Reply to RC2 and RC3

Thank you for your time and effort that you put into reviewing our paper, and for your insightful comments, all of which are addressed below. In the following we reproduce the original comments in black regular font*, and include our responses in blue italic font. Revisions made to the manuscript are indicated in* **blue bold font.**

This manuscript by Parrish et al. describes the analysis of trends in tropospheric ozone at mid-latitudes in both Northern and Southern Hemispheres over the last few decades, using the outputs from community efforts such as HTAP and TOAR. The manuscript is generally well written and presented, and provides a good account of the many factors affecting observed ozone concentrations, and of the challenges associated with modelling them. I have one major reservation on how the sites discussed are assumed to be representative of mid-latitudes for both hemispheres (see Major Comments below). There are also minor points (see below) that, once addressed, would make the text clearer. I recommend publication of this manuscript only once the comments below, especially the major ones, are suitably addressed.

Major Comments

My main reservation about the conclusions drawn by the authors of a reversal in the interhemispheric tropospheric ozone gradient in the pre-industrial era is that they mainly rely on the extrapolation of the fits to two series of observation from two background sites, one per hemisphere, and on the assumption that the polynomial fit derived from European monitoring sites (Fig. 2a) is representative of ozone trends for all NH mid-latitudes.

This is an important comment. It is true that we rely on an extrapolation of the fit to the Cape Grim data to estimate past ozone changes in the SH; with regard to the uncertainty that results from this extrapolation, please see our response to a more extensive comment on this issue made by Cooper et al. (acp-2020-1198-CC1). However, it is important to note that ***we do not simply extrapolate the Mace Head data in the NH****. Throughout northern mid-latitudes long-term changes in average baseline ozone concentrations are found to be the same, within narrow statistical confidence limits, when quantified in relative terms. This uniformity extends over all longitudes and through all altitudes from the surface to at least ~ 9km. The reasons for this uniformity and the evidence that supports it are discussed further in our responses following the next comment paragraph. Therefore, within a relatively small, quantified uncertainty we determine the past ozone changes that occurred at Mace Head before measurements were initiated, by normalizing that uniform relative change to the ozone concentrations measured at Mace Head from 1987-2017. The difference between a simple extrapolation and this normalization process may seem like a minor distinction, but it is of critical importance for accurately evaluating the uncertainty of our analysis.*

While the authors make a compelling case for their conclusions, for instance by showing how most European monitoring sites exhibit similar (relative) trends, it is also evident that some of the data shown in Fig. 2a (notably Arkona) exhibit deviations from the overall assumed trend. Furthermore, comparison with a non-European site (US Pacific MBL) shows potentially different temporal trends from its European counterparts (Fig 3). As the increase in tropospheric ozone in the NH is driven primarily by enhanced emissions of ozone precursors and nitrogen oxides (as the authors rightly point out), how do the authors justify assuming that these increases followed the same trend in Europe, North America and mid-latitude Asia throughout the time period considered here (1950-present)?

This is another important comment. We do not simply ***assume*** *that these increases followed the same trend throughout northern mid-latitudes;* ***simple transport and ozone lifetime considerations, observations, and model simulations*** *all support the* ***conclusion*** *that baseline ozone concentrations followed the same relative long-term changes throughout northern mid-latitudes****.*** *This allows us to confidently normalize the quantified relative long-term change to the recent Mace Head measurements. The observations and model simulations supporting this conclusion have been extensively discussed (HTAP, 2010; Parrish et al., 2012; 2014; 2020) and are illustrated in Figures 2 and S1 of our paper. Simple transport and ozone lifetime considerations support this conclusion; in the free troposphere at northern mid-latitudes the net lifetime of ozone is estimated as 100 days,*

which is considerably longer than either the circum-global transport time (~30 days) or the vertical overturning time scale (~ 20 days). Consequently, even though the many sources and sinks of ozone are heterogeneously distributed, and each may possibly change over long time scales, the relatively rapid mixing and transport ensures that those changes are reflected in average baseline ozone concentration changes throughout northern mid-latitudes. In the presence of relatively rapid transport and mixing, there simply is no mechanism that can maintain heterogeneity in the long-term changes in the zonal baseline ozone concentrations. Parrish et al. (2020; 2021b) discuss these considerations in more detail.

The authors convinced me that the sign of the ozone difference at Mace Head and Cape Grim (the two sites considered representative of their respective hemispheres) might have been reversed in the pre-industrial period for those two sites (or at least that their ratio might have converged to unity, as I have my reservations on the extrapolation of a 4th degree polynomial), however I'm not convinced that this can be extrapolated to all mid-latitudes based on these two sites alone. The authors point out that extrapolation to pre-1988 times (i.e., before the Mace Head record started) is apparently confirmed by measurements (Fig. 3), but as the authors point out, 3 out of 4 are from NH sites of questionable reliability, and the remaining site exhibited large variations in Fig 2 (Arkona). Why not show more points from the Arkona time series in Fig. 3? And perhaps the polynomial fit to the Arkona time series?

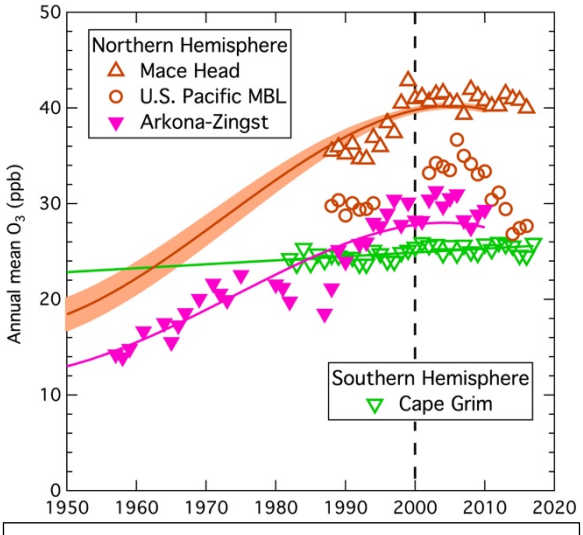

Figure 1. Annual mean ozone mixing ratios measured at 4 low elevation coastal sites. Three sites with fits are taken from Figure 3 of our manuscript. The Arkona-Zingst data and fit are taken from Figure 1 of Parrish et al., 2021.

Figure 1 shows Figure 3 of our paper with the Arkona time series and its polynomial fit included. We do not include the Arkona data in Figure 3 of the paper, because it does not represent undisturbed MBL baseline air; it is a coastal site on the Baltic Sea that receives air with a strong continental influence, which reduces the mean ozone concentrations through surface deposition. The US Pacific MBL data are included to show that the Mace Head data are not abnormally low; in fact they are among the highest NH MBL baseline ozone concentrations observed.

As for the SH, why was the Cape Point dataset not considered? It would be useful to include the Cape Point data series in Fig 3 as a term of comparison with the Cape Grim data, mirroring what the authors did for the NH data with Mace Head and the US Pacific MBL data.

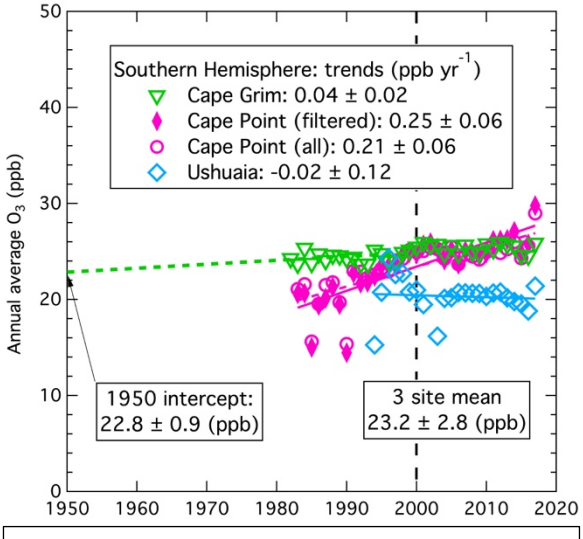

Figure 2. Annual mean ozone mixing ratios measured at 3 SH, low elevation coastal sites. The green line is a standard linear regression to the Cape Grim data, with extrapolation back to 1950. The 3 site mean with standard deviation is indicated in the annotation.

Figure 2 shows Figure 3 of our paper without the NH data, but with the 3 data sets that we considered for the MBL of the SH. Inclusion of the Cape Point and/or Ushuaia data sets in Figure 3 does not provide additional information as these data sets are generally consistent with the Cape Grim data as discussed by Cooper et al. (2014). Importantly, two approaches to estimating the preindustrial ozone concentrations at southern mid-latitudes (extrapolation of the Cape

Grim record to 1950, and simple averaging of all data reported from the three sites, under the assumption of no significant SH trend) give nearly identical results, as indicated in the figure annotations.

I appreciate how some of the points raised here may be difficult to address due to the paucity of pre-1980 data, as the authors point out. However the paper would still make a valuable addition to the current literature on tropospheric ozone if the language in the discussion/conclusions section was adjusted to account for the sources of uncertainty in their analysis, as outlined below.
In writing our paper, we attempted to carefully select language that struck the appropriate balance between clearly stating our quantitative conclusions and properly considering the uncertainty in our analysis. We appreciate the reviewer identifying specific instances where we failed to properly strike that balance. We respond to each of the reviewer's suggestions below.

Line 20 (abstract): replace "likely" with "potentially"
In Figures 3 and 4 of our paper, the shaded area about the NH curve indicates estimated confidence limits for the curve. That shaded area is entirely below the linear extrapolation of the SH line. The SH line is a conservative estimate of the past behavior (see our response to a comment included in the community comment acp-2020-1198-CC1). Thus, the NH curve is below the SH line well outside our estimated 95% confidence limits. Fully considering all of the issues involved, we believe that "likely" is the correct description in the full sentence on Line 20: "The available measurements indicate that this interhemispheric gradient was much smaller, and was likely reversed in the natural troposphere with higher concentrations in the SH."

Line 221: replace "were" with "may have been"
Following the same logic presented in the above response, **we have replaced "were" with "likely were".**

Line 226: replace "must necessarily have been" with "might have been"
Following the same logic presented in the two responses above, **we have replaced "must necessarily have been" with "likely were".**

Minor Comments

line 22: replace "natural" with "pre-industrial"
Replacement made

line 34: add "However, tropospheric ozone…"
"tropospheric" added

line 47 (and again 53): I don't think "inconsistency" is the correct word in the context given. "aspect" would sound better in the paragraph in its current form. I'm assuming the authors think the findings of recent analyses (described in lines 47-55) are inconsistent with the NH always being thought of as the hemisphere with higher levels of pollutants? I would suggest either replacing "inconsistency" or revising the paragraph.
We have changed "inconsistency", and now refer to a "quantitative aspect" on line 47. On line 53, we have changed "resolve this inconsistency" to "explain this issue".

Line 51 – add "higher in the NH than in the SH"
Addition made

Lines 145-150: Need to stress how this is strictly only valid across the temporal range for which measurements are available
The final sentence on these lines has been modified to read: "Figure 2 indicates that to estimate the long-term ozone change at Mace Head (or any other baseline representative site in western Europe) over the 1950 to 2010 period for which measurements were analyzed, one needs only

to quantify the year 2000 mean ozone at the site, and then calculate the product of that intercept with the polynomial fit."

Line 221: change to "Discussion and Conclusions"
Change made

Lines 222-226: You need to add the rate of increase in the SH for this sentence to make sense.
Thank you for identifying this issue; **this sentence has been revised to read: "First, the sparse measurement record at baseline sites indicates that between 1950 and 2000 ozone concentrations increased by a factor of 2.1 ± 0.2 in the NH (Figure 2 and Parrish et al., 2021) and a factor of 1.09 ± 0.04 in the SH (Cape Grim fit in Figure 3), which would imply an approximate doubling (factor of 1.93 ± 0.20); however present NH ozone concentrations are less than a factor of 2 greater than those in the SH (factor of 1.60 ± 0.03, based on year 2000 Mace Head to Cape Grim ratio), indicating that ozone concentrations likely were lower in the NH than the SH in 1950."**

Figures 3 and 4: Can you crop the y axis as starting from 10 ppb?
One goal of these figures is to compare the absolute ozone concentrations between hemispheres in Figure 3 and between measurements and model simulations in Figure 4. Inclusion of the origin of the y-axis in the figures eases that comparison for readers, so we have not cropped the y-axis to start at 10 ppb.

Further comments I missed in my first review:

Line 227: replace "were" with "may have been"
Following the same logic presented in earlier responses above, **we have replaced "were" with "likely were".**

Line 238: replace "were" with "may have been"
The full sentence on this line reads: "Thus, although the measurement record is sparse, observations and the wider considerations discussed above are all consistent with the conclusion that in the preindustrial troposphere, mid-latitude ozone concentrations were higher in the SH than the NH." We believe that "were" is correct in this context, because we are correctly stating that "observations and the wider considerations" are consistent with a particular conclusion, not that the conclusion is necessarily true.

Additional References

Parrish, D.D., et al. (2020), Zonal similarity of long-term changes and seasonal cycles of baseline ozone at northern mid-latitudes. J. Geophys. Res.: Atmos., doi: 10.1029/2019JD031908. (updated reference – cited as 2020b in our original manuscript.)

Parrish, D.D., R.G. Derwent, and J. Staehelin (2021a), Long-term changes in northern mid-latitude tropospheric ozone concentrations: Synthesis of two recent analyses, Atmos. Environ., 248, https://doi.org/10.1016/j.atmosenv.2021.118227. (updated reference - cited as in review, 2020a in our original manuscript.)

Parrish, D.D., R.G. Derwent, and I.C. Faloona (2021b), Long-term baseline ozone changes in the Western US: A Synthesis of Analyses, Preprint https://www.essoar.org/doi/abs/10.1002/essoar.10506269.1