Reply to the comments from the editor (Jens-Uwe Grooß):

Editor Review of the Manuscript "Stratospheric ozone trends and attribution over 1984-2020 using ordinary and regularised multivariate regression models" by Li et al.

As one of the reviewers of this manuscript did not commit a review and the other review was quite positive, I decided to base the decision of this manuscript on only one regular review and this editor review. Although I am not an expert on regression methods, I find the paper written clearly and understandable. Especially, the uncertainties of the derived ozone trends depending on the regression methods seem important to me. Also the depiction of the contribution of the natural processes to ozone changes is described well.

I would, however, suggest some more discussion of the results: To me it is not clear, in how far the shown differences in regression methods are now explaining the discrepancy in the lower stratosphere that was first pointed out by Ball et al. (2020). Besides the variability induced by the regression method, is there a model improvement with respect to the Multi-model-mean shown by Ball et al.? Or is this only the difference between free running CCMs and the CTM shown here. What can be learned from the machine-learning results (ML-TOMCAT)? Does the similarity with SWOOSH suggest that the basic mechanisms are well understood or would you expect this similarity as it is constructed by machine-learning using the observations? Why are the trends in the tropics so different between the two re-analyses? Is this due to the vertical velocities?

Therefore I suggest minor revisions to include this discussion, that would potentially bring the shown results better into the context of the present literature.

Reply: We thank the editor for his useful comments and suggestions about more discussion of the results, which have helped to improve the manuscript. The editor's comments are given below in black text, followed by our point-by-point responses in blue text.

(1) how far the shown differences in regression methods are now explaining the discrepancy in the lower stratosphere that was first pointed out by Ball et al. (2020).

Reply: Indeed, regression model methodologies do play some role in the trend estimation. Ball et al. (2020) used dynamical linear regression model (DLM) that attempts to determine time varying trends. Basically, DLM takes into account the temporal relationship between the dependent and independent variables, whereas OLS-type models assume that temporal relationship between dependent and independent variables is not important. So, to some extent somewhat larger negative trends shown in the study of Ball et al. (2020) are most probably due to the regression methodology adopted in their study. On the other hand, even with OLS/Ridge regression, models used here also show negative (though smaller in magnitude) in the lower stratosphere and exact causes for those trends are still not well understood. Although many recent studies (e.g., Chipperfield et al., 2018, Dietermüller et al., 2021, Li et al., 2022) attribute negative ozone trends in the lower stratosphere to dynamical changes, usage of reanalysis forcings (ERAI, ERA5, MERRA) are also not consistent, adding uncertainty in our understanding about the dynamical changes (e.g. Davis et al., 2023).

We have added some discussion about the differences in regression methods and comparison with the previous results (e.g. Ball et al., 2020) about the lower stratospheric ozone trends in the revised manuscript: "Negative trends with larger uncertainties are observed in the lower stratosphere, which are most pronounced in the tropics (-6.1±12.0 % per decade at 100 hPa), followed by the decrease at NH mid-latitudes (-1.6±3.2 % per decade at 100 hPa). The largest difference between OLS and Ridge regression methods occurs in the tropical lowermost stratosphere with a difference of ~9 % per decade at 100 hPa (but with larger uncertainties >10 % per decade for both regression methods), followed by the NH mid-latitudes with >2 % per decade difference at 100 hPa (~3 % per decade uncertainties). Note that, despite the large differences between OLS and Ridge-based trends, they are still within the uncertainties of the individual trends. The observed ozone decreases in the lower stratosphere are similar to recent records (e.g. Ball et al., 2019; 2020; Godin-Beekmann et al., 2022), which could be explained by the increased tropical upwelling and mid-latitude mixing (Wargan et al., 2018; Ball et al., 2020; Orbe et al., 2020; Davis et al., 2023). Nevertheless, the modelled lower stratospheric trends do not match those derived from observations. " (Lines 295-304)

"Compared to the trend estimates from simulation ERA5 in Figure 3, the ML-TOMCAT data set shows more consistent results with the SWOOSH data, with negative ozone trends in the tropical and NH mid-latitude lower stratosphere. Largest differences between SWOOSH- and ML-TOMCAT-based ozone trends appear in the SH mid-latitude lower stratosphere where ML-TOMCAT shows positive trends, and in the tropical mid- and lower stratosphere with close to zero trends near 60 hPa (although these trends have large uncertainties). On the other hand, trends from model simulation ERA5 show largest inconsistencies with respect to SWOOSH-based trends in the lower stratosphere. Simulation ERA5 shows positive trends for all three latitude bands that are more pronounced in the SH mid-latitudes (5.4±2.0 % per decade at 100 hPa for Ridge regression). These differences between satellite-based datasets and model simulation suggest there are still large uncertainties in the lower stratosphere where dynamical processes dominate (Dietmüller et al., 2021; Li et al., 2022). Ball et al. (2020) reported significant discrepancies in observation-model lower stratospheric ozone trends by using various satellite-based data sets and chemistry–climate models (CCMs). Although the inconsistencies vary with various datasets and fit methods (Dietmüller et al., 2021; Bognar et al., 2022), models generally do not reproduce the observations and the reason remains an open question." (Lines 305-319)

(2) Besides the variability induced by the regression method, is there a model improvement with respect to the Multi-model-mean (MMM) shown by Ball et al.? Or is this only the difference between free running CCMs and the CTM shown here.

Reply: The TOMCAT 3-D off-line chemical transport model (CTM) shown here is forced with meteorological fields from ECMWF ERA5 reanalyses (Hersbach et al., 2020) with a coherent historical assimilation of observations and full stratospheric chemistry scheme to reproduce the behaviour of ozone as closely as possible.

Objectively, there is no model improvement with respect to the Multi-model-mean (MMM) shown by Ball et al. (2020). The inconsistencies in observation-model lower stratospheric ozone trends shown in this study show some differences with those in previous study (Ball et al., 2020), which results from the difference between free-running CCMs and the CTM shown here. As replied above, we have added some discussion about the discrepancy of the ozone trends in the lower stratosphere.

(3) What can be learned from the machine-learning results (ML-TOMCAT)? Does the similarity with SWOOSH suggest that the basic mechanisms are well understood or would you expect this similarity as it is constructed by machine-learning using the observations?

Reply: The ML-TOMCAT data set we use here is a long-term chemically (and dynamically) consistent, satellite-data-based global gap-free stratospheric ozone profile data generated by applying a supervised machine-learning (ML) algorithm to the random-forest (RF) regression analysis (Dhomse et al., 2021).

The similarity or better agreement with SWOOSH is not surprising, as also mentioned by the reviewer (Dr Mark Weber), since satellite corrections used for ML-TOMCAT are derived from the same MLS data which are also part of SWOOSH, i.e. using 20 years of UARS-MLS (1991–1998) and AURA-MLS (2005–2016) measurements as a training period. However, it is also important that for the non-MLS time period, SWOOSH relies on limited (~30 profiles per day) observations from SAGE II and HALOE solar occultation instruments. So, monthly zonal means in SWOOSH data would have a limited set of observations but ML-TOMCAT would have means from all the model grid points. Dhomse et al. (2021) have demonstrated that ML-TOMCAT ozone concentrations are well within uncertainties of the observational data sets at almost all stratospheric levels, and there are significant improvements compared to the TOMCAT 3-D chemical transport model. Here we aim to illustrate that even with a limited set of denser observations to construct machine-learning based data, it still shows remarkable consistency with purely satellite measurement-based data in terms of ozone trends.

We have also added some sentences and comments about the similarity/agreement of the ML-TOMCAT data set with SWOOSH in the revised manuscript, e.g. "Comparing the ML-TOMCAT-based trend estimates with the ERA5-forced model simulation, we find ML-TOMCAT shows significant improvements with much better consistency with the SWOOSH data set, despite the ML-TOMCAT training period overlapping with SWOOSH only for the Microwave Limb Sounder (MLS) measurement period." (abstract, Lines 44-46) & "The better agreement between ML-TOMCAT and SWOOSH, due to satellite corrections derived from the same MLS measurements, show some improvements in this machine-learning based data set compared to TOMCAT chemical transport model." (Lines 306-308).

(4) Why are the trends in the tropics so different between the two re-analyses? Is this due to the vertical velocities?

Reply: Due to the differences existed between ERA5 and ERA-Interim reanalyses (e.g. vertical and horizontal resolutions, radiative transfer models and measurements assimilated, and changes in number and type of observations adopted), the differences of the trends in the tropics between the two re-analyses can be attributed to many reasons, including the different vertical velocities.

A detailed comparison between the model simulations forced by two re-analyses (ERAI and ERA5), including the difference of the stratospheric ozone profiles trends, has been reported in Li et al. (2022). From the discussion about the differences in age-of-air (AoA) tracer between two simulations, there exist some fundamental differences in the representation of Brewer-Dobson circulations between two reanalysis data sets.

As for TOMCAT setup, simulation ERA5 shows improvements in the TCO biases in the tropics compared to simulation ERAI. A possible explanation is that the finer vertical resolution in ERA5 alters vertical transport pathways that are critical for controlling ozone concentration as, within a few kilometres in the stratosphere, the ozone lifetime changes from days to a few years. Besides, simulation ERA5 shows increasing AoA trends in the whole stratosphere, while simulation ERAI shows a hemispheric dipole trend pattern with increasing AoA in the NH and decreasing trend in the SH lower stratosphere.

As differences between TOMCAT simulation forced with ERAI and ERA5 are already discussed in Li et al., (2022), the reviewer (Dr Mark Weber) suggested to omit results and discussion of ERAI. Hence, all the related comparison between the two model simulations has been removed from the revised manuscript.

References:

Ball, W. T., Alsing, J., Staehelin, J., Davis, S. M., Froidevaux, L., and Peter, T.: Stratospheric ozone trends for 1985–2018: sensitivity to recent large variability, Atmos. Chem. Phys., 19, 12731–12748, https://doi.org/10.5194/acp-19-12731-2019, 2019.

Ball, W. T., Chiodo, G., Abalos, M., Alsing, J., and Stenke, A.: Inconsistencies between chemistry–climate models and observed lower stratospheric ozone trends since 1998, Atmos. Chem. Phys., 20, 9737–9752, https://doi.org/10.5194/acp-20-9737-2020, 2020.

Bognar, K., Tegtmeier, S., Bourassa, A., Roth, C., Warnock, T., Zawada, D., and Degenstein, D.: Stratospheric ozone trends for 1984–2021 in the SAGE II–OSIRIS–SAGE III/ISS composite dataset, Atmos. Chem. Phys., 22, 9553–9569, https://doi.org/10.5194/acp-22-9553-2022, 2022.

Chipperfield, M. P., Dhomse, S., Hossaini, R., Feng, W., Santee, M. L., Weber, M., Burrows, J. P., Wild, J. D., Loyola, D., and Coldewey-Egbers, M.: On the cause of recent variations in lower stratospheric ozone, Geophys. Res. Lett., 45, 5718–5726, https://doi.org/10.1029/2018GL078071, 2018.

Davis, S. M., Davis, N., Portmann, R. W., Ray, E., and Rosenlof, K.: The role of tropical upwelling in explaining discrepancies between recent modeled and observed lower-stratospheric ozone trends, Atmos. Chem. Phys., 23, 3347–3361, https://doi.org/10.5194/acp-23-3347-2023, 2023.

Dhomse, S. S., Arosio, C., Feng, W., Rozanov, A., Weber, M., and Chipperfield, M. P.: ML-TOMCAT: machine-learning-based satellite-corrected global stratospheric ozone profile data set from a chemical transport model, Earth Syst. Sci. Data, 13, 5711–5729,https://doi.org/10.5194/essd-13-5711-2021, 2021.

Dietmüller, S., Garny, H., Eichinger, R., and Ball, W. T.: Analysis of recent lower-stratospheric ozone trends in chemistry climate models, Atmos. Chem. Phys., 21, 6811–6837, https://doi.org/10.5194/acp-21-6811-2021, 2021.

Godin-Beekmann, S., Azouz, N., Sofieva, V. F., Hubert, D., Petropavlovskikh, I., Effertz, P., Ancellet, G., Degenstein, D. A., Zawada, D., Froidevaux, L., Frith, S., Wild, J., Davis, S., Steinbrecht, W., Leblanc, T., Querel, R., Tourpali, K., Damadeo, R., MaillardBarras, E., Stübi, R., Vigouroux, C., Arosio, C., Nedoluha, G., Boyd, I., Van Malderen, R., Mahieu, E., Smale, D., and Sussmann, R.: Updated trends of the stratospheric ozone vertical distribution in the 60° S–60° N latitude range based on the LOTUS regression model , Atmos. Chem. Phys., 22, 11657–11673, https://doi.org/10.5194/acp-22-11657-2022, 2022.

Li, Y., Dhomse, S. S., Chipperfield, M. P., Feng, W., Chrysanthou, A., Xia, Y., and Guo, D.: Effects of reanalysis forcing fields on ozone trends and age of air from a chemical transport model, Atmos. Chem. Phys., 22, 10635–10656, https://doi.org/10.5194/acp-22-10635-2022, 2022.

Orbe, C., Wargan, K., Pawson, S., and Oman, L. D.: Mechanisms Linked to Recent Ozone Decreases in the Northern Hemisphere Lower Stratosphere, J. Geophys. Res-Atmos., https://doi.org/10.1029/2019JD031631, 2020.

Wargan, K., Orbe, C., Pawson, S., Ziemke, J. R., Oman, L. D., Olsen, M. A., Coy, L., and Emma Knowland, K.: Recent decline in extratropical lower stratospheric ozone attributed to circulation changes, Geophys. Res. Lett., 45, 5166–5176, https://doi.org/10.1029/2018GL077406, 2018.

WMO: Scientific Assessment of Ozone Depletion: 2022, GAW Report No. 278, ISBN 978-9914-733-97-6, 2022.