Reply to Referee #1

This paper explores the effects of SF6 loss on the mean age values and trends derived from SF6 model output and observations. The conclusion, that SF6 loss can reverse the sign of the mean age trend over recent decades compared to mean age from a tracer without loss, is important in helping us understand the long running discrepancy between modeled and measured trends. The explanation is actually incredibly straightforward, that the higher the concentration of a species in the atmosphere, the more absolute loss will occur and therefore the bias will grow over time. It's kind of hard to believe none of us studying this topic hadn't thought of this possibility before, but there you go. Nice work by the authors of this paper to recognize the importance of SF6 loss and quantify the effect.

There are still outstanding issues, the trend in CO2 derived mean age is one of them and the SF6 lifetime derived in the EMAC model is another since it seems to be higher than other recent estimates. Nevertheless, this paper is an important step forward in our understanding of mean ages derived from observations and how models can be used to help put them in context. My main comments described below revolve around how best to use this information to help us make the observationally derived SF6 mean ages more accurate. It doesn't make sense to cast the measurements aside when it comes to mean age estimates so we need to be careful to frame the current understanding in a more inclusive way.

I do recommend this paper be published with consideration of the comments below.

Main comments

It would be really nice to see a plot of latitude vs. altitude trend differences between REF(WS, SF6) and REF(NS, SF6). I realize you've focused the trend analysis on the locations of the Engel et al. papers in the 4-6 year mean age range, and that's important to see. But at what mean age does the trend discrepancy emerge from the uncertainty to become significant? How does this vary with location? You hint at it in Figure 1 where the NS and WS profiles diverge but it would be good to see more. I think this is important for the community to know what age ranges we can reliably use SF6 as a mean age tracer. For instance, in my 2014 paper I showed mean age trends at four altitude ranges and in the bottom layer, where mean ages were 2-3 years, the SF6 measurement trends agreed within uncertainties with the modeled mean age trends. Is this expected based on your work?

Reply #1:

Thank you for your comment, and this is a very good suggestion. We included the relative difference of both the mean AoA climatology as well as the trends as latitude vs altitude plots. For the climatologies, we included the relative difference between AoA from REF(WS, SF6) and REF(NS, SF6) for two periods, namely 1970-1980 and 2000-2010. These two figures have been added to Figure 3. The caption has been amended to include:

"The relative difference between AoA(WS, SF6) and AoA(NS, SF6) from the reference simulation for the periods 1970-1980 and 2000-2010 are shown in panels g) and h), respectively. The black contours depict AoA(NS, SF6)REF for the respective time period. "

and the following text has been added to line 229:

"The relative effect of the sinks on AoA derived from the SF6 tracers with non-linear growth can be seen in Figs. 3g) and h). Mean AoA values derived from SF6 in the early period (1970-1980) are moderately affected by the sinks, with a difference of around 20-25% in the polar middle stratosphere and above (i.e., for mean AoA values above 5.5 years). Differences are small (less than 10%) for mean AoA values below 4 years. However, as will be discussed (see Section 3.5), the effect of the SF6 sinks increases over time and for the later period (2000-2010), mean AoA derived from SF6 is affected by the sinks considerably. Differences greater than 20% can be seen in Fig. 3h) for

AoA above about 3 years, and in the extratropical lower stratosphere, differences are larger than 10% even for mean AoA values of 2 years and above."

Furthermore, we plotted the latitude vs altitude trend differences between REF(WS, SF_6) and REF(NS, SF_6) and added this to the revised manuscript (Figure 5). The caption reads as:

"Relative difference between the trends of AoA(WS, SF_6) and AoA(NS, SF_6) over the time periods (a) 1965-2011 and (b) 2000-2010. The black contours depict the trend of AoA(NS, SF_6) over the respective years. Dotted regions indicate where the relative trend difference is not significant, i.e. the statistical significance lies below the 5% threshold."

This figure indicates that the long-term trend (over 1965-2011) for SF_6-based AoA with sinks deviates beyond 100% from the long-term trend of SF_6-based AoA without sink consideration, which implies that SF_6-based AoA is unsuitable for AoA trend calculations in the whole stratosphere. These differences are generally less pronounced for the shorter period of 2000-2010, and the relative trend differences in AoA(WS, SF_6) and AoA(NS, SF_6) lie under 100% in the tropical and extratropical lower stratosphere.

This has been added as a paragraph at the end of Section 3.4, in line 294:

"As shown in Fig. 5, the strong deviation of SF_6-derived AoA trends holds almost throughout the stratosphere: Figure 5a) indicates that the long-term trend (over 1965-2011) for SF_6-based AoA with sinks deviates beyond 100% (i.e., by a reversal of sign) from the long-term trend of SF_6-based AoA without sink consideration, which in turn implies that SF_6-based AoA is unsuitable for AoA trend calculations in the whole stratosphere. These differences are generally less pronounced for the shorter period of 2000-2010 (see Fig. 5b), and the relative trend differences in AoA(WS, SF_6) and AoA(NS, SF_6) lie under 100% in the tropical and extratropical lower stratosphere."

As a corollary to your findings, couldn't a correction to SF_6 mixing ratios be made in the calculation of mean age to account for loss? This would be similar to the way CO_2 is adjusted for production from CH_4 oxidation before calculating mean age. This correction would have to vary with time and location, which you could derive from the differences between your REF(NS, SF_6) run and the MIPAS data.

Reply #2:

Thank you for your thoughts on this – while this topic will require more in-depth analysis than there is room for in this paper, and we aim to study this in more detail in a follow-up study, we have now extended the mathematical discussion in Section 3.5 (starting at line 329) to its own Section 4: "Theoretical considerations and concept for sink correction methods". In this section, first analyses of a possible corrective approach to account for SF_6 loss are put forward, and we conclude that it might be possible to apply a linear correction to AoA of up to 4 years.

As such, we have amended line 19 at the end of the Abstract to read as:

"We conclude the study with a first approach towards a correction to account for SF_6 loss and deduce that a linear correction might be applicable to values of AoA of up to 4 years."

The final paragraph has been extended in line 407 in the Conclusion and now reads as:

"Furthermore, we put forward a first approach towards a method for SF_6 loss correction, which in the future shall be further developed and applied on observational data. From our first analyses, we can conclude that a linear correction (that is dependent on both time and effective lifetime of SF_6) can likely be applied to AoA values up to 4 years. However, further studies with more comprehensive approaches are required for a precise quantification of these values."

Specific comments

Line 6: I would add 'in the NH extratropical middle stratosphere' after 'positive trend'. Or something similar to give some reference to the Ploeger et al. (2015) and Stiller et al. (2017) hemispheric shift effect on the age trends that you mention in the introduction. Also, in Ray et al. (2014) I showed that the balloon age trends are negative in the lower extratropical stratosphere.

Reply #3:

Thank you, we have incorporated your comments in the Abstract. The sentence we added reads:

"Satellite observations show much older air than climate models and whilemost models compute a clear decrease of AoA over the last decades, a thirty-year timeseries from measurements shows a statistically non-significant positive trend in the NH extratropical middle stratosphere."

We also amended the section in the Introduction to read as:

"… show a (statistically non-significant) positive trend (note that Ray et al. (2014) also shows negative balloon AoA trends in the lower extratropical stratosphere)."

Figure 1: It seems like you could do a better job of averaging model data to match the March 2000 balloon profile location. This balloon flight was clearly in the vortex as shown in Figs. 2 and 3 in Ray et al., 2017 with equivalent latitudes from 68-75N. By averaging model locations at all longitudes and latitudes as low as 60N you are mixing locations in and out of the vortex. It's relevant to the lifetime estimates to see how the polar vortex model and measurement profiles compare since this is the region most impacted by the SF6 loss.

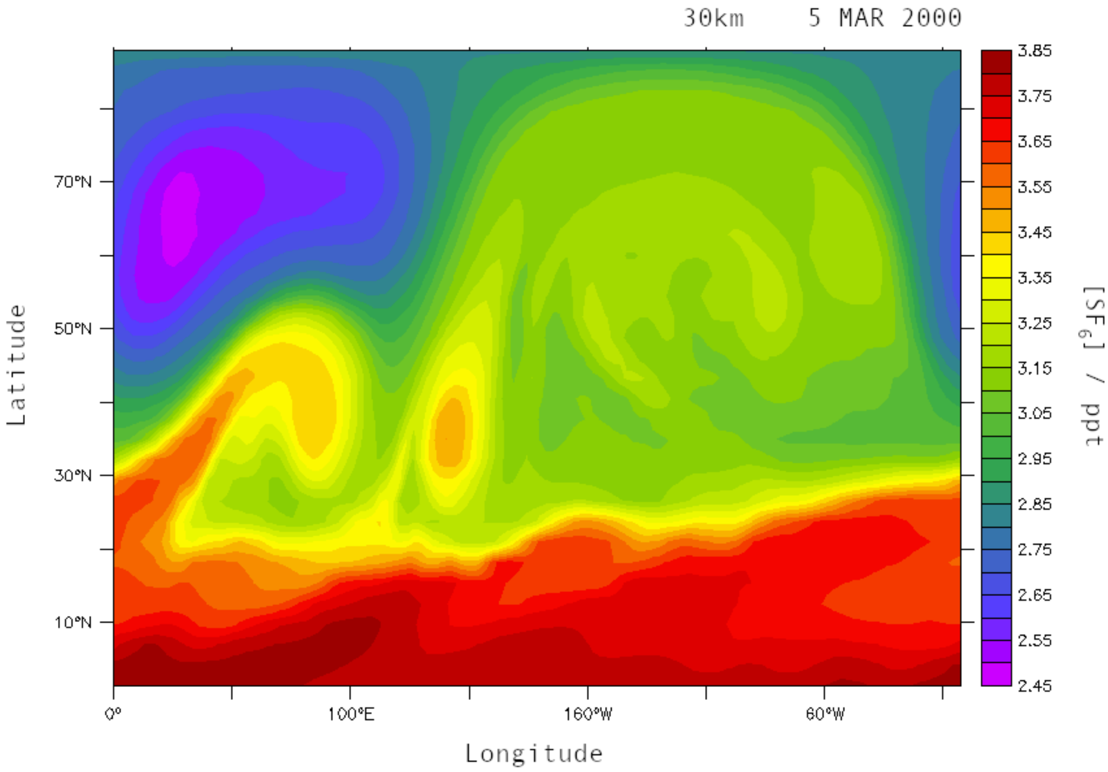
Reply #4:

Thank you for your constructive comment. We have examined the polar vortex in the SD simulation for 5th March 2000 and find that it is slightly displaced and centered over Northern Europe. To ensure that we show the SF6 profile based on air masses from within the vortex, we changed the latitude and longitude selection to 65N-80N and 0E-100E. We have attached a figure of SF6 as a function of latitude and longitude at 30km at the end of this reply (see below).

Figure 1 has been changed accordingly and updated in the revised manuscript.

Lines 171-173 have been updated to read as follows:

"To ensure that the SF6 profile is based on air masses from within the vortex, modelled SF6 values are averaged over 0E-100E and 65N-80N, which corresponds to the area of the vortex core for the given day. The standard deviation of SF6 for 0E-100E averaged over the respective latitude range is shown as error bars."



Line 130: 'approximately' is misspelled, change 'turned out to be' to 'are'
Line 199: 'definition of Braesicke…'
Figure S4: No y-axis labeling.
Figure S5: In the caption you state 'Values averaged over the region 30-50N and 2007-2010'. I think the 30-50N part is not needed here since these are latitude vs altitude plots.

Reply #5:
Thank you for pointing out these errors, we have incorporated your comments in the relevant sections.

Reply to Referee #2

General comments:

1. The introduction suggests that main objective of the study is to reveal the cause of discrepancy between modelled and observed AoA and their trends. If it is the case, it should be stated explicitly. The lack of clearly formulated objectives and research questions makes it difficult to understand e.g. the model experiment design and justification for the choice of specific setups for the model experiments.

Reply #1:
Thank you for your comment. The main objective of our study is to further the general understanding of the effect of SF6 sinks on the derivation of SF6 based AoA, namely the resulting discrepancies between modelled and observation-based AoA, and investigate the effect on the long-term trend. To make this clearer, we added the following to line 64:
"… with the aim to understand the effects of SF6 sinks on tracer derived AoA and its long-term trends. Specifically, we calculate for the first time the effect of the sinks on the long-term trend of AoA."

2. Further, the introduction states that "a comprehensive explanation for the trend differences between models and observations is still missing". The effect of the SF6 destruction on the apparent AoA has been already pointed by Waugh and Hall (2002, Sec 3.2) and addressed by Kouznetsov et al. (2020), who has simulated the effects of the mesospheric sink of SF6, and concluded that "The apparent over-ageing introduced by the sink is large and variable in space and time. Moreover, the over-ageing due to the sink increases as the atmospheric burden of SF6 grows". (For more details, please refer to Sec.6.3 of the latter paper.) Therefore, it should be clearly stated what makes the existing explanations non-comprehensive.

Reply #2:
Thank you for your comment. Kouznetsov et al. (2020) carried out simulations to study the SF6 sinks in a chemistry transport model, while our study uses a chemistry climate model to examine the effects on the long-term trend of AoA. We also provide a correction of the non-linear growth of SF6 in the calculation of AoA in our chemistry climate model.
As such, we have amended line 49, it now reads:
"… In a study based on a chemistry transport model, Kouznetsov et al. (2020) showed that changes in SF6-derived apparent AoA over one decade are highly influenced by the SF6 sink, and can even turn positive. However, a comprehensive understanding of what contribution the individual effects have on the AoA trend depending on altitude and latitude is still missing."
Furthermore, line 51 has been amended to read:
"… SF6 sinks lead to older apparent AoA (see e.g. Waugh and Hall (2002) and Kouznetsov et al. (2020)), …"
And line 66 has been extended to read:
"… We apply a correction for the non-linear growth of SF6 in the calculation of AoA, based on Fritsch et al. (2019), which allows a more quantitative view of the effect of the SF6 sinks."

3. The conclusions are formulated quite vaguely. It should be clearly stated what are the findings of the present paper, and how they agree/disagree with earlier results, and what of the findings are new. It might make sense to have separate "Discussion/summary" (all references to earlier results etc.) and "Conclusions" (the concise statements that the authors are ready to defend).

Reply #3:
Thank you for your comment. We extensively discuss the results in the various subsections of Section 3 and so we have decided on refraining from having a separate discussion section.
We have changed the Conclusion to state clearly that

- we show the effect of SF6 sinks on the AoA trend for the longer time series, not only for the short MIPAS time series which has large variability issues
- we rule out that other effects have an impact: we quantify these other effects and thus can say that the sinks do indeed have this big an effect on tracer derived AoA and its trends

For this, we have added to line 393:
"In this study we show the effect of SF6 sinks on the tracer derived AoA trend for both longer time series as well as for the short time series corresponding to the MIPAS time frame, whereby issues in variability are associated with the latter."
Line 401 has been amended to read as:
"Our sensitivity studies quantified that..."
Furthermore, as mentioned in Reply #2 and Reply #4, we have changed line 373 to highlight that EMAC is a chemistry climate model:
"… using the chemistry climate model EMAC…"
and line 374 has been amended to read:
"...This submodel allows for explicit calculation of SF6 sinks and we applied a correction for the non-linear growth of SF6 in the calculation of AoA (Fritsch et al., 2019)."

4. Modelling studies of long-term evolution of SF6 distribution in the atmosphere have been reported by e.g. Reddmann et al (2001), Kovacs et al. (2017), Kouznetsov et al. (2020). The need for the present study and its similarities and differences from earlier ones should be clearly indicated.

Reply #4:
Thank you for your comment. While the above mentioned modelling studies may bear some similarities, they also report a wide range of SF6 lifetimes, for example. The SF6 model used in this study is based on the reaction scheme of Reddmann et al. (2001).
We extended line 51 to read as:
"SF6 sinks lead to older apparent AoA (see, e.g. Waugh and Hall (2002) and Kouznetsov et al. (2020),..."
Line 56 has been extended to read:
"A more recent model study by Kovács et al. (2017), who used the Whole Atmosphere Community Climate Model (WACCM) to determine the atmospheric lifetime of SF6, reported a…"
Line 58 has been amended to:
"… Kouznetsov et al. (2020), who performed simulations of tracer transport with a chemical transport model, shows…"
Lines 372-374 have been changed to highlight the differences from earlier studies as well as clarify our aims:
"While previous studies (see, e.g. Kouznetsov et al. (2020)) showed that the chemical sinks can strongly influence SF6-derived AoA in terms of absolute values and decadal changes, we investigate for the first time how longer-term trends are affected in a consistent manner, and investigate the different contributions from circulation changes, changes in abundance of reaction partners, and trends induced by constant destruction rates. To make this step towards understanding the reasons for these discrepancies, we thus study ..."

Specific comments

Sec 2.1: A brief characteristic of the model is missing., e.g. "online spectral chemistry-climate model with hybrid sigma-pressure vertical layers".

Reply #5:

Thank you for your comment. To meet your point, we extended the beginning of Sect. 2.1 (EMAC model), it now reads:

"For this study, we use the EMAC (ECHAM MESSy Atmospheric Chemistry, v2.54.0, Jöckel et al., 2010; Jöckel et al., 2016) model, a numerical Chemistry and Climate Model (CCM) system. It contains the General Circulation Model (GCM) ECHAM5 (ECMWF Hamburg, Roeckner et al. 2003), with its spectral dynamical core, as well as the MESSy (Modular Earth Submodel System, Jöckel et al. 2005; Jöckel et al., 2010) submodel coupling interface. The latter is a modular interface structure for the standardised control of process-based modules (submodels) and their interconnections. We apply the model in a T42 ...."

Sec 2.2: The description of the SF6 sub-model is very unclear. Probably, most of the reactive species from Table 1 were not implemented as actual tracers in the model. One has to indicate which species were taken as climatology, which were forced from other models, and which were actual tracers. Was the submodel implemented as prescribed destruction rate as a function of altitude, latitude and season, or was it something more sophisticated? The description should be sufficiently detailed to allow for an independent reproduction of the experiment with another model.

Reply #6:

Thank you for this comment, we improved the description of the SF6 submodel now such that it is clear what had been done and the simulations are reproducible. For this, we reduced Table 1 to only those chemical reactions that are actually considered in the module (following Reddmann et al., 2001). As the text already states, the submodel is indeed more sophisticated than prescribed destruction rates, it considers the individual chemical reactions for SF6 chemistry. As stated in Sect. 2.2, for the reactive species of SF6, tracers can be used in EMAC, but as noted in line 97-98, we here prescribe the tracer fields from an EMAC simulation with activated comprehensive interactive chemistry.

We additionally added in line 97-98:

"... The reactant species involved in the SF6 chemistry (HCl, H, N2, O2, O(3P) and O3) and the radiatively ..."

And moreover, we add to Sect. 2.3

"To compute the photodetachment rate of SF6-, we follow Reddmann et al. (2001) using the extraterrestial solar photon flux with no attenuation of the UV-photon flux, as provided by WMO (1986)."

To make clear that the submodel is part of the official MESSy distribution for all users, we further add in line 81:

"... in the mesosphere, and the submodel is operationally available for all users in MESSy from version 2.54.0 onwards."

Contrary to stated in ll. 94-95, Fig.S1 does not show the relative importance of various reactions for SF6 destruction, but rather shows same reactions as in Table 1, but in a graphical form.

Reply #7:

Thank you for your correction. We have amended the sentence in the revised manuscript to read as:

"… For a general overview of the various reactions see Fig. S1 in the Supplementary Information."

Sec 2.3. This is by far not the first simulations of this kind. What are the similarities and differences from the setups used in earlier modelling studies? What is justification for specific model experiments, i.e. research questions to be addressed with each of the setups?

Reply #8:

Thank you for your comment. To meet your point, we have added to the beginning of Section 2.3, starting at line 97:

"The simulations performed in this study include a more comprehensive approach for the calculation of the SF6 sinks. We use a climate chemistry model (as opposed to studies based on chemistry transport models, see e.g. Kouznetsov et al. (2020)) and use a more comprehensive SF6 submodel than previous chemistry climate model studies (see, e.g. Marsh et al. (2013) for the Whole Atmosphere Community Climate Model (WACCM))."

Regarding the CSS sensitivity simulation, line 108 now reads as:

"...simulation (year 1950 on repeat). With this simulation we aim to address the influence of the reactant species involved in the SF6 sink reactions. …"

For the time slice simulation TS2000, we added to line 111:

"This will allow us to investigate the effects of the SF6 sinks under a constant climate."

Sec 3.1: The section has one comparison against 3-year-mean MIPAS profile for a latitude belt of 30N-50N, and one in-situ profile. Those are nice for illustrations, but are insufficient to judge on the model performance in reproducing SF6 distribution in the stratosphere.

Reply #9:

Thank you for your comment. We agree that it would be desirable to carry out a more in-depth comparison to observations. However, we focus on the general effect of SF6 sinks on AoA, and so our aim here is to ensure a realistic-enough representation of SF6 by the EMAC model. Further studies would be required for better tuning of these, whereas we illustrate that the EMAC SF6 profile agrees well with SF6 profiles that are readily available - in this case, data provided by Ray et al. (2017) and Stiller et al. (2020).

We have added this in the beginning of Section 3.1, where line 167 now reads as:

"...observational data. This study does not perform a detailed comparison of SF6 profiles, as the major aim is not an in-depth evaluation of the SF6 submodel, but rather a quantification of the potential effects of the SF6 sinks on AoA and its long-term trends. However, to ensure that SF6 values in the EMAC model are within the range of observational estimates, we perform selected comparisons to data from Ray et al. (2017) and MIPAS SF6 (Stiller et al., 2020, paper in preparation. Fig 1a …"

Fig.1a: It is not clear why this specific latitude belt, and these specific years were selected. The MIPAS error bars show standard deviation of individual MIPAS profiles. How those are related to the uncertainty of the average (of millions?) profiles that are shown? The MIPAS averaging kernel and spatio-temporal collocation notably affect the comparison (Kouznetsov et al. 2020). This issue has to be at least discussed.

Reply #10:

Thank you for this comment. We are interested in the mid-latitude region and chose the specific years of 2007-2010 as there is a complete dataset in this time span. To make this clear, we add in line 170:

"… zonal mean SF6 averaged over 30°N-50°N and 2007-2010. These years have been chosen as the dataset is complete in this period."

With regards to your comment on MIPAS: In our approach we checked whether the mean model profile could be part of the ensemble of MIPAS measurements. We have not compared the mean profiles of the measurements and model. For our purpose, we need the standard deviation of the zonal mean ensemble. This is the error bar that is plotted in Fig. 1a. The standard deviation of the observed zonal mean ensemble consists of the measurement noise error of the MIPAS measurement, further random error sources from the retrieval (e.g. temperature uncertainties), the natural variability over the longitudes of the latitude band, and the four years of averaging (2007 - 2010). The model is not in contradiction to the observations if the standard deviations overlap. Averaging kernels and the error covariances both describe the interdependence of the vertical profile grid points in the case of the measurements. Such an interdependence tends to smooth the profiles in the retrieval application chosen for MIPAS SF6. Since we expect rather smooth profiles anyhow, partly because of the nature of this trace gas, and partly due to averaging over a large number of profiles, further smoothing does not greatly change the resulting standard deviation. If we wanted to compare the mean profiles to each other, the correct approach would be to apply the averaging kernel of each single MIPAS profile to the model profile at the geolocation of the MIPAS measurement, and to average over the smoothed model profiles. This, however, was not our intention.
We now include this in the discussion of Fig. 1a, in line 170:
"The error bars represent the standard deviation of the zonal mean ensemble, which consists of the measurement noise error of the MIPAS measurement, further random error sources from the retrieval (e.g. temperature uncertainties), the natural variability over the longitudes of the latitude band, and the four years of averaging (2007 - 2010)."

Fig 2a. Interestingly, Kouznetsov et al. (2020, Fig 5 there) shows very similar offset of the model profiles with respect to the in-situ one. I wonder if it is a coincidence, or an indication of a similar issue with both model setups.

Reply #11:
Indeed, this is an interesting point! As there are many differences between the model types and setups, and also considerable uncertainty in the satellite derived SF6, it is hard to judge whether the similarity might stem from a similar issue.

Sec 3.2: The methodology for the lifetime estimate is quite unclear. Instead of explaining the method used, the authors put a reference to a 600-page report (Braesicke et al, 2019). The method, probably, refers to the equation in p. 1.20 of the report. The equation assumes well-mixed atmosphere and implies that the destruction of SF6 is proportional to its burden, which is not the case: the destruction does not depend on the tropospheric content of SF6, but rather on its content in mesosphere. In a situation when the change of SF6 burden is substantial at a time scale of ~10 years (AoA in the stratosphere) the difference leads to "surprising" results like reduction of SF6 lifetimes by 25% over 100 years. Given the slow destruction of SF6 one could still define the lifetime in terms of well-mixed assumption, but that would require a long-term simulation without emissions, to let the mixing ratio relax to its equilibrium distribution and get to the exponential decline of the total burden. Alternatively, as it was done by Kouznetsov et al. (2020), one could use a total burden that corresponds to the mixing ratio next to depletion layers.

Reply #12:
Thank you for this comment, we improved the description of the SF6 lifetime calculation such that it is clear what has been used. To make this clear, we add in line 198:
"… We calculate the lifetime following Equation 1 in Section 3 of Reddmann et al. (2001), namely: [see Manuscript]. It is based on the reaction rates (ki) of the chemical reactions (Ri) marked in Table

1, the branching fraction e (taken as 0.999, see Reddmann et al. (2001)), and the efficiency of the SF6-recovery reactions (n), where n is calculated as: [see Manuscript]."

Sec 3.4 -- 3.5: Same behaviour of trends in the apparent SF6 AoA has been pointed out by Waugh and Hall (2002), Waugh et al.(2003) and demonstrated with extensive model simulations (Kouznetsov et al., 2020) for various latitudes and altitudes. Please specify what is a new finding here with respect to those studies.

Reply #13:

Thank you for your comment. As addressed in Replies #2, #4, #8. Kouznetsov et al. (2020) uses a chemisty transport model and considers the AoA trend only over 10 years, whereas we look at long-term trends and use a general circulation model with atmospheric chemistry. While our results bear similarities to the Whole Atmosphere Community Climate Model (WACCM) , see e.g. Marsh et al. 2013, they do not use a precise correction for the non-linearity of SF6 emissions and are therefore not in line with Engel et al. (2017) nor with Fritsch et al. (2019). We have, for example, amended the Conclusion (see Reply #4) and Section 2.3 (see Reply #8) to clearly state these differences.

l.329-365: The fact that SF6 destruction causes the positive trend in the apparent AoA follows from the simple fact that the SF6-AoA is proportional to a difference between stratospheric and tropospheric SF6 mixing ratios. Since the destruction is proportional to the SF6 mixing ratio, the difference increases together with the increase of the atmospheric SF6 burden. There is no need to involve any equations or advanced concepts (like Green functions etc.) to explain that.

Reply #14:

Thank you for your comment. We find the additional mathematical presentation of the effect of SF6 sinks on AoA could be of interest for the reader. Furthermore, as discussed with Referee #1 (see Main Comment #2), we extended Section 3.5 to now include an examination of the possibilities to include a correction for the SF6 sinks.