). This is an indication of well mixed air inside the polar vortex due to the long isolation in polar winter. […]”

- the mid-latitudes vertical gradient is weak and more variable --> the vertical gradient in mid-latitude N2O is weak and more variable

  We have taken the suggestion from the reviewer.

- L146-147: "Towards tropopause altitudes, the N2O profiles of vortex and mid-latitudes merge and differentiation becomes difficult." Near the tropopause the vortex proper – and the transport barrier it represents – is no longer defined; the region in which chemical processing still takes place but confinement is weak (below ~350–380 K in SH, ~400–450 K in the NH, depending on the year) is often termed the "subvortex" [see Santee et al., JGR 2011, and numerous references therein]. So it is not appropriate to refer to "vortex profiles" in this region.

  We have rewritten this sentence for more clarity, that the vortex barrier vanishes towards the tropopause and no vortex profile can be defined.

  "[…] Towards tropopause altitudes, the transport barrier of the polar vortex disappears, and a classification is not possible. […]"

- L149: at best --> ideally

  Done.

- L152-153: "Stratospheric transport and mixing is related to the isentropic surfaces whereas mixing at the extratropical tropopause affects the lowest 25 K relative to the local tropopause." This sentence needs work.
  - The wording "is related to" is not clear. I assume that the first half of this sentence is referring to the fact that adiabatic flow in the stratosphere largely occurs along isentropic surfaces, but this should be clarified.
  - is --> are; also, add a comma after "surfaces"
  - References are needed, especially for the point that mixing affects the lowest 25 K above the tropopause (see below).

  We have reworded the sentence into two sentence and added necessary information on stratospheric transport and mixing. References were added for mixing at the extratropical tropopause.

  "[…] The composition of the lowermost stratosphere is affected by the diabatic decent inside and outside the polar vortex and quasi-isentropic mixing with air from lower latitudes. In addition, mixing at the extratropical tropopause affects the lowest 20-25 K above the local tropopause (Hoor et al., 2004, 2005) […]"

- flights, which --> flights that; contact to --> contact with

  This sentence was rewritten for more clarity.

"[…] The vortex reference profile (see Fig. 2) was generated from all flights that are assumed to contain measurements within vortex air. […]"

- L155-157: "Data from these flights were pre-filtered by taking only the measurements polewards of 60°S equivalent latitude and 20 K above the local tropopause." There are several issues with this sentence.
    - The concept of equivalent latitude should be defined and a suitable reference for it provided (e.g., Butchart & Remsberg [JAS, 1986]).

      We have included the suggested reference. However, we will refrain from defining the equivalent latitude at this point in the text, given that it is a well-known concept in stratospheric research.

    - Presumably the EqL is being calculated based on PV from a meteorological reanalysis, but this information needs to be provided. The reanalysis being used in a study is typically identified in the "Data and Methods" section – here that section is entitled "The SouthTRAC Campaign", but the meteorological data is also an important component of this study and probably merits its own subsection. I have more to say on this point later.

      We support this statement. It is not well captured, where the meteorological data come from, although they are important for the further analysis in this manuscript. Therefore, we include a new subsection "Meteorological data" in the section "The SouthTRAC campaign" with the necessary information.

      "[…]
      2.1 Meteorological data

      HALO was equipped with a wide range of in situ and remote-sensing instruments. In addition to the scientific instruments installed for the measurement campaign, the Basic Halo Measurements and Sensor System (BAHAMAS) is part of HALO. BAHAMAS is installed permanently and provides meteorological and aircraft parameters along the flight trajectory (DLR, 2020). The local tropopause information along the flight tracks of HALO was created using the Chemical Lagrangian Model of the Stratosphere (CLaMS) (e.g., Grooß et al., 2014). The underlying meteorological data are taken from ECMWF ERA-5 data (Hersbach et al., 2020). In this work, the potential vorticity (PV) based dynamic tropopause is used (e.g., Gettelman et al.,2011), taking the 2 PVU (potential vorticity unit) for the dynamical tropopause. Since the PV tropopause is not physically meaningful in the tropics, potential temperature level of 380 K was taken as tropopause if the 2 PVU level lies above. […]"

    - Similar to the above point, it needs to be made clear how the local tropopause is being determined. The Fig. 2 and 8 captions mention the "WMO tropopause", but more detailed information should be provided in the main text. In addition, it seems that the results may be highly sensitive to the exact definition of the tropopause used, and some discussion of the associated uncertainty in the results would be appropriate.

Information about the local tropopause can be taken from the newly introduced subsection 2.1 (see point before).

- The Antarctic vortex frequently extends to EqLs lower than 60°S EqL. Have the authors made sure that imposing the 60°S EqL cutoff has not eliminated vortex profiles in 2019?

  In fact, filtering to 60°S equivalent latitude can remove some measurements that may be counted as vortex. This pre-filtering is necessary for the iterative filtering procedure to find the lower envelope. The loss of these few data points should not affect the final profile. We checked graphically if the cut at 60° leads to a substantial loss of data points at the lower envelope. This was not the case.

- The previous paragraph states that mixing affects the region within 25 K of the tropopause, so it is not clear why the cutoff here was chosen to be 20 K.

  The previous paragraph has been changed with updated information, that mixing affects the region within 20 -25 K above the local tropopause (e.g. Hoor et al., 2004, 2005).
  In the Southern Hemisphere, the extratropical transition layer seems to be shallower (Hegglin et al., 2009). We therefore chose the lower value of 20 K for the cutoff.

- L157-158: This sentence ("The lowest levels ...") is repetitive, unnecessary, and out of place – it should be combined with the similar sentence in L152-153.

  This sentence has been removed and the necessary information and reference has been provided earlier in the text as suggested.

- Fig. 2 caption: to the local --> from the local; criterion on --> criterion of; the the --> the

  Done; Done; Done.

- L158-164: One general comment is that the creation of reference profiles is a key point on which much of the following analysis rests, and its description should not be relegated to the separate Supporting Information, which many readers will not make the effort to obtain. I would prefer to see it in the main text, but it should at least be moved to an Appendix included at the end of the main paper file (unless ACP no longer allows such Appendices). Another general comment is that the main text, SI, and figure captions together fail to clearly describe the method, as specified in more detail in the points below.
  - L160: At this point, the reader has no idea what is meant by the term "vortex profile function." Also, Werner (2006) is a PhD thesis for which no download information is given, and thus it is not a suitable reference.

We have changed the reference to a more suitable one:
Werner et al., 2010.

- L161: Elsewhere the convention “60°S EqL” is used, so for consistency “-40° and -60° EqL” here should be “40° and 60°S EqL”.

  We changed it accordingly.

- S1, L2-9: The first two paragraphs of the SI are fully redundant with the discussion in the main text. Apart from this repeated material, the description of the procedure in the SI is only one paragraph long, so again I would argue that it would be better to edit, merge, and rearrange the discussions in S1 L10-28 and L158-164 to produce a single compact paragraph in the body of the paper.

  We removed the first two paragraphs, which are indeed repeated material. The description of the filter procedure was moved to the main manuscript. We added an appendix A describing the filter procedure:

  “[…]
  Appendix A. Filter procedure for vortex and mid-latitude profiles

  A filter procedure was used to derive the lower envelope for the vortex profile and the upper envelope for the mid-latitude profile. Figure S1 in the supporting information displays the procedure for the task using either ΔΘ or Θ as the vertical coordinate. The process is initialized by binning the N_2O measurements into intervals of e.g., ΔΘ. The bin size must be adjusted to the number of measurements available for the vortex and mid-latitude profile to make the filter procedure work properly. For every bin, the mean value, standard deviation, and relative standard deviation are calculated. This is necessary as the condition for the filter needs a binned profile to begin with. While the maximum relative standard deviation is larger than the preset outlier limit, the measurements that are not flagged as outliers are binned in intervals of ΔΘ (this is done twice in the first iteration step, since the binned profile is already needed for the initialization and no outliers are set for the beginning of the filtering process). Every bin is checked whether the relative standard deviation is larger than the outlier limit. In this case, all measurements of N_2O which are higher (or lower, if the upper envelope is requested) than the mean of the respective bin are flagged as outliers and removed from further iterations. The iteration process stops when the maximum relative standard deviation is below the preset outlier limit. For the vortex profile, bin size was set to 5 Kelvin. The variability of N_2O on a constant Θ surface inside the vortex is about 6 ppb (Greenblatt et al., 2002). For the range of N_2O mixing ratios in this work, this corresponds roughly to about 3 % and was thus set as the outlier limit. Four iterations were done to get the lower envelope (grey samples in Figure S 2 a) and b) in the supporting information). For the mid-latitude profile, the bin size was set to 2 Kelvin. Strahan et al. (1999) showed that the variability of N_2O in the Southern

Hemisphere lower stratosphere of the mid-latitudes is approximately between 5 and 15 % (see plate 6 therein). Therefore, a value of about 10 % is set for the outlier limit, which leads to two iteration steps for the remaining measurements. For the profiles only those measurements are used which are not marked as outlier […]"

In the main text, we replaced the sentence "A more detailed description of the creation of reference profiles can be found in the supporting information." With the following information from the supporting information:

"[…] As an intermediate step to the final profiles the measurements of the lower and upper envelope are binned in 5 Kelvin intervals of Θ or ΔΘ (see Figure S 2).
Mean values of the binned profiles are then used to generate a polynomial fit function for the vortex profile and the mid-latitude profile (Figure S 3). […]"

- Fig. S1: I did not find the flowchart to be particularly helpful, so it could remain in the SI.

  Flowchart remains unchanged.

- Fig. S2-S7: It would be much easier on the reader if all of the vortex figures (S2, S4, S6) were combined into one 3-panel figure, and the same for the mid-latitude figures (S3, S5, S7). In fact, it would probably work to combine them all into one 2-row, 3-column figure. o S1, L13: is calculated --> are calculated

  The figures are rearranged. We combined Figure S2 to S5 into one 2-row, 2-column figure. Therefore, it is more compact, and the reader gets a good overview. Figure S6 and S7 are combined to one 2-column figure.
  We have refrained from a 2-row, 3-columns figure, as the individual figures become much smaller, and the reader may have difficulties looking at them.

- S1, L14 and L21: pre-setted --> preset

  Done.

- S1, L14: measurements, which --> measurements that

  Done.

- S1, L17: Is this the case --> In this case

  Done.

- S1, L21: shows --> show

  Done.

- S1, L23: latiudes --> latitude

  Done.

- S1, L22-23: More discussion is needed on the 3% and 10% "outlier limits" for vortex and mid-latitude profiles, respectively. How were these preset outlier limits and Δθ bin sizes determined? What factors drove the differences between the values of these quantities for the vortex and mid-latitude profiles? How sensitive are the results to these choices?

  Greenblatt et al. (2002) shows that the variability of N_2O inside the vortex is about 6 ppb. By looking at the lowest N_20 mixing ratios within the potential temperature range (see Figure S2), 6 ppb refers to roughly up to 3% variability. For the mid-latitudes, Strahan et al. (1999) investigated a variability in the lowermost stratosphere ranging from around 5% to 15%.
  We chose the mean of 10% as representative of the variability of the mid-latitudes. The size of Δ θ was chosen accordingly the amount of measurements to make the filter procedure work. We have added this information at appropriate places in the text.

  Regarding Δθ:
  "[…] The bin sizes must be adjusted to the number of measurements available for the vortex and mid-latitude profile to make the filter procedure work properly. […]"

  Regarding the 3% outlier limit:
  "[…] For the vortex profile, bin size was set to 5 Kelvin. The variability of N_2O on a constant Θ surface inside the vortex is about 6 ppb (Greenblatt et al., 2002). For the range of N_2O mixing ratios in this work, this corresponds roughly to about 3 % and was thus set as the outlier limit. […]"

  Regarding the 10% outlier limit:
  "[…] For the mid-latitude profile, the bin size was set to 2 Kelvin. Strahan et al. (1999) shows that the variability of N_2O in the Southern Hemisphere lower stratosphere of the middle latitudes is approximately between 5 and 15% (see plate 6 therein). Therefore, a value of about 10 % is set for the outlier limit, which leads to two iteration steps for the remaining measurements. […]"

- Fig. S2: What causes the "staircase" pattern between the points discarded in iterations 1 and 2? Also, left over --> leftover, but "remaining" would probably be a better word here (also in the Fig. S3 caption)

  The "staircase" pattern is a result of the filtering procedure and is always as large as the bin size. For every bin, all measurements which are higher (for vortex profile) or lower (for mid-latitude profile) than the N_20 mean value of the bin are flagged, in case the relative standard deviation of the bin is larger than the outlier limit.

- L167: decent --> descent

  Done.

- L168-169: Greenblatt et al. [2002b] quantifies descent inside the Arctic vortex and thus has only limited applicability for SouthTRAC, since the characteristics and seasonal evolution of descent are somewhat different in the two hemispheres, as discussed by Manney et al. [JAS, 1994], which would be a more appropriate reference. Manney et al. note that parcels in the SH lower stratosphere generally cease to descend in mid-October, so descent might still be ongoing in September, contrary to the statement made here. Also, a further --> further.

  The referee is correct. Greenblatt et al., 2002b is not a suitable reference for the diabatic descending of the Southern Hemisphere since it differs in this aspect from the Northern Hemisphere. We therefore rearranged this paragraph and used the suggested reference instead.

  "[…] In the lower stratosphere of the Southern Hemisphere, the descending stops around mid-October (Manney et al., 1994). However, N_2O data of the SouthTRAC flights did not reveal strong diabatic descent during the time of the campaign (below θ= 400 K). Therefore, only one reference vortex profile was generated for the campaign. […]"

- L173-174: I am confused about the "prescribed cutoff value" for vortex profiles and "associated variability" for mid-latitude profiles mentioned here and specified (20 ppb and 15 ppb, respectively) in the Fig. 2 caption. Where did these values come from? Why is one characterized as a "prescribed cutoff" and the other as an "associated variability"? Do these values have anything to do with the outlier limits discussed above? In addition, it seems that all points falling "below" (i.e., to the left of) the grey "cutoff" curve are deemed to be vortex points, even when they are "above" (i.e., to the right of) the black N2Ovor curve, but that is not what is said in "if the mixing ratio is below the respective N2Ovor with a prescribed cutoff value, then it is assigned to the vortex" (and similarly for the mid-latitude case).

  The prescribed cutoff was adopted from Greenblatt et al. (2002). Ivanova et al. (2008) also used the value 20 ppb by Greenblatt et al. (2002) for measurements in the Antarctic, thus we assumed a usage of the value for both Arctic and Antarctic. We derived the associated variability from the 10% variability of the mid-latitude measurements, leading to a $\pm$15 ppb variability.
  In the text, we include theses information into the last paragraph of section 3.1:

  "[…] For the prescribed cutoff value, the value of 20 ppb proposed by Greenblatt et al. (2002) was used. The associated variability of the mid-latitude profile was set to 15 ppb, as the variability N_2O in the mid-latitudes is roughly 10% (see supporting information and reference therein). […]"

  For more clarity regarding the classification procedure, we changed some wording in the following same paragraph:

"[…] if the mixing ratio is below the respective N_2O_{vor} added by a prescribed cutoff value, then it is assigned to the vortex. Otherwise, if the mixing ratio is above the respective N_2O_{mid} minus an associated variability, then it is assigned to mid-latitudes. Mixing ratios above the respective N_2O_{vor} added by a prescribed cutoff value and below the respective N_2O_{mid} minus an associated variability are assigned to the boundary region. For the mixing ratios where N_2O_{vor} added by a prescribed cutoff value and N_2O_{mid} minus an associated variability overlap, these measurements cannot be assigned to one region. […]"

- L176-177: It is stated that measurements in the overlap region cannot be "fully assigned to one region". So what was done with them? Were they included in the analysis or discarded? This question is answered later in section 4.3, but the reader should not be left in suspense here. Also, delete the comma after "ratios", and "cannot" is one word, not two.

  We added a sentence to make clear, that these measurements were not discarded. Further, we delete the comma after "ratios" and changed "can not" to "cannot".

  "[…] For this reason, these measurements are assigned to both the vortex and the mid-latitudes in later analysis. […]"

- L181: has been --> was

  Done.

- L183: "timed" sounds intentional, whereas I believe that the campaign fortuitously took place shortly after the SSW. This sentence is also grammatically awkward. I suggest instead: "... November; thus they occurred shortly after the minor SSW event and captured the late winter evolution ..."

  We agree that "timed" is misleading and changed the sentence similar to the suggestion.

  "[…] The SouthTRAC campaign flights took place from early September to early October and in the first half of November; thus they took place shortly after the minor SSW event and captured the late winter evolution of the Antarctic polar vortex. […]"

- L187: Other flights besides the 11 September flight are omitted from Fig. 3, since the number shown does not add up to the total given in section 2.

  That is correct. Figure 3 only shows the local flights for which the classification was possible. We added this in the text.

  "[…] Figure 3 displays an overview of air mass classification of the local flights of the SouthTRAC campaign (classification in θ-coordinates) […]"

- L189: It is not quite true that “flights sampled mostly inside the vortex or vortex boundary region” during Phase I – in fact, Fig. 3 shows that few or no such measurements were taken on nearly half of those flights. Also, extensive --> extensively; add a comma after “phase”

  The meaning of this sentence is that there were flights in the first phase, which measured predominantly vortex and vortex boundary region. To show this better, we have divided the sentence and added this information:

  “[…] Vortex and boundary region were sampled in both phases of the campaign. The first phase contains some flights that have predominantly sampled vortex or vortex boundary region (e.g., flight ST15 on 29 September or flight ST16 on 30 September). […]”

- L191: more than half of the ... air 54% --> more than half (54%) of the ... air

  Done.

Section 4

- L194-197: These sentences about EESC seem out of place here. Perhaps they would fit better at the beginning of subsection 4.2, or in the Introduction. Also, matter --> manner

  We agree with the Referee and moved the sentences about EESC to the introduction. In addition, we reworded the first sentence at the advice of the Referee #1.

  “[…] Equivalent effective stratospheric chlorine (EESC) is a simple metric that sums the effect of ozone depleting substances (ODS) as an equivalent amount of inorganic chlorine in the stratosphere (Newman et al., 2007; Daniel et al., 1995). Changes to the EESC are mainly due to Cl_y, as Br_y makes up a smaller fraction (Strahan et al., 2014) […]”

- L197-199: These two sentences are redundant with the paragraph at the start of section 4.1. It would be better to merge / edit to avoid such repetition from one paragraph to the next.

  As this information are redundant with the first paragraph of section 4.1, we removed these sentences and part of one sentence was included in the first paragraph of section 4.1.

  “[…] This could lead to a rather coarse resolution where fine structures like filaments and small-scale dynamical perturbations are sometimes not well resolved. […]”

- L205-206: Measuring CFC-12 on both the ECD and MS channel of the instrument allows to up-sample the measurements of the organic source gases by using the higher resolved --> Measuring CFC-12 in both the ECD and MS channels of the instrument allows the measurements of the organic source gases to be up-sampled by using the better-resolved

Done.

- L206: CFC-12 on --> CFC-12 in; throughout the manuscript (including in figure captions), “measurements on” should be changed to “measurements in”

  We changed “CFC-12 on” to “CFC-12 in”; There is no “measurements on” in the text to change.

- L207: but also a better precision than on the MS channel --> but they also have better precision than data from the MS channel

  Done.

- L209: add a comma after “ratios”

  Done.

- L209: It might be good to add “linear or polynomial” in front of “fit function”.

  Done.

- S2, L30-31: For up-samling the GhOST-MS measurements, pre-required are good correlations between CFC-12 and the other --> Up-sampling the GhOST-MS measurements requires good correlations between CFC-12 and the other

  Done.

- Fig. S8: It should be made clear in the caption that all of the data shown are from the GhOST-MS channel. Also, the small font makes the axis labels on these panels very hard to read.

  We filled in the information, that the data are from the GhOST-MS channel.
  In addition, we have created two figures from this one figure so that the plots are easier to look at.

- Fig. 5: In red, original data, whereas is black, measurements were up-sampled using CFC-12 measurements of the ECD channel --> Original data shown in red, measurements up-sampled using CFC-12 from the ECD channel in black

  Done.

- L216-217: It is not that “the original data were not well captured”; rather, the original lower-resolution data did not capture well the abrupt transitions between regimes.

  We rewritten this sentence for more clarity.

"[…] Especially with the sharp gradients, e.g. at 04:10 UTC and at 05:50 UTC in Fig 5, the original lower-resolution data did not capture well the transitions between the regimes, compared to the up-sampled […]"

- L224-225: I find this wording unclear. It would be better to rewrite as: "Organic chlorine (CCly) ... up-sampled GhOST-MS measurements. Thus, Cly can be calculated from Eq. 1 if the mixing ratios of the major chlorine-containing substances at the stratospheric entry point (Cltotal) are known. Air enters the stratosphere predominantly ...".

  We rewritten this part of the section accordingly.

- L227: can not --> cannot

  Done.

- L228: times in the stratosphere since they entered the stratosphere --> times since they entered the stratosphere

  Done.

- L236: previous --> previously

  Done.

- L238: ratio, which --> ratio that; degradation and thus --> degradation, which thus

  Done; Done.

- L239: It would be appropriate to add "estimated" in front of "entry mixing ratios"

  We have inserted "estimated" as suggested.

- L241: For the case no --> For the case where no; also, to make the distinction between the so-called "semi-direct" and "indirect" methods more clear, it would help to add "indirectly" between "calculated" and "based on".

  Done; included "indirectly" between "calculated" and "based on".

- L244: add a comma after "trends"

  Done.

- L246: where --> when

  Done.

- L247: delete the commas after both "showed" and "tracers"

Done.

- L254: respective entry --> respective estimated entry

  Done.

- L256-257: rations with --> ratios by

  Done.

- L259: It is difficult for the reader to keep track of exactly what is meant by the "direct", "semi-direct", and "indirect" methods. To help clarify this sentence, it would be good to add "based on previous balloon observations transferred to 2019" after "indirectly determined correlations" (assuming that I am interpreting the approaches correctly).

  We have attached the proposed additional information to this sentence.

- L259: indirectly determined values are not only based on observations which have been performed about 10 years earlier but also are from the --> indirectly determined values are based on observations that were not only performed about 10 years earlier but that were also from the

  Done.

- L261: This wording is unclear. Replace "They" with "The balloon-based correlations".

  Done.

- Fig. 6: I assume that the SouthTRAC (black) points in this figure show the up-sampled (not raw) GhOST-MS measurements for CFC-11, etc., but this should be stated explicitly. Also, I may not be interpreting this figure correctly. Do the red symbols represent the correlations between balloon CFC-11 (for example) and balloon CFC-12 data, or between balloon CFC-11 (for example) and GhOST-ECD CFC-12 data? Please clarify. In addition, the term "retrended" is used only in the figure caption and legend, not in the main text. Although it is somewhat ambiguous, it is fine to use this term as long as it is defined in the body of the paper.

  Figure 6 displays the raw GhOST-MS measurements of CFC-12, CFC-11, CH_3Cl, and HCFC-142b in black. The red symbols represent the balloon observations in 2009 and 2011 transferred to the time of the SouthTRAC campaign in 2019, thus not using GhOST-ECD CFC-12 data. We want to show that despite the different hemispheres and the time interval of 10 year, the correlations of the long-lived chlorinated substance are comparable. These correlations can be used to determine organic chlorine (CCly) from CFC-12 alone. For more clarification, we extended the description in the caption and avoid the term "re-trended":

"[…] Correlation between CFC-12 and CFC-11, CFC-12 and CH_3Cl, and between CFC-12 and HCFC-142b. In black the raw measurements by GhOST-MS, in red the balloon observations scaled to the time of the SouthTRAC campaign using mean arrival time. […]"

We also extended the sentence in the text, where figure 6 is mentioned:

"[…] Fig. 6 displays scaled correlations from the balloon observations (red) and correlations form the SouthTRAC data (black) of three long-lived substances against CFC-12. […]"

- L262-265: I am not following the logic here. I understand that a subset of the components of CCly "retrended" from earlier balloon data match well correlations with CFC-12 measured by SouthTRAC (Fig. 6). But why does that necessarily mean that CCly based on correlations with CFC-12 can be used as a good proxy for Cly? I feel that a step is missing. And why is it relevant to mention here again that Cltotal can be derived to calculate Cly – that information is not being used for Eq. 2 and the indirect method, is it? In general, I feel that the relationship between Eq. 1 (at the heart of the semi-direct approach) and Eq. 2 (the basis of the indirect method) is not clearly explained. Please clarify this discussion.

  We agree that there is a missing step in between the explanation on how to get to equation 2. Thus, we expanded the description, why we need Cl_{total} again for the indirect Cl_y and how we get Cl_{total}. The following changes are done in the manuscript:

  "[…] The balloon-based correlations correspond well to the correlations measured during the SouthTRAC campaign. Thus, the balloon-based correlations can be used to determine CCl_y from CFC-12 alone. As already mentioned earlier, Cl_{total} is also needed for the calculation of Cl_y. For this, the mean age values derived for the balloon measurements are used and Cl_{total} is calculated for the conditions during SouthTRAC hereafter. Cl_y is then derived as the difference between Cl_{total} and CCl_y. A correlation function for the conditions during Antarctic late winter 2019 has then been derived for the indirect calculation of Cl_y as a function of CFC-12 mixing ratios (Eq. 2). […]"

- L266-267: Since the previous sentences have been discussing balloon-based correlation functions, this sentence about the GhOST-ECD CFC-12 data seems out of place and confusing to me – maybe it belongs at the end of the following paragraph rather than here, or perhaps I have misunderstood the role of those data in the foregoing discussion, as noted above.

  With the before mentioned method, we generated the correlation function for Cl_y with the reference substance CFC-12 based on the balloon data. We now need CFC-12 measurements during the SouthTRAC campaign to calculate Cl_y. This was done using CFC-12 from the GhOST-ECD channel. We slightly changed the sentence to make clear, that CFC-12 of GhOST-ECD was used for indirect Cl_y.

  "[…] In the following, CFC-12 from the GhOST-ECD channel is used for the indirect determination of inorganic chlorine. […]"

- Table 2: To enhance clarity, it would be good to add “indirectly” in front of “derive”.

  Done.

- Fig. 7: Are the indirect results shown here from the “retrended” balloon data or from SouthTRAC? I assume the former but this should be stated explicitly. And why not show comparisons for both data sets, since you also provide SouthTRAC coefficients in Table 2? Also, in the caption: Indirectly and directly determined ... (green and back) and --> Indirectly (green) and directly (black) determined ... and

  This figure shows results of the indirect method from the “re-trended” balloon data. We clarify this in the caption of the figure.

  “[…] Figure 7. Indirectly (green) determined Cl_y based on balloon observations in 2009 and 2011 and semi-directly (black) determined Cl_y as a function of age of air. In red, the absolute difference between these methods. […]”

- L275: delete the comma after “Hemisphere”

  Done.

- L276: Cly, where --> Cly in cases where

  Done.

- L277-278: “Since it was possible during SouthTRAC to measure the organic source gases, the Cly from the direct measurements was used for further evaluation.” In fact, while reading this section I wondered why the authors bothered to pursue the indirect approach when they actually have direct measurements of CCly. The discussion of the indirect Cly calculation is particularly confusing, and it is not at all clear at this point what value it brings. So it would be helpful to add a pointer to section 4.4, where the indirect method is needed for the comparisons with Cly in the Arctic, to better justify the inclusion of this discussion here. In addition, I think that, rather than “Cly from the direct measurements”, it would be more appropriate to say “Cly determined semi-directly from the measurements”.

  You are right. For SouthTRAC the indirect method is not needed.
  However, the measurements during SouthTRAC offer to compare the semi-direct method and the indirect method to show that the indirect method leads to comparable results, which is crucial later. We added a pointer to section 4.4 to justify the comparison.
  We also changed “Cl_y from the direct measurements” to “Cl_y determined semi-directly from the measurements”.

  “[…] However, the good comparability of the two methods offers the possibility to compare Cl_y from different measurement campaigns, which differ regarding the number of measured chlorinated substances (see section 4.4). […]”

- L279: Another thing that is not clear to me is why the fit coefficients for N_2O are given in Table 2 if they are not being used here at all. On the other hand, several previous studies have used N_2O to derive Cly, so I think it might be useful to expand the discussion of the N_2O correlations. An obvious question that arises is: How well does the Cly derived from fits with N2O agree with that based on the correlation with CFC-12, for both the balloon data and SouthTRAC? A figure similar to Fig. 7 could be added for N2O.

  We included the fit coefficients for N_2O as N_2O was used to calculate Cl_y in several publications before. Thus, if one wants to calculate Cl_y for this time using N_2O measurements, the given correlation coefficients can be used. As they are already mentioned in the paragraph before (before equation 2) we expand the discussion of the N_2O correlation at this point by the following (including a similar figure to Fig 7, as suggested by the referee):

  "[…] N_2O shows a compact correlation to long-lived chlorinated substances and has been used in many publications for the determination of Cl_y (e.g., Schauffler et al., 2003; Strahan et al., 2014; Strahan and Douglass, 2018). Using CFC-12 from the GhOST-ECD channel and N_2O from the UMAQS instrument, we obtain comparable values for Cl_y (see figure S 7 in the supporting information). In the following, CFC-12 from the GhOST-ECD channel is used for the indirect determination of inorganic chlorine. […]"

  We have refrained from a figure for the coefficients from the SouthTRAC data, since in corresponding coefficients in the table are derived from the same semi-directly determined Cl_y values.

- L284: only measurements were taken, which are polewards of 40° equivalent latitude --> only measurements polewards of 40° equivalent latitude are used

  Done.

- L287: an air mass --> air mass

  Done.

- L289-290: It seems to me that it might be better to exclude the measurements in the overlap region from the analysis rather than "double count" them. How many measurements fall into this category, and how would omitting them change the results?

  When using the classification in Θ-coordinates, there are 263 Cly measurements in the overlap region for the SouthTRAC campaign. The number of measurements is even larger with 592 measurements using ΔΘ as the vertical coordinate. Omitting these measurements would affect the variability of the profiles in the range of potential temperature, where the vortex cutoff and variability of the mid-latitude profile overlaps. Since, for example, the cutoff of 20 ppb is no longer considered, a narrow

range of mixing ratios is considered for the vortex profile, and it shifts toward higher Cly values. The measurements in the overlap area can, however, originate from the vortex or the mid-latitudes. Thus, we would like to include them for both profiles.

- Fig. 8: I have a number of comments / questions about this figure.
  - It is stated that data are averaged over -40° to -90° – is the filtering of measurements obtained equatorward of 60° equivalent latitude (mentioned in section 3.1) only applied in calculating the vortex reference profile? Please clarify (here and in section 3.1).

    Indeed, the filtering of measurements poleward of 60° equiv. latitude is only applied to generate the vortex reference profile as well as the filtering between 60°S and 40° S equivalent latitude for the mid-latitude profile. The consideration of measurements south of 40° S equiv. latitude serves on the one hand to ensure that the classification was only made in this range and to filter out the data of the transfer flights to the Northern Hemisphere.
    The information given in the text, both in section 3.1 and in section 4.3, should already make clear that only for the vortex profile, there is a limited number of flights and the cut at 60°S, whereas in the analysis in 4.3 all flights are used and from this all measurements south of 40°S are taken.

  - The colors denoting the different regions have been changed. Previous figures used a consistent set of colors for these classifications, and it would be easier for readers if that same color scheme was used in all figures for which those classifications are relevant.

    The colors have been changed to match the colors of the previous figures in terms of classifications.

  - The lack of tick marks on the top and right-hand axes is annoying and makes it difficult to judge the values given in the text. In addition, the tick marks (especially the minor ones) that are present on the bottom and left-hand axes are too small to be easily seen.

    Tick marks were added to the top and right-hand axes. In addition, ticks were made larger, both major and minor ticks.

  - Why does Cl_{total} vary with altitude?

    The mean value of Cl_{total} between around 280 and 400 K potential temperature in figure 8 is 3074 ppt with a maximum of 3097 ppt and a minimum of 3054 ppt. Thus, the range of Cl_{total} is about 43 ppt.
    Furthermore, the relative standard deviation of all bins is around 0.41%.
    The standard deviation of each individual bin is of the same order of magnitude. The variability of Cl_{total} is therefore very small. Nevertheless, a very small increase in Cl_{total} with altitude in the stratosphere can be expected, considering the temporal delay (the age of air) with which Cl_{total} propagates into the stratosphere (e.g., Engel et al., 2002).

- Elsewhere total chlorine was written "Cl_{total}", and that should be the case here too.

  Cl_{tot} was changed to Cl_{total} in the figure legend.

- For consistency with the text, "-90° to -40°" should be "40°–90°S".

  The figure title was changed to:

  "[…] Chlorine partitioning between 40 and 90°S eq. lat [°] […]"

- "mean averaged" is redundant

  We deleted "mean".

- L291: the measurements --> the SouthTRAC measurements

  Done.

- L292: A reference is needed for the AGAGE results.

  For not misleading at this point, we have rewritten this sentence, as this is not at statement based on AGAGE result but based on the chlorine input values which are based on the AGAGE time trends.

  "[…] The SouthTRAC measurements of the long-lived chlorinated substances are consistent with the total chlorine of these substances, based on time trends from the AGAGE Network. […]"

- L294-295: add commas after "330 and 390 K" and "390 and 400 K"

  Done.

- L299-300: "at this altitude" – which altitude? Where Cly is maximum? Please clarify.

  We changed "at this altitude" to "at highest measured potential temperatures"

- L301: Accompanying the minor SSW --> As a consequence of the minor SSW

  Done.

- L301-302: It is not clear that the 16.4 million km2 value quoted here refers to the maximum daily ozone hole area. Moreover, it is not true that the 2019 hole was "the smallest since its discovery" – other holes in the mid-1980s had smaller maximum daily area values. In any case, rather than quoting this value from the Ozone Watch web site, a better approach would be to reference the Wargan et al. [JGR, 2020] paper (already cited elsewhere in this manuscript); their Fig. 1d puts the area of the

2019 hole into climatological perspective. Perhaps more importantly, it is not clear what the point of these sentences is. Are the authors trying to imply that Cly levels in the 2019 vortex played a role in the weak ozone hole that year? Although Cly abundances inside the vortex do vary from year to year as discussed previously by Strahan et al. [JGR, 2014], variations in lower stratospheric temperatures are the primary driver of variations in the strength of polar ozone depletion.

We took Wargan et al. 2020 as the reference for the size of the Antarctic ozone hole in 2019. In addition, we have rewritten the sentences. We do not want to create a link between the amount of Cl_y and the size of the ozone hole. Instead, we wanted to note that the minor SSW event led to an early chlorine deactivation.

"[…] Thus, in late winter and early spring at this altitude about half of the recorded chlorine is found in inorganic form. Despite this amount of inorganic chlorine in the lower stratosphere, the total polar ozone column was higher than usual in September 2019. As a result of the minor SSW event, chlorine deactivation began earlier in 2019 and the ozone hole was about 10 x 10^6 km^2 in size, thus only 20 % of that in 2018 mid-September (Wargan et al., 2020). […]"

- L303: The title of this subsection suggests that the SouthTRAC and PGS comparison focuses on the polar region, but Fig. 10 and associated discussion includes mid-latitudes as well. It may be true that comparison of Cly in the Antarctic and Arctic polar vortices has not been done previously, as the authors assert, but it is not the case that no such comparisons have been performed in the midlatitudes. In fact, total column Cly (or, rather, HCl+ClONO2) in the NH and SH mid-latitudes (Jungfraujoch and Lauder, respectively) and the trends therein are compared in the Ozone Assessment (e.g., Fig. 1-13 of WMO 2018). It would be good to place their findings into the context of these (and possibly other) midlatitude results.

  We included mid-latitudes comparisons from the Ozone Assessment 2018 to section 4.4. of our manuscript.

  "[…] In addition, Mahieu et al. (2014) reported long-term total column data for HCl and $ClONO_2$ (representing Cl_y) in the stratosphere, at Jungfraujoch (46.5°N) and at Lauder(45°S), though the end of 2016. A negative trend of Cl_y is observed at both stations but with a non-significant trend for the Jungfraujoch data over the last decade and a slightly larger negative trend from the Lauder data. Furthermore, lower-stratosphere HCl from the Global Ozone Chemistry And Related trace gas Data records for the Stratosphere (GOZCARDS) shows larger decreases at southern latitudes and increases at northern mid-latitudes (Froidevaux et al., 2015}. Thus, higher values of Cl_y in the mid-latitudes during PGS seems to be plausible. […]"

- L310: A separation --> The separation

  Done.

- L311: is based on the above mentioned method based on --> is based on the above-mentioned method using

Done.

- S3, L34: the vortex and the mid-latitude profile during PGS is needed --> the vortex and the mid-latitude profiles during PGS are needed

  Done.

- S3, L35: Phase --> phase

  Done.

- L312-314: As shown in section 4.2, the indirect method ... possible, proves to be comparable --> In section 4.2, the indirect method ... possible, was shown to provide results comparable to those obtained by the semi-direct method

  The sentence was rewritten as suggested.

- L317: I assume that "2019" is a typo and that the same balloon data from 2009 as used for SouthTRAC were again used for PGS?

  Yes indeed, this is a typo and we changed it to 2009.

- S4, L37: Correlations function --> Correlation function

- L319 and L326: Cl_{tot} --> Cl_{total}

  Done.

- L319-320: This sentence ("The vertical coordinate of the classification was selected according to the displayed vertical coordinate.") is confusing and, if I understand it correctly, completely unnecessary, as the very next sentence makes clear that the two panels display the results as a function of θ and Δθ. Also, I am curious why both vertical coordinates are shown here but not in Fig. 8. Then in Fig. 10 only the tropopause-relevant coordinate is shown, with the argument that it allows for better comparison of Cly in the two hemispheres. It seems to me that it might have made more sense to show both views in Fig. 8, and then use only the tropopause-relevant coordinate in Figs. 9 and 10 for the reasons stated.

  We agree that this sentence is unnecessary and removed it.
  Regarding then use of Δθ *as a vertical coordinate: Only when comparing the two campaigns* Δθ is used as *by using* Δθ *instead of* θ *the comparison of the campaigns leads to slightly different results especially in the lowest stratosphere. We would therefore leave the vertical coordinates as they are for the figures 8 to 10.*

- L322: as it was done --> as was done

  Done.

- L323-325: Several issues arise in this sentence.
  - As I noted above, information about the meteorological data on which this study depends needs to be provided much earlier in the manuscript, ideally in section 2.

    As mentioned earlier, a new subsection "Meteorological data" was included in the section "The SouthTRAC campaign" with the information about the meteorological data for this study.

  - I was surprised to discover that the analysis is based on NCEP reanalyses. Insufficient information is provided here to identify exactly which NCEP reanalysis is being used (NCEP-NCAR R1, NCEP-DOE R2, CFSR, or CFSv2), but that needs to be specified and the corresponding reference (i.e., published journal article) cited.
  - I have concerns if either NCEP R1 or R2 have been used for this analysis. Although both are still in widespread use, these reanalyses have been shown in several studies, including some recent papers stemming from the SPARC Reanalysis Intercomparison Project (S-RIP), to be unsuitable for most stratospheric studies (as noted in the S-RIP overview paper by Fujiwara et al. [ACP, 2017]).
  - What exactly is meant by "climatological" here? That is, how many years have been considered in the averages? Were the climatological means also calculated over the days covered by the respective campaigns, or are they monthly averages, or ...? Are the climatological tropopauses being calculated over 40° to 90° (this latitude range is stated in the caption, but it is not clear exactly what it is referring to)?

    Regarding the three points, we excluded the comparison with climatological mean tropopause values for the SouthTRAC and PGS campaign.
    The NCEP/NCAR Reanalysis 1 (R1) was used with monthly means. Due to the change to PV-based tropopause and the revision of the manuscript, we no longer performed a comparison with climatological tropopause values. This information is no longer included in the revised script.

- L326-327: "... abundance of total chlorine (Cltot) was lower in the stratosphere from the time of PGS (2015/2016) to the time of SouthTRAC (2019)" – the wording of this sentence is unclear. It should be rewritten to state that the abundance of total chlorine in the stratosphere decreased between the two campaign periods. It is very difficult for the reader to precisely judge the magnitude of this decline from the figure. Is the difference in the estimated PGS and SouthTRAC values of Cl_{total} consistent with expectation given the known decreasing trend in stratospheric chlorine loading? This is a key point.

  We slightly changed the wording of this sentence to the following:

  "[…] Independent of the vertical coordinate, total chlorine (Cl_{total}) in the lower stratosphere decreased from the time of PGS (2015/2016) to the time of SouthTRAC (2019) […]"

We further expanded the discussion about the decrease of Cl_{total} between the two campaigns by including results from the WMO Report 2018:

"[…] The difference is on average about 60±9.6 ppt, thus roughly a rate of change of -16±2.6 ppt year-1. This rate is higher than the average rate of change of -12.7±0.9 ppt year-1given by Engel et al. (2018b) between 2012 and 2016, considering the long-lived chlorinated substances. […]"

- L329-330: the SouthTRAC profile increased stronger and values become more than 435 ppt larger than during PGS within the vortex at equal potential temperatures --> the SouthTRAC profile increased more steeply, reaching values more than 435 ppt larger than those during PGS at the same potential temperatures

  Done.

- L330: Differences become --> Differences are

  Done.

- L331-332: This sentence ("Inside the vortex ... during PGS.") is entirely redundant with the second sentence of this paragraph and should be deleted.

  We excluded this sentence.

- L332-333: Although close together between 20 and 25 K Δθ, the difference of Cly increased to 565 ppt at 65 K Δθ --> Although the two Cly profiles lie close together between 20 and 25 K Δθ, the differences between them increase to 565 ppt at 65 K Δθ

  Done.

- Fig. 9:
  - Again, please add tick marks on the top and right-hand axes.

    Done.

  - 40° to 90° --> 40° to 90° equivalent latitude

    Done.

  - Delete "and as a function of potential temperature difference to the local tropopause" in line 3 – this information is provided in the description of panel (b) below.

    Done.

- “mean averaged” is redundant

  Done.

- SouhTRAC --> SouthTRAC

  Done.

- L335: the latitude --> the geographic latitude

  Done.

- L337-338: and better allow for --> and allows for better

  Done.

- L339-340: It might be interesting to know how many points contribute to each latitude- altitude bin in both hemispheres. Is there a minimum threshold for the number of points in each bin? Perhaps bins with very disparate numbers of points contributing in the NH and the SH could be marked in some manner.

  The threshold for the number of measurements in each bin is five measurements. If this is not reached, the bin is not used for the evaluation. We add this information in the text as followed:

  “[…] Only bins which contain at least five data points were considered in this analysis. […]”

- L341: add a comma after “latitudes”

  Done.

- L344: Highest levels of Cly reach 386 ppt more Cly during PGS --> The highest values of Cly reached are 386 ppt greater during PGS

  Done.

- L345: vortex of each hemisphere is --> vortices of the two hemispheres are

  Done.

- L347: I do not believe that “sporadic”, which means “infrequent” or “intermittent”, is the right word here, especially as no time information is conveyed in this plot. Perhaps “weak” or “moderate” would work, if I have understood the point the authors wish to make.

  We changed “sporadic” to “weak”.

- L348: it is not clear what “it” is referring to here -- Cly?

  “It” is referred to the potential temperature range of up to 20 K above the local tropopause. As the sentence before begins with this range of potential temperature, we change “it” to “this range”.

- L349: for both --> in both; add a comma after “hemispheres”; there is no need to introduce the acronym for ExTL since it is not used again in the manuscript

  Done; Done; “(ExTL)” was removed from the sentence.

- Fig. 10: Please add tick marks on the top and right-hand axes as well as minor tick marks. Also, the color bar label should indicate that these are differences, not raw Cly values.

  We added the tick marks and changed the label of the color bar.

Section 5

- L352: Using an extended method according to Greenblatt --> Extending the method of Greenblatt

  Done.

- L353-355: It is stated that, compared to coarser-resolution PV, the method to define the vortex used here allows small structures such as filaments to be resolved. First, as noted earlier, modern meteorological reanalyses provide PV at fairly fine resolution. Second, no evidence is presented in this paper that any such filaments were actually resolved using their approach. So I am not convinced that a PV-based definition would not have been adequate.

  We agree that a comparison with the PV-based classification, which has not been done in this manuscript, is not appropriate. Although it is not clear whether vortex filaments are detected, the classification based on N2O measurement does indeed reveal small-scale structures, as can be seen in Figure 5.
  We replaced this statement with a more suitable one:

  “[…] The classification of air masses based on high-resolution in-situ measurements of N_2O offers the possibility to detect and account for even small structures and follows well the sharp gradient between the regimes. […]”

- L358-360: The authors are correct when they point out that the dynamical tropopause would be more appropriate for this kind of study than the thermal tropopause. Unfortunately, the use of the WMO tropopause raises questions about the value of this investigation. I do not really understand how the authors can say that “no dynamical PV tropopause data is yet available for the SouthTRAC campaign”. In fact, high-resolution PV fields are available from multiple reanalyses. There is abundant literature discussing which PV values are most appropriate for defining the

dynamical tropopause, depending on the hemisphere and isentropic surface, etc. So it is not clear to me why the authors could not have chosen representative PV values and performed their own interpolations to the in situ measurement locations to determine the local tropopause. But even if the authors are not set up for those calculations, they could still do more to reassure readers that use of the thermal tropopause does not substantially affect their conclusions. Keber et al. used the dynamical (2 PVU) tropopause, so that information is readily available for PGS. Some simple comparisons between the WMO and PV tropopauses for the period of the PGS campaign and examination of the impact the differences between them have for Figs. 9 and 10 would be informative.

We have consulted the working group that determines the local tropopause height in potential temperature along the flight trajectory for the HALO measurement campaigns PGS and SouthTRAC. We now have the PV-based dynamic tropopause height. Accordingly, the evaluation is now carried out with these values.
The sentences regarding the missing PV tropopause values will be removed from the manuscript.

- L364: CFC-12 on --> CFC-12 in

  Done.

- L365: channel --> channels

  Done.

- L372: add a comma after “SouthTRAC”

  Done.

- L374: add a comma after “2015/2016”

  Done.

- L375: “At the time of publication, it is not known that such a comparison has already been made”. First, this statement is somewhat ambiguous. I think the authors mean “To our knowledge, such a comparison has not been published previously.” Second, they should be a bit more precise in the language here, focusing on Cly in the polar vortices, given the discussion of mid-latitude Cly in the WMO Ozone Assessment as noted above.

  We now use the suggested wording “To our knowledge”.
  Furthermore, we have rewritten this sentence to point out, that we did not find a comparison of vortices Cl_y in previous publications.

  “[…] To our knowledge, a comparison of Cl_y of the Arctic and Antarctic vortex has not been published was previously. […]”

- L382: would be negative to about --> are estimated to be negative at about

  Done.

- L382-383: The difference of Cly inside the respective vortex is significant and even larger than the inter annual variations reported by Strahan et al. (2014) --> The differences in Cly values inside the two vortices are substantial and even larger than the interannual variations reported by Strahan et al. (2014) for the Antarctic. (See earlier comments on a similar statement in the abstract.)

  Done.

- L384: of the respective --> in each

  Done.

- L385: respective campaign only shows a section of the respective winter seasons. These sections do not match --> respective campaigns only show a portion of the winter seasons. These intervals do not correspond

  This sentence was partly rewritten as suggested.

  "[…] Furthermore, the respective campaigns only show a part of the winter seasons. These intervals do not correspond completely […]".

- L386: the respective polar vortex --> the two polar vortices

  Done.

- L389: add a comma after "SouthTrac"

  Done.

- L391: First, citations need to be added to support this statement about the BDC being stronger during NH winter than during SH winter. Second, I think it would be more appropriate to move the conjecture about a possible cause for the interhemispheric disparity in Cly to section 4.4, where these results are discussed, rather than have it in the "Summary and Conclusions" section. In addition, I'd like to see the discussion of the discrepancy and its possible causes developed a bit more, and put into context of the midlatitude results in WMO 2018 (mentioned above). Also, I suggest some wording changes: on the northern winter hemisphere than in the southern winter hemisphere due to stronger Brewer-Dobson circulation --> during winter in the Northern Hemisphere than during winter in the Southern Hemisphere due to the stronger Brewer-Dobson circulation.

  On the advice of the referee, the presumption is included in Section 4.4. The necessary source for the hemispheric difference of the Brewer-Dobson circulation was added. In addition, the discussion regarding the difference of Cl_y in the mid-latitudes of SH and

NH has been extended. This also includes results from the WMO report 2018, mentioned in a previous comment. It was no longer possible to change the wording because a change was made to the sentence.

"[…] Nevertheless, possible reasons for the observed differences can be derived from the hemispheric difference of the Brewer-Dobson circulation, using the age of air as a common metric for transport. Konopka et al. (2015) showed, that north of 60 °N, age of air is always younger than south of 60 °S in the same season, implying a stronger residual circulation in the Northern Hemisphere. Analysis of Haenel et al. (2015) revealed differences in age of air trends in the lowermost stratosphere of the mid-latitudes of Northern and Southern Hemisphere with a positive trend in the Northern Hemisphere and a negative trend in the Southern Hemisphere. […]".

- L393: in higher --> at higher

  Done.

- L394: exhibits a larger variability as it is more effected --> exhibits larger variability as it is more affected; also, capitalize "Southern Hemisphere"

  Done; Done.

- L395": side -- > hand; is less effected --> is typically less affected

  Done; Done.