Comment on acp-2022-182
Anonymous Referee #2

Referee comment on "Effects of Reanalysis Forcing Fields on Ozone Trends and Age-of-Air from a Chemical Transport Model" by Yajuan Li et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-182-RC1, 2022

Review of Li et al., 2022

Li et al. compare in their manuscript ozone, simulated with the CTM TOMCAT, based on the reanalyses ERA-Interim and ERA5 with different observational collections for total column ozone and ozone profiles in the time frame 1979-2018. They find that ERA5 better reproduces measured TCO anomalies, while trends and profiles are better reproduced by ERA-Interim.

The paper discusses an important topic, which is certainly within the scope of ACP. However, major revisions are necessary before publication in ACP. In particular the structure of the paper should be revised to have it more concise, with some of the current figures moved into a supplement. In addition, some background information of the usage of the TOMCAT model and the used methods are missing. In the second part of the paper, it is a major weakness that the observational SWOOSH data is barely used. The AoA part seems to be not connected to the preceding part of the paper.

General points:
- In the methods, the PWLT regression model is introduced, but later in the analyses it is not mentioned in which way the shown data has been treated by this model. Probably, the model has been used to calculate the anomalies, but this should be mentioned.
- Section 3.1 is lengthy for my taste. Information from some of the figures is repetitive, so maybe some of the figures could be moved to a supplement? Further, the structure of this section is not very clear to me. Maybe it could help to have sub-subsections? Or maybe it is sufficient to shorten this part?
- In Section 3.2, the observational SWOOSH data is barely used in this article. Instead the profile information of both model runs is compared with each other. It would be more helpful to assess which reanalysis data set performs better by comparisons to the observations.
- Similar to the comment on Section 3.1, Section 3.2 would benefit from more external structure.
- It would be very helpful to know how the two different observational data sets compare to each other, since the outcome of the comparisons to the simulation results are quite different. I think it would be a good idea to calculate TCO from the SWOOSH profiles and include these data in the comparisons of Section 3.1 (or in the supplement) to show to which degree these observations are consistent.
- Section 3.3 seems to be a related, but different topic. It is also not reflected by the title of this paper. I would suggest to remove this AoA part from this manuscript, and rather focus on the ozone trends. Otherwise a lot of more work would be necessary to properly introduce the measurement data sets used to compare the model with. However, this would result in an even longer paper, which may be less focused.

Specific points:

**Introduction:**
- Line 45: "Besides the decrease in ODSs, cooling induced by increased greenhouse gases (GHGs) slows the rate of ozone loss [...]": This view is a little too simplistic, as it is known that cooler stratospheric temperatures may lead to larger PSC areas and thus more catalytic ozone loss (see e.g., Rex et al., 2004, 10.1029/2003GL018844)

**Section 2.1:**
- It is noted that the TOMCAT CTM uses ODSs, GHGs, aerosols from volcanic eruptions, but it is not mentioned how these variables are initialized, and how emissions are handled.
- For the TOMCAT run with ERA5, it is not stated how the considerably higher resolved ERA5 data is degraded to the lower resolved TOMCAT horizontal and vertical resolution. It should be further discussed, how this downsampling could influence the comparison with ERAI, which has a more similar horizontal and vertical resolution as TOMCAT.
- Further, it is not clear to me, if for the ERA5 run, the hourly output of ERA5 has been used. Again, the influence of different temporal sampling from the driving fields should be discussed.
- It would be interesting to know what kind of observations have been assimilated to ERAI and ERA5. I guess they also both used some kind of ozone information to get radiation and dynamics.

**Section 2.2:**
- The URL to the C3S data directs to a page which states, that it should not be used any more. Instead this data set should be used: https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-ozone-v1?tab=overview. It looks like, the authors used the data set with known bugs for this manuscript. The authors should discuss these known bugs and how these are affecting the analysis.
- For the SWOOSH data some information of vertical resolution would be helpful, in particular in comparison to the vertical resolution of TOMCAT.

**Section 2.3:**
- Equation (1) is written down in a confusing way. After reading all the text, it becomes clear, why there are two linear terms, but it would be better to define the function piecewise instead (see https://en.wikipedia.org/wiki/Piecewise).
- I guess the proxy functions have been used earlier. It would be helpful to have a reference to a publication, where the reader could look up the $X_i(t)$ functions used for
this work.
- Line 154: Typo "Steinbrecht et al. 2017"
- Line 166: Missing subscript in Xi

Section 3:
- I do not see the point of the text here. It reads like an introduction to Section 3.1, so it could be moved there directly.

Section 3.1:
- Figure 1 (caption): Please also mention that these are monthly means displayed here.
- Line 182: "Anomalies are calculated by subtracting the long-term monthly average from each monthly mean value.": Does that mean that C_0 from equation (1) is subtracted? Otherwise, how is the PWLT regression model used here?
- Line 193: "[...] obtained from merged C3S and TOMCAT simulations, A_ERAI and B_ERA5 [...]": For me, this reads like C3S and TOMCAT simulation data are merged in some way, but the figure looks more like these individual data sets are displayed together in the left panels of Fig.1.
- Line 208: "[...] such as representation of dynamical processes in the ERA-Interim reanalyses.": Please give some examples of such dynamical processes.
- Figure 2: Color bars are missing.
- Figure 2 and corresponding text: Again, it is not clear to me, how the PWLT regression model has been used here, as it was stated in the beginning of Section 3.
- Line 221: "[...] B_ERA5 shows larger positive biases (more than 15 DU) during NH winter-spring seasons [...]": Why is that so? It seems like Arctic ozone depletion is underestimated here. Are the Arctic vortices not represented well in ERA5 (and ERAI in a less pronounced way)?
- Line 239: "Both CTM simulations capture the observed seasonal characteristics of TCO variations averaged across all latitude bands considered here": How can one make such a statement based on DJF and JJA timeseries over several years?
- Figure 3: The caption should warn the reader that y-axes change for the different panel rows.
- Figure 3: Are the authors really sure about their correlation coefficients? For panel 3e, the red line seems to be further away from the black line almost for the whole time series than the blue line, but its correlation coefficients is said to be better. Also, for the other panels, the rather good correlation coefficients look strange to me. But maybe the offset is confusing my eyes.
- Discussion of Figure 3: Both models are overestimating DJF O3 in the north, and in particular ERA5 is underestimating DJF O3 in the south. Please discuss.
- Figure 4 caption: "The absolute differences with the standard deviations averaged over the whole period between simulation A_ERAI (B_ERA5) and C3S are presented in blue (red) text.": This is confusing. Maybe the authors mean: "The absolute differences averaged over the whole period between simulation A_ERAI (B_ERA5) and C3S are presented together with the standard deviations in blue (red) text."?
- Discussion of Figure 4: This is very repetitive compared to the discussion of Figure 3. I suggest to move one of the figures to the supplement and bring both discussions together.
- Line 277: "The regression model used here is identical to that used in Li et al. (2020), except for the different explanatory variables considered for different latitude bands.": Is this regression model also identical to the one described in Section 2.3? In that case, this sentence should be moved there, and in Section 3.1, it should be referred to Section 2.3. Otherwise, this different regression model should be introduced in more detail in a new section 2.4.
- Table 1: The information of this table could be integrated into Figure 5 (similar to the
regression coefficients in Figures 1 or 3).
- Line 297: "[...] with more significant decreases at NH and SH mid-latitude bands for the simulations than C3S." : Does that mean that there is a general problem in TOMCAT or in both reanalyses?
- Line 299: "The recovery since 1998 (Trend2) from C3S is quite different to that from the simulations in terms of its magnitude and significance." : This is not really true for the tropics, where Trend2 seems to agree between C3S and ERA5.

Section 3.2
- Figure 6: I got confused by the titles "1984-2018" for all panels but the first column. Since in all of the panels the same time range is shown, this information could be replaced by a more meaningful title (e.g. "Rel. Diff to SWOOSH" in column 2).
- ERA5 seems to be a bit worse than ERAI for all latitudinal bands, except for the southern UTLS region. But in general, both TOMCAT runs agree in underestimating the upper stratosphere, and overestimating the lower stratosphere. Can the authors exclude that this is a general problem of TOMCAT? How much is the vertical resolution of the model, since it only has 32 vertical levels?
- Figure 7: Why are the models not compared to the measurements over time? I find it hard to judge which model performed better concerning the profile just by comparing profiles averaged over 34 years! In Fig. 6 it can be seen that both simulations having a hard time to reproduce the measurements. It would be helpful to know if this is an overall problem or limited to a certain period! Figures S1 and S2 are more important than Figure 7!
- Figure 8: Again, it would be more helpful to see a comparison with the observations here.
- Line 354: "Large biases in temperature anomalies between two simulations [...] confirming that some of the inhomogeneities seen in ERA-Interim upper stratospheric temperatures [...] have been corrected in ERA5." : How is this statement supported by the figures shown here? The reader can only see that both model runs are different in the stratosphere. Without comparison with the observations, such a statement is not really solid. This problem continues throughout this section.
- Line 370: "Figure 9 shows the PWLT trends for the zonal mean ozone anomalies over the periods 1984-1997 and 1998-2018 obtained from SWOOSH, A_ERAI and B_ERA5 simulations." : I think for these trends the variables C_1 and C_2 have been introduced in equation (1). Please use these variable names here, too.
- Line 375: "It is also important to note that much smaller ozone concentrations in this region means larger retrieval errors for satellite measurements that are used in SWOOSH data set." : What are the errors in this region for the SWOOSH data set? How do these errors translate into these trends? Are these errors able to explain the differences to the model? Please provide some numbers from an error analysis here!
- Please give the reader some help to read the figures and tell what is meant with "upper stratosphere" or "lower stratosphere". Some pressure ranges would be very helpful here.
- Line 380: "In the lower stratosphere, both SWOOSH and A_ERAI show negative trends in the tropical and NH extratropical regions, while B_ERA5 shows increasing trends throughout almost the whole extratropical region." : For me, ERAI and ERA5 look very similar in the lower stratosphere, with ERAI having more pronounced features (e.g. 20°N, 40 hPa) compared to ERA5. SWOOSH trends instead look very differently and do not have such kind of a feature at all, which means that ERA5 performs better in this region. However, this is barely discussed here. Further, the very strong trend seen by ERAI in the Antarctic (which does not agree with SWOOSH) is not mentioned here.
- Line 395: "The trends derived using simple Ordinary Least Square (OLS) method are generally in good agreement with those derived from MLR": I am confused here: Are the trends in Figure S3 computed in a different way? If so, this method should be introduced in the methods section, and it should be mentioned just when introducing the figure in the
text. At this point, this sentence is out of place.
- Line 397: "Hence, these results show that ozone trends from B_ERA5 should be considered with care." I do not agree with this statement. I rather would say that both of the simulation results should be considered with care, depending on which region you look at. For the zonally averaged profiles, ERAI may outperform ERA5, but after seeing Figure 9, I would be very careful with such general statements.

Section 3.3
- As mentioned in the general part, I would suggest to remove this section and prepare it for a different publication. In that case, the observational data could be introduced properly (which is missing now completely), and additional, such as satellite data sets of AoA could be considered in a more detailed comparison. This would also help to reduce the number of figures in this paper: 12 figures are really a lot and I fear that the reader will lose its focus at this point.
- If the authors decide to keep this section, they should carefully restructure the whole paper, in order to connect the ozone trends with the AoA trends. This should also be reflected by the title, the introduction and the conclusions. Right now, these parts are rather separate.

Section 4:
- Line 504: "The PWLT-based regression model shows that both SWOOSH and A_ERA1 show negative trends since 1998 in the NH extratropical lower stratosphere where, in contrast, B_ERA5 shows increasing trends.": As mentioned above, I do not see the picture that clear as the authors do. From the figures shown in this manuscript, I would be more careful with such statements.
- Line 516: "Our results show that although B_ERA5 shows better agreement with observed TCO than A_ERA1, they do not confirm that B_ERA5, based on the newer reanalyses, performs better in simulating stratospheric ozone overall.": Again, I think this broad statement is not backed by the findings of this paper.