Reply to Reviewer #1

Weber et al., Global total ozone recovery trends attributed to ODS changes derived from five merged ozone datasets, doi:10.5194/acp-2021-1058

Reviewer comments are provided here with our replies written in italics

1. Short resume

Weber et al. present a comprehensive analysis of trends in total ozone, focusing primarily on the period since the turnaround in ozone-depleting substances. This is an update and extension of earlier work published in 2018. In contrast to latter publication, the authors now claim the detection of increases (0.4%/decade) in near-global (60S--60N) total ozone since 1996, with high confidence (>3-4σ). Positive trends over broad mid-latitude region in both hemispheres (35N--60N and 35S--60S), about 0.5--0.7%/decade, are significant as well although close to the 2sigma detection threshold.

The dynamical process terms (Arctic and Antarctic Oscillation, Brewer-Dobson circulation) in the regression model play a central role in this detection, especially at northern mid-latitudes. The authors deliberately chose not to detrend the dynamical terms prior to regression, in order to account for any long-term changes in AO, AAO and BDC. In doing so, they find that trends become less negative before 1996 and more positive since 1996 across large regions of the low- and mid-latitudes. This more complete attribution results in a higher significance of the trends, especially in the northern hemisphere where the 2sigma detection threshold was passed. Hence, the authors conclude that dynamical changes appear to counterbalance the recovery of ozone in the mid-latitude NH.

The authors furthermore explain the positive recovery trend of total ozone as a result of changes in ozone-depleting substances. Indeed, the ratio of the rate of increase and decrease in ODS concentrations is consistent with the rate of depletion and recovery of total ozone across all 5° latitude bands between 60S and 60N.

2. Recommendation

This paper provides an important update to previous assessments of long-term changes in total ozone. It is very well written and accessible to a large scientific audience. The methodology is sound and the presented results support the claims made by the authors. I highly recommend publication of this work in ACP if my remarks below have been addressed.

3. Major comments

Ordered in order of appearance in the text.

3.1 Extension of GOME-type backwards in time (Sect. 2.7)

I understand the importance of covering a sufficiently long period, but is this backwards extension for GOME-type data records still needed now that more than two solar cycles have been completed since 1995? Doesn't this break the independence between SBUV and GOME-type estimates? By how much does the negative trend in the SBUV period influence the recovery trend estimates during the GOME-type period? Have you tested the sensitivity of the resulting trend to the choice of NASA COH or NASA MOD, and without the extension?

→ *We try to stay consistent with the W18 paper, where we also applied for these extensions. The main idea of extending to the full-time period is to have as close as possible the same impact from all proxy terms, not only the solar term, during the full-time period. Since we use independent linear trends before and after the ODS peak, the impact of early trends on the late period trends is minimised.*

Avoiding data gaps is important but preserving data quality/stability is perhaps even more important under high aerosol backgrounds. Could you elaborate why gaps are more important or, if that is not the case, comment on the stability of both SBUV records after Pinatubo?

→ *Calculated annual means were accepted as valid if at least 80% of the monthly means were contributing (10 months minimum) and 80% of the 5° zonal means were available in the broader zonal bands. If these conditions are not met we consider the annual mean as not representative for the given year and should be excluded from the MLR. The consistency of the SBUV data records with other total ozone data has been documented, e.g. Chou et al, 2014.*

3.2 No reference to how trend errors are estimated (Sect. 3)

Many trend estimates (Fig. 3) are close to the 2sigma threshold. The computation of MLR coefficient uncertainties, therefore, deserves some attention, this is missing right now. Please explain how MLR parameter errors are computed or refer to relevant publications. → *All uncertainties are given as 2sigma and sigma is calculated from the least-squares fit. This is a standard approach and is described in many statistics textbooks.*

Somewhat related to this, was there any consideration of including reported measurement errors in the regression? → *Measurement errors were not accounted for. Not for all merged datasets, an uncertainty estimate is provided from merging the data.*

3.3 Annual time series

p.7, l.180: Could you motivate the choice for analysis of annual mean time series instead of monthly mean data? Is there an impact on the trend estimates and their significance? Please refer to relevant publications. → *The main reason for using annual means is that this does not require corrections for auto-correlation (mentioned in the text). Adding auto-correlation terms in the regression will not alter the trends but increases uncertainties in the fit coefficients (and trends). The short-term variability is not the focus of our MLR except that we try to minimise the residual of the regressed timeseries.*

3.4 Robustness of attribution to dynamical processes (Sect. 5)

Previous work by the authors (Weber et al, 2018) also considered terms for dynamical processes in the MLR. At the time, however, no significant positive trends were detected (Fig. 9). → *see our general comment in the beginning of the reply to reviewer #2*

It would be enlightening to discuss whether the four additional years of data have truly helped to attribute ozone changes more robustly to dynamical changes. Or, whether it is plausible that the current attribution is subject to geophysical variability (and measurement uncertainty). → *see our general comment in the beginning of the reply to reviewer #2*

4. Minor comments

p.1, l.12-13: Near-global trend values disagree with quoted values in Section 4. Please revise. → *Numbers in the abstract have been adjusted to the values shown in Fig. 1.*

p.3, l.82: "Annual mean timeseries of all five merged datasets are in very good agreement". Somewhat subjective, please add a number. → *add: "… to within a few DU"*

p.5, l.132: The evolution in satellite quality has been described adequately. This is missing in the WOUDC section. Surely, there must have been progress in the calibration of these instruments or the coherence of the network since the work by Fioletov in 2008. If so, could you update this section accordingly? → *The ground-based network calibration procedures have been established a long time ago and there are no major changes in the network operation. The same is true for the WOUDC data set that is regularly generated by the WOUDC. We added a reference to a recent paper where the differences between satellite and ground-based data are discussed on a global scale. We added a reference to Garane et al.*

p.6, l.142-143: "[...] can be estimated with a precision comparable with satellite-based data sets (~1%)." A reference would be appropriate. → *Comparison of satellite and ground-based data sets is discussed in the following paper: Chiou, E. W., Bhartia, P. K., McPeters, R. D., Loyola, D. G., Coldewey-Egbers, M., Fioletov, V. E., Van Roozendael, M., Spurr, R., Lerot, C., and Frith, S. M.: Comparison of profile total ozone from SBUV (v8.6) with GOME-type and ground-based total ozone for a 16-year period (1996 to 2011), Atmos. Meas. Tech., 7, 1681–1692, https://doi.org/10.5194/amt-7-1681-2014, 2014. It was referenced in line 140 and we added this reference to line 143 (at the end of the last sentence)*

p.6, l.150: Remove "from the past into the future" as the statement "between 1960 and 2100" is more than sufficient. → *done*

p.6, l.154-156: I am sorry, I did not get the point of "The multi-dataset mean was then added back to each dataset, such that all bias corrected timeseries are provided in units of the total column amounts (W18). However, the trend results derived from them are identical to those derived using anomaly timeseries." Could this be clarified a bit better for the non-expert? → *This procedure means that the bias-corrected time series differ from anomaly timeseries by a constant offset (multi-instrument mean). The bias correction has no influence on the calculated trends but makes the data more legible in the plots.*

p.6, l.154: "to the mean". The 1998-2008 mean at the global or local level? → *all data are annual mean zonal means and for each zonal band considered an average for the period 1998-2008 was calculated for each dataset and a mean over all datasets (multi-instrument mean) calculated*

p.6, l.165: See comment below, the second term in Eq. 1 should be b_1 (t-t_0) → *this is not correct, since t_0-t is positive (t_0>t), b_1 will be negative if ozone declines.*

p.6, l.166: "coefficients b_1 and [...]" This is inconsistent with the notation in Eq 1. Sign of first trend term (t0-t) implies that positive b_1 values represent a decline in ozone. Please change this. The factors X_1(t) and X_2(t) define the decline/recovery periods. → *see previous statement*

p.6, Eq.2 and 3: Figure 1 suggests that the "recovery" period starts in 1996, so the turnaround is defined as $t_0=1996$. If this is correct, then the notation in Eq. 2 and 3 should be changed to $X_1(t)=1$ for $t < t_0$ and $X_2(t)=1$ if $t \leq t_0$ (and vice versa for $X_i=0$). The trend model is not continuous at t_0, hence $<$ or $\leq$ do make a difference. → *This was indeed not consistent and has been corrected at several places. The early period is t < t0, late period t >= t0. So the first period includes t_0=1996 and the late period starts with 1996. As mentioned in the text the shift of t_0 back and forward did not change the trend estimates.*

p.7, l.185-187: Is there any particular reason why you haven't used GloSSAC v2 (Kovilakam et al., 2020)? → *We actually tested the Glossac dataset, but we found only negligible differences in the trend estimates. This is likely due to the fact that the El Chichon and Mt Pinatubo eruptions dominate the stratospheric aerosol optical depth proxy timeseries. This effect is even enhanced since we use two proxies to separate both major volcanic events.*

p.8, Table 2: EHF is missing from this list. Where can it be downloaded? → *As mentioned in the text, the eddy heat flux was calculated by us from the ERA5 reanalysis data and was not taken from an external source. A description of how to derive the eddy heat flux from reanalysis data is given in W18.*

p.8, Eq.4: "BDCn" and "BDCs" should be explained in the text. → *added, "The BDCn and BDCs are 100 hPa eddy fluxes in the northern (n) and southern hemisphere (s)."*

p.8, l.208: "the linear trend terms best approximate EESC related trends". Can a match between ozone trend and EESC expectations really validate the choice of terms in the MLR? There is a risk of a circle reasoning here. If the improved agreement with EESC expectations is motivating the choice of terms in the MLR model then you can't use this same agreement again to conclude a causal relation between trend and EESC. → *We only assume that all trends not related to ODS changes are contained in the proxy terms. The linear trend before and after the ODS peak is independent, but It turns out that the trend ratio before and after the ODS peak is consistent with the rate changes of EESC to within the uncertainties from the regression. However, we know that there are feedbacks between ODS (ozone) and climate (dynamics). Therefore, the linear trends will only approximate the ODS related contribution to ozone changes.*

p.8, l.215-216: This phrase is not entirely clear on whether or not you use the detrended proxy. This choice is so central to this paper that it must be very clearly stated. → *We added, "For these reasons, we do not detrend the proxy timeseries in this study".*

p.9, Fig. 1: chi^2 is the sum of "the squared differences median timeseries minus MLR" → *We changed to ".. sum square of differences between median and MLR timeseries' divided by ..."*

p.9, l.219: "MLR prediction after fitting" would be clearer than "MLR result from applying". → *better: "MLR timeseries derived from"*

p.9, l.220: To me, "after 1996" suggests 1996 is not included. What about replacing "after 1996" by "since 1996" throughout the manuscript? → *see earlier comment. It should be "after 1995" or "since 1996", similarly "before 1996" and "until 1995". We changed accordingly.*

p.9, l.224: "recovery from reductions in ODS" would be more clear on the effect of ODS on ozone. → *done*

p.11, l.260: Replace "from applying" by "when applying"? → *leave it as is.*

p.11, l.260: It is somewhat unexpected to regress a "super" merged timeseries rather than average the trends from individual records. What is the rationale? Also, the sample size is just N=3, for 1979-1995, so won't the "super"merge-then-regress method lead to more uncertainty in the MLR parameters than the regress-then-average approach? → *In Table 3 we present the trends of the median timeseries' as well as the median and mean of the individual trends. The numbers are nearly the same.*

p.12, Table 3 (caption): The periods in the caption are inconsistent with information in Figs 1 and 2. The first trend period stops in 1995, the second starts in 1996. Hence, it should be 1979-1995 and 1996-2000. → *done*

p.12, Table 3: For each latitude belt, the occurrences of "mean/median trend >1996" should be >=1996, in order to be in line with Fig. 1 and 2. → *changed to t>1995*

p.12, Table 3: The error notation was confusing for me, I haven't seen this specific notation very often. For instance, what does -1.9(13)" mean? Is it -1.9+-0.13 or -1.9+-1.3 or ...? I find an explicit notation such as "+0.4+-0.2" much more effective. I recommend using this throughout this table and also the manuscript. → *It is a common way to put uncertainties in the brackets, but I agree that this is not so widely used in the atmospheric science community. In order to keep the table compact, we will remain with our notation.*

p.12, l.265: "One notable change from W18 is that the tropical trends during the ODS rising phase are now more negative (down to -1%/decade) while before they were mainly close to zero. This may be caused by the additional proxy terms used in this study". The pre-1996 data have been available for a very long time now. Has this effect never been looked into before? If so, please refer to relevant work. - → *see our general comment in the beginning of the reply to reviewer #2*

p.12, l.270: Please replace the "maybe" (conditional) by an "is" (certainty). Trend uncertainty scales with n^(-3/2} (e.g., Weatherhead et al., 2000) so the eight more years in the recovery period already lead to ~45% smaller trend error. This seems not too far from the observed factor 2 reduction of the error in Table 3 and Fig. 3. → *done*

p.13, l.274: "The expected tropical recovery [...]". Estimated mid-lat NH recovery trends are too small compared to EESC prediction as well. → *added "In the NH extratropics the expected ODS related*

recovery is slightly higher than the observed trends, but also agree within the uncertainties of the observed trends."

p.15, l.320: "NH total ozone has been steadily declining..." conflicts with the first phrase of this paragraph "stable ozone levels in NH since 2000". Please clarify the text. → *The stable levels refer to annual means at NH middle latitudes as shown in Figure 2 (added "middle latitude"), while the decline in Figure 6 is shown for March only and also includes polar latitudes.*

p.15, l.324: "with larger springtime polar ozone losses"? → *done*

p.15, l.325: Remove "recent" from "A recent downward trend". Perhaps you meant that this was recently reported? Ball et al report a continuous decline since the 1980s, not a recent decline. → *done*

p.18, l.332: Quoted recovery trend value (11%/decade) conflicts with that in Figure 7 (12%/decade). Please correct. → *done*

p.19, Table 4: Same comment on error notation as in Table 3 (p.12). → *see earlier comment.*

p.19, l.367: The Gaudel paper is about differences between tropospheric ozone data records. So probably not the best reference when the message is about consistency between tropo/strato/total ozone. → *We removed this sentence, as we did not mention tropospheric ozone at all in the paper. When using annual mean zonal mean averages contribution of tropospheric ozone is likely to be small, but may become more important when looking at regional trends.*

5. Technical corrections

→ *all done*

p.1, l.10: Remove "on" from "[...] is indeed on slowly [...]".

p.1, l.12: Remove "in absolute numbers".

p.1, l.15: Add "-" to "chemistry-climate models".

p.2, l.30: Typo "stratosphere".

p.2, l.38: Remove "agreement" from "Montreal Protocol agreement".

p.3, l.75: Replace "in large part" by e.g. "largely".

p.3, l.79: Replace "Observations Zénithales" by "Observation Zénithale".

p.4, l.87: Replace "are processed using the same V8.7 retrieval algorithm" by e.g. "are retrieved using the same V8.7 algorithm".

p.4, l.108: Type "[...] shift to an equivalent [...]".

p.5, l.130-132: Double occurrence of ground-based. First one could be removed, e.g. "The WOUDC zonal mean ...".

p.7, l.175: Add "." after "W18)".

p.7, l.189: Replace "there are not sufficient number of months" by e.g. "there are not enough months" or "there is not a sufficient number of months".

p.7, l.194: Replace MLR "equation" by MLR "model"?

p.8, l.212: Remove "the possibility", as it is a bit redundant.

p.8, l.212: Replace "MLR results" by "MLR fit residuals" perhaps? This is a bit clearer as the MLR parameter estimates are MLR results as well.

p.9, l.218: "five bias-corrected" instead of "bias-corrected five".

p.11, l.242-243: Maybe you forgot to remove the newline between paragraphs?

p.11, l.251: Add a “+" sign to the quoted values at start of this line.

p.11, l.256: Remove ' after "timeseries".

p.11, l.261: Add "/decade " after +0.5%”

p.12, Table 3 (caption): Remove "and" from caption "[...] in bold have an absolute [...]"

p.12, Table 3 (caption): Add "prediction" at the end of "and mod_ithe MLR".

p.12, Table 3: Add $+$ to trend value $\geq$1996 for median time series near-global.

p.12, Table 3: The quoted r^2 value for WOUDC in 20S-20N band is single digit (0.7), should be double (0.70).

p.13, l.276: Remove "on" from "elucidate further on".

p.13, l.285: Type "Fig. 4a" should be "Fig. 4".

p.15, Fig.5 (caption): There is a missing word in "Negative values an anti-correlation [...]".

p.15, l.311: Add "s" to "chemical effect"?.

p.15, l.316: Add full stop at end of phrase.

p.17, Fig.7 (caption): Capitalise "See".

p.18, l.331: "Earlier signs of ozone recovery have been", should be plural.

p.18, l.331: Add "," in between "Now with".

p.18, l.332-333: "During September, the Antarctic ozone hole usually grows and [...]".

p.18, l.340: Remove "as shown in Fig. 7". A bit redundant, you already referred to the figure in the previous phrase.

p.18, l.344: Replace "globally" by "global"?

p.18, l.352: Add "," in between "tropics recovery".

p.18, l.354: Add "," in between "Arctic large".

250 p.19, l.363: "chemistry-climate models".