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 (see also comment to lines 132–133 that is resolved by the new 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

- *Although the description of the previous implementation of Polar SWIFT, the one requiring additional tracers advected by the model, and how this differs from the revised version would be helpful.*

  There is no previous implementation of SWIFT into a GCM that does not use the transport parameterization, and hence, this is not a revised version, but the original implementation.

  The confusion may arise since it was not discussed in Wohltmann et al. (2017) how the the transport would be handled in an actual implementation in a GCM. This was an obvious omission, and part of the motivation of this manuscript was to rectify this omission. At the time of writing of Wohltmann et al. (2017), this omission was deliberate, since we were still in the process of implementing SWIFT into the GCMs at that time.

  We have now added text to the introduction describing how SWIFT is implemented into a GCM (see reply to your next comment). This may also help with this comment.

  We were required by the journal to put a version number into the title. Since the transport parameterization has not really a version number, we chose the same version number used for Polar SWIFT in Wohltmann et al. (2017), since the transport parameterization has always been part of the complete Polar SWIFT model. We hope that has not caused too much confusion, but we are not allowed to remove this number from the title.

  It may also have caused confusion that it was not mentioned that $ClONO_2$, HCl and HNO_3 are not included in the transport parameterization. Tests with a transport parameterization for these species showed that the changes induced for O_3 were small. Only O_3 is returned to the GCM as input for the radiation module. We have added discussion along these lines to the manuscript.

  We have change "evolution" to "chemical evolution" in line 19 to remove a potential source of confusion. Similar changes have been applied throughout the manuscript.

- *In particular, while there is a discussion of how Polar-SWIFT is implemented in ATLAS (Section 2.2), it is difficult to get an idea of how exactly Polar-SWIFT with the newly developed transport parameterization is implemented in a GCM.*

  We have added additional discussion to the introduction to give a general idea about the parts of the implementation into a GCM that are common to all GCMs. However, we would like to keep this discussion short.

A discussion of the implementation into the GCMs was actually not our main aim. Since the implementation of SWIFT is different for every GCM, we felt it would be out of the scope of this manuscript to describe the implementation in detail here. We would like to refer to the individual papers which describe the implementation of SWIFT into the respective GCMs. Unfortunately, as stated in the abstract, the manuscripts for ECHAM, AFES and ICON-NWP are still in preparation, and the existing publication for ECHAM (Romanowsky et al.) does not contain a very detailed account on the model implementation.

A detailed dicussion will also involve the chemistry part of Polar SWIFT, which is why we think a more thorough discussion leads too far away from the focus of this paper (a paper with a scope similar to Wohltmann et al., 2017, might be more appropriate here).

- *What prognostic variables does the model require...*

  If we understand the comment correctly, you mainly refer to the chemical part of the model described in Wohltmann et al. (2017) here. Since the chemical part is described in detail in Wohltmann et al. (2017), and this paper focuses on the transport part, we would like to leave the short description in lines 20–22 as is. This description is sufficient to understand the following and more detail feels out of the scope of this manuscript.

  As stated above, it may have caused confusion that $ClONO_2$, HCl and HNO_3 are not included in the transport parameterization. We now state that in the manuscript.

  ...and how does this differ from the previous version that required the explicit calculation of transport.

  There is no previous version, see above.

- *One other, overarching question I am left with is how does the GCM specify the concentration of ozone outside of the polar vortex when using this version of the Polar SWIFT ozone parameterization? This should be described here.*

  Outside of the polar vortex, the values of the internal ozone climatology of the GCM, which can vary with season, are used as input for the radiation module. 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 paragraph on the GCM implementation to the introduction.

Specific comments

- *Lines 4–5: "Many GCMs do not include a usable general scheme for the transport and mixing of chemical species in the stratosphere." As many*

GCMs and Earth System Models now contain a prognostic treatment of aerosols, I would think a serviceable transport scheme would be more generally available. Is it perhaps more true that many GCMs do not specify a high-enough model lid and a sufficient number of model levels in the stratosphere to adequately resolve the dynamics of the stratosphere?

We have changed the statements in the abstract to reflect our motivation in a better way. The original formulation may have been somewhat misleading. 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.

Isn't this point also a little beside the point because the parameterization being presented here is for the vortex averaged transport effect and would not be the kind of quantity readily calculated by a 3-D advection scheme in a GCM?

As can be seen by the implementation of SWIFT in the ATLAS-SWIFT model, there is a straightforward way to implement full 3D transport into a GCM that uses the chemistry part of Polar SWIFT, although SWIFT only calculates vortex averages (see description in lines 48–52 of the original manuscript and Wohltmann et al., 2017). But maybe this is not the point you are aiming at?

The transport parameterization presented here may only be of benefit when a full 3D transport scheme is not available or when there are deficiencies in the representation of transport in the full 3D scheme of the GCM. 3D transport from the GCM does not necessarily perform better than the parameterization, since transport from the GCM could have deficiencies that are not present in the transport parameterization, which is based on reanalysis data. There may be cases where a parameterization of the vortex-averaged transport effect may only be a poor replacement for a full 3D transport scheme.

- *Lines 23–34: The section beginning "When using the Polar SWIFT model. . . " is actually describing the present work but it is a bit disorienting to the reader because it appears in the introduction and really does not describe previous versions of Polar SWIFT or the motivation for the current work.*

  We rephrased the abstract to make our motivation more clear (see comment lines 4–5). A benefit of our approach includes the easy and self-contained coupling to a GCM. Another advantage can be that a transport parameterization based on reanalysis data and measurements can avoid deficiencies in the representation of transport in the GCMs.

  There is no previous version of Polar SWIFT without the transport parameterization, see above in "general comments". The chemistry part of Polar SWIFT is only described in a few sentences, since a detailed description can be found in Wohltmann et al. (2017) and this paper focuses on the transport. We hope that the description is sufficient to put things into context.

- *Lines 65–76: It took me a couple of readings to understand what is going on – just too many variations of 'SWIFT's. Maybe an introductory sentence around Line 67 would help, stating that the transport parameterization is derived from an analysis of the total and chemical tendencies of ozone from a simulation of the ATLAS-SWIFT model?*

The complete section 3 was rewritten to consider a comment of the editor (see above). Please check if it reads better now and is more clear to the reader.

- *Lines 132–133: "For the temperature variable ΔT_{fit}, we use the vortex-averaged temperature difference in a layer at a given date compared to the start date (vortex formation date, see Table 1)". Since the ΔT_{fit} term also involves subtraction of the estimate of the time-evolving radiative equilibrium temperature, as discussed a bit later, what is the reason for including the (somewhat arbitrary) temperature at the vortex formation date? Given the radiative relaxation timescale, the temperature at the vortex formation date I think would become irrelevant as an estimate of dynamical forcing with increasing time since the start date. And, as pointed out at lines 133–136 "Equation 2 suggests that the difference in temperature to the start date roughly corresponds to the deviation of the temperature on this day from the radiative equilibrium temperature by the effects of the Brewer-Dobson circulation. According to Equation 2, this would be exactly true when the temperature at the start date would be the radiative equilibrium temperature." And, of course, there is no guarantee that the temperature at the start date will be the radiative equilibrium temperature.*

  This has been largely resolved by the new method for obtaining the fit, which does not need the temperature difference to the vortex formation date anymore. Now, only the differences from day to day are used.

  We also have added the following sentence to the description of Equation 4 (previously Equation 2 in the original manuscript): "On short time scales, the change in temperature is directly correlated to the corresponding downwelling in that time period ($\overline{T}_{\mathrm{R}}$ terms cancel out)."

  We agree that temperature will lose its memory to transport effects on time scales longer than the radiative relaxation time scale and that this could affect the quality of the results of the temperature-dependent term for long-term changes.

  We added the following discussion to the end of 3.3: "As a note of caution we have to stress here that while this method will work well for short-term changes in temperature and ozone, it might not work well for changes on a longer time scale. On longer time scales, temperature will start to lose memory of the transport in the past due to the radiative relaxation time scale of about 1 month." We have currently found no way to implement a method that will also correctly model the interannual variation of the very long-term changes in ozone.

- *Lines 221–223: "The differences between Polar SWIFT and MLS cannot be explained by the transport parameterization, since they are much larger than the differences of about 0.2 ppm between the transport parameterization and the transport term of ATLAS discussed in the last section 4.1." But in Figure 6, particularly for the Northern hemisphere the Polar SWIFT model without the transport parameterization (the blue line of*

Figure 6) does a good job of estimating ozone for the two cold years. Is the argument that there should always be some positive contribution from transport so that the chemistry-only simulation should be even lower than it is, particularly for cold years? Do the authors have any reason to believe the chemistry parameterization underestimates the amount of ozone chemical destruction?

This is a good point. This probably should be phrased more carefully. This statement implicitly relied on the assumption that the transport is represented well in the ECMWF ERA5 data and in the transport scheme of ATLAS, which, however, can't be guaranteed.

And actually, you are right that there is not really reason to believe from Wohltmann et al. (2017) that the chemistry parameterization underestimates the amount of chemical ozone destruction in cold winters (more than in warmer winters). There is just no clear indication from the results in this paper (compare Figure 15 from Wohltmann et al. (2017) to this study). The problem is also much less pronounced at other levels.

Changed the text in the manuscript to "The model overestimates ozone at 54 hPa by about 0.7–0.8 ppm compared to the MLS measurements in two winters with low ozone values (2010/2011 and 2019/2020). That is, in cold winters with large ozone depletion and a weak Brewer-Dobson circulation, while warmer years are simulated relatively well (however, the overestimation in cold winters is much less pronounced at the levels 3–5, see Figure S24 in the supplement). The differences in cold winters between Polar SWIFT and MLS at 54 hPa might not be explained by the transport parameterization alone, since they are much larger than the differences of about 0.2 ppm between the transport parameterization and the transport term of ATLAS discussed in Section 4.1 (however, this relies on the assumption that the transport is represented well in the ECMWF ERA5 data and in the transport scheme of ATLAS). Hence, the differences between Polar SWIFT and MLS in cold winters could also be a deficiency of the chemistry model of Polar SWIFT. However, there is no clear indication from Wohltmann et al. (2017) that this could be the case. A detailed discussion of the chemical model of Polar SWIFT is outside the scope of this paper."

We have to admit that it is tempting to just assume that there was almost no transport in cold years, which could explain the good agreement of Polar SWIFT without transport and MLS in these years. However, there are too many unknowns here to constrain things enough to come to a certain conclusion here.

A minor source of uncertainty will also be uncertainties in the MLS measurements, which we did not mention because they are probably small compared to other uncertainties.