Dear reviewer,
many thanks for your helpful comments!

Important note

There is a major change unrelated to the reviewer comments: To be able to resolve a comment of the editor, we had to change the method of obtaining the fit coefficients. The editor asked for error bars in Table 2 (Table 1 in the revised manuscript). Unfortunately, it was mathematically not obvious how to obtain these values with the method used in the original manuscript. For this reason, we introduced 2 major changes to the method:

1. The fit is now based on the non-accumulated ozone time series (i.e., changes per day) and not on the accumulated time series.

2. Instead of fitting the parameters for each individual year and then averaging the fit parameters over the years, we now concatenate the time series of all years before the fit to obtain a single fit parameter.

These changes prompted the following changes to the manuscript:

1. A complete rewrite of section 3 to reflect the changes in the method.

2. Addition of new Figures 2 and 4 (figure numbers from new manuscript) and deletion of old Figures 2 and 3 (figure numbers from original manuscript).

3. The contents of all figures and the numbers given in the text (fit parameters etc.) have changed sligthly throughout the paper.

The fit parameters obtained by the new method are very similar to the fit parameters of the old method. That means that the results do not change qualitatively and that the conclusions remain the same.

The new method is more elegant and gets rid of some (partly arbitrary) assumptions of the old method. However, a disadvantage is that the actual fit does not "look" as clear and intuitive in new Figure 2 as this was the case with old Figures 2 and 3.

Note on dates for vortex formation and breakup

During the preparation of the revised version, we noticed some slight inconsistencies in the definition of the vortex formation and breakup dates. A few dates were not consistent with the 15 million km^2 criterion (see "tracked changes" version of Table 3), and the validation in Figure 6 and 7 of the revised manuscript did not use exactly the same dates as the fit. This has been corrected. In addition, we changed the vortex formation date for the southern hemisphere from 1 May to 15 May to exclude some time periods with a weak vortex.

Note on plots

For technical reasons, we had to change the software used to create the plots. That means that colors, font sizes, axis tick marks etc. may have changed.

General comments

- *I am a bit surprised that the models implemented Polar SWIFT (e.g., ECHAM6, AFES4.1, ICON-NWP, ATLAS) do not have a usable scheme for tracer transport and mixing in the stratosphere. Is it because the model top is too low to resolve the stratosphere, or their scheme is so inefficient that adding four tracers will make the model unusable? I suggest the authors provide enough information about this, for example the cost of carrying four additional tracers. Otherwise, their motivations of this work are not clear to me.*

  We have changed the statements in the abstract to reflect our motivation in a better way. The original formulation may have been somewhat misleading. All of the models mentioned in the introduction have a tracer transport scheme, and we did not want to imply that these schemes can't be used. It is perhaps better to say that our scheme can be used as an alternative to the schemes for tracer transport and mixing that usually exist in GCMs, and that our scheme may have benefits.

  The initial idea of Polar SWIFT was to develop a fast and self-contained module to determine polar ozone depletion, with the aim of an easy and straightforward coupling of this module to a GCM. The concept of parameterizing the transport was our first approach, because it kept the technical interface between SWIFT and the GCM very simple. We have added this as an additional motivation to the abstract.

  There may be better methods of simulating the transport of stratospheric ozone than our parameterization and these methods are successfully used in existing models (see e.g. the models and validation in Dietmüller et al, doi:10.5194/acp-18-6699-2018). However, since our transport parameterization is fitted to reanalysis data based on measurements, it may actually perform better than the transport scheme in an existing GCM, which may e.g. suffer from deficiencies in the gravity wave parameterization that influence the Brewer-Dobson circulation in the model. This was one motivation for our parameterization that we state now more clearly in the abstract.

  For instance, we implemented tracer transport for SWIFT in ECHAM6. ECHAM6 (and also the AFES GCM) is a hydrostatic model and the tracer transport is based on a Lin-Rood scheme (Giorgetta et al., 2014, https://mpimet.mpg.de /fileadmin /publikationen /Reports /WEB_BzE_135.pdf). We tested Polar SWIFT in ECHAM6 with tracer transport and found that the tracer transport of ECHAM6 overestimated the ozone concentrations inside the vortex, especially in the southern polar vortex. The results obtained by the transport parameterization actually were an improvement over the version with tracer transport. A reason for the bad performance of the tracer transport may be the overestimation of horizontal transport, which is a known issue in ECHAM6 (Stevens et al. 2013, doi:10.1002/jame.20015). A GCM with a more advanced tracer transport scheme (e.g. ICON) and improved vertical wave propagation will certainly compensate for some of these deficiencies.

The computational cost of adding more tracers to the GCM was not a serious issue. While the running time increased somewhat, this was not the main bottleneck in the computation.

Temperature biases in the GCM might influence the transport parameterization via the temperature dependent term. This issue and its solution are addressed in the SWIFT coupling paper, which is currently in preparation.

We do not expect that the model lid poses an issue. The GCMs that Polar SWIFT was coupled to have a model top at 1 Pa, which covers the domain of the Brewer-Dobson circulation.

Note that ATLAS is not a GCM and therefore was not included in our statement that there is no usable tracer scheme in some GCMs. ATLAS actually has a dedicated scheme for tracer transport in the stratosphere, and this scheme has been used in Wohltmann et al. (2017) for the validation runs of the chemistry part of the SWIFT model.

- *How are the tracers handled outside the polar vortex in these models? Are they advected?*

  The ozone values outside the polar vortex are taken from the internal ozone climatology of the GCM, which varies with season. Tracers are not advected outside the polar vortex. There is no interpolation applied between the two domains, since the edge of the polar vortex often forms a strong barrier between air masses and strong gradients in species concentrations are common.

  We have added discussion along these lines in an additional short section on the GCM implementation in the introduction.

- *When more efficient tracer advection scheme is available, GCMs do not need the transport parameterization. I am not familiar with the models mentioned in this manuscript, but the GCMs I have experience with all have tracers (trace gases, aerosols, and artificial tracers) transported and mixed in the stratosphere. A couple of years ago, one additional tracer roughly adds 1% overhead computational cost to the atmospheric model I used. It was a little slow, but still usable if one can keep the tracer number relatively small. Recently with the latest breakthrough of more efficient tracer advection scheme (e.g., Bradley et al., 2019; 2021), the MPI communications can be reduced by 9x in the model I use. Adding $\approx$30 tracers only costs $\approx$15% overhead. So, the tracer advection is not the primary bottleneck anymore (at least for some GCMs). If possible, I* *would suggest the models adopt more efficient advection schemes, which provides a better solution in the long term. The method here may be used as a temporary fix when needed.*

We agree that tracer transport is not necessarily a computational bottleneck for GCM simulations. The computational cost of tracer transport was not our main motivation to develop the transport parameterization (see above). A main motivation for the transport parameterization was an easy-to-use and self-contained model. Another important motivation was to improve on the quality of the tracer transport in GCMs, which sometimes has deficiencies as described for ECHAM6 above, related e.g. to problems with the gravity wave parameterization that affect the Brewer-Dobson circulation. The fact that our model is based on a fit to reanalysis data might improve on this situation. We are also aiming to use tracer transport schemes in the future, but more validation work needs to be done.

Minor comments

- *L1, what does SWIFT stand for?*

  In the original manuscript (Rex et al., 2014, doi:10.5194/acp-14-6545-2014), which introduced SWIFT as a "proof-of-concept" model, SWIFT stood for "Semi-empirical Weighted Iterative Fit Technique". However, the method used for the fit in the proof-of-concept version was already replaced by a different method in the first operational SWIFT version (Wohltmann et al., 2017). Therefore, we would suggest not to consider SWIFT as an acronym, since we feel spelling out the original acronym would only cause confusion.

- *L37-38, why these 5 levels? Is that where the polar vortex is simulated? Some clarifications are helpful for the readers to understand the choice.*

  This has historical reasons. The levels roughly encompass the vertical range in which ozone depletion is observed (see Wohltmann et al., 2017). The choice of the pressure levels was guided by the pressure levels of the ECHAM (EMAC) model in this altitude range, which was the first model in which Polar SWIFT was implemented (see also Wohltmann et al., 2017). We have added some short discussion and Wohltmann et al. (2017) as a reference to section 2.1.

- *L49, do you mean a single number of ozone change rate from Polar SWIFT is used for a given layer? It is not clear if "the rate of change of ozone calculated by Polar SWIFT" is a single number or different for the grid boxes in the same layer. I feel it's a single number as SWIFT calculates the vortex-averaged value, but the sentence is a bit unclear. Please clarify.*

  The rate of change is a single number. We have added "vortex-averaged" to the sentence to make this more clear.

The sentence continues to read: "...for a given layer to the ozone value of every air parcel inside the vortex and inside this layer. Note that this means that the ozone field does still vary across the vortex.". We hope this clarifies that the ozone field itself (compared to the rate of change) is not a single number, but different for every air parcel. Note that ATLAS is a Langrangian model and not a Eulerian model. There are no grid boxes, but individual air parcels (locations of trajectories).

We have also added some explanation that the initialization and the transport can lead to variability in the ozone field itself.

- *Figure 6b: years 2010-2012 show a large increase in ozone observation but totally missed by the models. Do you know why?*

  This is a good and interesting question. Looking at the time series of the ozone holes in the past, these winters do not seem to stand out (e.g. Bodeker and Kremser, doi:10.5194/acp-21-5289-2021). While this is interesting, we think a thorough investigation of the interannual variation of ozone in the southern hemispheric vortex and why this is not reflected in our model is outside the scope of this technical paper. May be part of the answer is that we are looking at values only from a particular day and a particular level, which may increase the variability of the measurements (and add some uncertainty to the comparison).

- *L227-230, the authors listed almost all the possible causes of why the model failed to capture the mean ozone at southern hemisphere as observed. This is not very helpful. Can the contributions from these potential factors be somewhat quantified?*

  It is very hard to disentangle these effects, because they add up in the end results, without the possibility of a clear attribution. E.g., a major problem here is that it is not clear how well the transport in ECMWF compares to reality. There are ways to investigate this further (comparing to conserved tracers like N2O in runs of the full ATLAS model, for instance, or comparing the results of the chemistry of Polar SWIFT to the full chemistry model), but we think a thorough investigation of this is a lot of work and is probably outside the scope of this technical paper. We know from ATLAS run that in some circumstances, tracers like N2O or CH4 cannot be reproduced as nicely as we would like it.

  Figure 16 of Wohltmann et al. (2017) shows an ozone bias at 46 hPa for the mean ozone values for a Polar SWIFT run with the full transport scheme of ATLAS that is similar to the bias observed at 41.6 hPa for run with the transport parameterization (Figure S49 of the supplement). This points into the direction that the transport parameterization is not a likely cause for the differences. We have added discussion along these lines to the manuscript.