Review of "Hemispheric differences in ozone across the stratosphere-troposphere exchange region" by Sequel et al.

The manuscript compares ozone mixing ratios of stratospheric ozone in the UTLS between the northern and southern hemisphere mid-latitudes, and between high and low ozone depletion years. The analysis is mainly based on in-situ measurements (aircraft and ozonesondes), but is complemented with CAMSRA model output for a better spatiotemporal representation.

I first want to thank the authors for following my suggestion to extend the time range of their original study period and to incorporate additional SH ozonesonde sites (Lauder and Macquarie Island). The manuscript is now also better organized and most of my questions for clarifications have been answered.

General comments

However, I'm still missing somewhat the purpose and focus of the paper. The purpose of the paper and the method followed by achieving this should be mentioned more clearly in the beginning of the manuscript. You describe some different elements of the puzzle in each paragraph of the introduction (STE, available data, SSW), but you do not lay the puzzle yourself by e.g. explicitly linking these phenomena (STE, SSW) with their possible impact on hemispheric ozone differences, why concentrating on mid-latitudes in this study, why it is important to discriminate between high and low ozone depleting years in this study. You assume that the reader will see the puzzle during the course of the paper. Also in the introduction, I'm missing some insight on what the manuscript wants to add to the current knowledge?

Therefore, 2 suggestions, for the introduction:

- add what is already known about UTLS ozone mixing ratio differences between the NH and SH, and what is known already about the impact of high/low ozone depletion in the stratosphere on UTLS ozone.
- Add a true roadmap for your study and give a short explanation for each step (instead of the paragraph from lines 85 to 96), for instance:
    - **We want to study hemispheric ozone differences in the UTLS at mid-latitudes during spring**. Why in the UTLS? Why at mid-latitudes? Why during springtime? Why are such possible hemispheric ozone differences important?
    - **We will only look at the UTLS ozone of stratospheric origin**. Therefore, in our analysis, we make a distinction between high and low stratospheric ozone depleting events/years.

Specific comments

- After 2 Data and before 2.1 SouthTRAC data, write a short introductory paragraph in the style of "The UTLS ozone measurements used in this study are available from (research + commercial) aircraft and ozonesondes and are complemented with chemical reanalysis vertical ozone

profiles. To determine the stratospheric or tropospheric origin of the ozone data, we used water vapour (or humidity) measurements from the aircraft and radiosondes coupled to the ozonesondes and from the chemical reanalysis, and additionally CO, HNO_3, HCl measurement from some of the aircraft data. In the next section, we give more details on these datasets.

- Caption Fig. 1: mention explicitly for which time period the number of observations is shown here.
- Again, after 3 Method and before 3.1 Study period & UTLS definition, it is important to provide some guidance to the reader. Therefore, write a short introductory paragraph in the style of "With the data available and described in the previous section, we will now describe how we will analyze springtime UTLS ozone differences at the mid-latitudes between both hemispheres. We first describe how our study period and the used UTLS definition in 3.1, mention how we distinguish between high and low stratospheric ozone depletion years in 3.2, and we show how we ascertain the stratospheric origin of the analyzed UTLS ozone concentrations in 3.3."
- Section 3.1 misses a real focus. The first two lines (176-177) belong to the SouthTRAC data description. Define the UTLS (300-200 hPa) and the free troposphere (700-300 hPa) clearly. Mention that your study focus on the springtime only, and already define this periods (4 Mar – 20 May, NH & 4 Sep – 20 Nov, SH) here.
- Section 3.2 and Table 2: Here, some major clarifications are needed. Low ozone depletion years are defined by the presence of a SSW event (last column of Table 2). However, for most NH years, the SSW central date lies well ahead of the study period (Mar – May), so it is not clear of the SSW event still occurs during the study period. If this is not the case, it should be mentioned what the expected impact of a SSW event earlier that year would be on the UTLS springtime ozone concentrations. To me, a more direct distinction between high and low ozone depletion years could be made by simply looking at the stratospheric (or total) ozone amounts, averaged over the 45-60° latitude bands, for the 4 Mar – 20 May (NH) or 4 Sep – 20 Nov (SH) period. Please discuss.
- Table 2: add a column with NH and SH to the left, add either "IAGOS" or "SouthTRAC" before flights and add that these latter flights only occurred in 2019. Also, I don't understand why the total number of IAGOS flights in the NH for the two periods (later winter-early spring & mid spring), resp. 456 and 835 flights, is not equal to the number of IAGOS flights in Table 1 (6315).
- To me, it makes more sense to incorporate parts of section 4.1 in section 3.2. In its current form, section 4.1 gives the impression of being a rather standalone section, and does not entirely seem to fit within the logical flow of the paper. You could solve this by transferring parts to section 3.2. Basically, in section 3.2, you want to isolate the air of stratospheric origin in the UTLS to look at its ozone concentration properties. So section 4.1 should make a link to this section, but now using tracer correlations to look at the origin of the UTLS air masses. Also the choice of the 4.1 section title could be better. The sentences at the end of page 8 might be replaced by a better guidance to the analyses done in sections 4.1 and 4.2, and referring more directly to these sections (rather than describing in rather general terms, as is done now).
- Figure 2: in addition to my previous comment: argue why these O3-H2O tracer correlations are important, and what you learn from them, and if there is difference between the LD and HD correlations. Also, specify in the caption which periods (2002-2022?, 4 Mar – 20 May NH?, 4 Sep – 20 Nov SH?) have been used for the correlation plots.

- Figure 3: specify in the caption which periods (2002-2022?, 4 Mar – 20 May NH?, 4 Sep – 20 Nov SH?) have been used for the correlation plots.
- Figure 4: specify in the caption that these are SouthTRAC flights only!
- Line 269: what does the comparison between Fig. 5 (stratospheric origin) and A1 (mixture of tropospheric and tropospheric origin) learn us? For instance, for late winter – early spring the SH LD ozone amounts are much larger than the SH HD ozone amounts in A1 (no RH filter), compared to the same comparison in Fig. 5 (RH filter). Any clue for this?
- On page 12, line 280, you define differences between high and low-depletion years as "interannual variability". I don't think that "interannual variability" is a good term for it; it is such a general term. Also you should define how you calculate the ozone difference between high and low-depletion years. It is simply the difference between the mean of the overall springtime ozone between 300-200 hPa for both the LD and HD years, or do you somewhat average out the mean values for every pressure level, as shown in Fig. 5? How are then the values mentioned in lines 280-282 calculated? Please specify! From Fig. 5, I would assume that the ozone differences between LD and HD years are nowhere significant, so I don't understand quite well how the values, and their uncertainties, in lines 280-282 have been obtained.
- In Fig. 6: can't the comparisons (vertical dashed lines) with the ozone medians in the NH obtained from IAGOS added to this figure, similarly as has been done for the SH?
- Related to previous comment: the analysis of this figure 6 is only used for comparing the SouthTRAC measurements (2019 only) with the CAMS reanalysis output (the low-depletion years 2002 and 2019?). Is this the most important message from this Figure? Shouldn't the (pattern) agreement between Fig. 5 (observations) and Fig. 6 (model) be discussed first? E.g. the higher NH ozone values, the slightly (but not significant) higher amounts for LD years compared to HD years? And how do the model values relate with the measurement values (for all measurement types, and in NH and SH)?
- Lines 314-325: those findings are really not obvious at all from the figure. Are you saying that the agreement between CAMS and SouthTRAC (SH LD, light blue) is better during late winter – early spring compared to mid spring? On which ground? For pressures lower than 270 hPa (instead of below 270 hPa)? What do you mean with higher "fluctuations" in the medians for the mid-spring period? In the vertical or looking at the percentiles of the medians (I don't see this for the latter)? Are fluctuations in the vertical relevant given the coarser vertical pressure grid for CAMS?
- Figure 7: in the discussion of Fig. 7, it might be informative to give the overall mean UTLS ozone value for each considered case (NH LD, NH HD, SH LD, SH HD).
- Conclusions: if you put so much weight on the analysis in section 4.1 (two paragraphs devoted to it in the conclusions), you should stress much more the importance of this analysis in the main text as well. As written before, for me, it really hampers the logical flow of the paper a bit in its present form.

Technical corrections

- Line 39: remove the dot before OH
- Line 86: remove the brackets before and after Neu et al., 2014
- Line 86: remove "diminished"
- Line 90: remove "from"

- Line 93: in situ measurementS. Futher: replace “air masses exhibiting stratospheric character” with “air masses of stratospheric origin”
- Line 135: replace so-called initiative with the Tropospheric Ozone Assessment Report – Phase II (TOAR-II) Focus Working Group HEGIFTOM (Harmonization and Evaluation of Ground-Based Instruments for Free-Tropospheric Ozone Measurements).
- Line 161: an identical instead of the identical
- Line 184: proposed by Charlton and Polvani (2007)
- Line 207: Section 4.1 instead of Section 3.2
- Line 237: give the pressure level instead of “above 9 km”.
- Line 266: and instead of y
- Line 305: add ozone between percentage and difference
- Line 358: replace “characterized by” with “of”.