Review of the Manuscript

Overall Comments

Overall I think that this is an interesting and well-written paper which is very suited to ACP. The paper describes a new method for estimating age-of-air which offers reduced uncertainty compared to using a single trace gas such as SF6. The method is described and illustrated in some detail. I think the paper will be suitable for publication after addressing the points below. My main comment is that paper uses modeled tracers to illustrate the method and I am not (at present) convinced that all sources of uncertainty have been accounted for. I.e. what about possible systematic errors in the model's ability to reproduce one or more of the individual trace gases used due to errors in kinetic parameters.

We sincerely thank the reviewer for their interest in our manuscript, their thorough evaluation, positive feedback, and constructive criticism. We have addressed all the major and minor comments in the revised version of the manuscript and provide a point-by-point reply below (with reviewer comments in black and replies in blue). The main changes in the revised manuscript are as follows:

(a) We slightly improved the AoA calculation method further by including an additional condition for the construction of the lookup tables to account for the deterioration of the tight correlations at lower mixing ratios,

(b) We added a comparison of the AoA-profiles derived from the GLORIA-B measurements with recently published AoA-profiles of an independent study (`https://egusphere.copernicus.org/preprints/2024/egusphere-2024-3279/`)

(c) We added a more thorough discussion of model uncertainties in the new section 4.4 "Model and method uncertainties", where uncertainties and possible biases of the CLaMS model are addressed (included in this document as section 4.4).

Main Points

1) A major aim of the paper is to argue that the new method will reduce uncertainty. However, quantification of the errors (e.g. uncertainty below 0.3 yrs as given in the abstract) is based on a model where it is possible to get the best possible agreement between the AoA from different methods. I think that the quoted uncertainty is mainly based on atmospheric variability. What about uncertainty and bias in the modeled tracers due to e.g. photochemical data, model-calculated loss rates. Related to that, how well does the model reproduce the individual tracer profiles (and tracer- tracer correlations) from the observations such as GLORIA. Please discuss these points.

   We agree that we need to address possible uncertainties and biases of the CLaMS model in more detail. We have therefore included a comparison of the AoA-profiles we have derived from the GLORIA-B measurements with the recently reported profiles of an independent study. The comparison was added as follows to Sec: 3.4 "Application to GLORIA balloon data":

> *The two profiles of "balloon borne Gimballed Limb Observer for Radiance Imaging of the Atmosphere (GLORIA-B) new method" are also in good agreement with the recently published mean age of air (AoA) profiles by Schuck et al. [2024b]. In their study, Schuck et al. [2024b] calculated AoA values using cryogenic whole-air samples of sulfur hexafluoride (SF_6) and carbon dioxide (CO_2) collected in Kiruna, Sweden, during August 2021. Similar to the AoA profiles of "GLORIA-B new method" the results reported by Schuck et al. [2024b] exhibit an increase in AoA with altitude, reaching approximately five years at around* 22 km. *Beyond this altitude, the AoA values remain approximately constant at their maximum of five years. There are good reasons to believe that the utilized model-based lookup tables do reflect the actual atmospheric conditions well enough to justify their application to the GLORIA-B measurements. These reasons will be discussed in the next section. On top of that, an alternative to the model in the form of global satellite measurements as the foundation of the lookup tables will also be discussed in the next section.* (Lines 475 to 483 in revised manuscript)

Additionally, we have introduced a new discussion section, "4.4 Model and Method Uncertainties," where we address the applicability of the improved model-based lookup tables to the measurements of GLORIA-B (added as Sect. 4.4 in this document). In this section, we highlight that previous studies have demonstrated the reliability of the Chemical Lagrangian Model of the Stratosphere (CLaMS) model for the trace gases used. We also emphasize that while errors in the model's kinetic parameters, which could introduce biases in the results, cannot be fully ruled out, a comprehensive analysis to quantify these uncertainties would extend the scope of our study.

2) The model simulation used here applies a HALOE-based mid-stratosphere upper boundary condition for the (non-zero) tracers. This will:

   (a) Impose a circulation speed on the tracers that may be inconsistent with the model dynamics (e.g., as pointed out for ERA5), and

   (b) Limit the use of this method to a period for which such boundary data could be derived.

   Please discuss the impact of these points.

   In general, we agree with the referee about the way the HALOE-based upper boundary condition could cause inconsistencies with the model dynamics as described in (a). However, the impact of this upper boundary condition on our results is expected to be very small, as effects of changes in the upper boundary on considered trace gas composition below 25 km are almost negligible. We evaluate the influence of the upper boundary of the model in the newly added section "4.4 Model and method uncertainties" in the discussion of the revised manuscript (added as Sect. 4.4 in this document) and added a short discussion on the limitations of our study there.

Minor Points

1) Figure 1 caption and elsewhere. "greenish". This is not a colour. Please choose an actual colour and use that for lines and descriptions.

The color of the profile "CLaMS clock tracer" and its uncertainty range in Figs. 1 and 10 (original manuscript) was changed to black in the revised manuscript.

2) Page 5, Line 109: The range 0 K to 70 K. Please explain more clearly the parameter you are using to determine the boundary layer up to 1.5 km.

   We have added more information about the hybrid vertical coordinate ζ and clarified that $\sim 1.5\,\mathrm{km}$ was chosen as the height of lower boundary of the model, and it roughly corresponds to the height of the planetary boundary layer

   > *ζ is equal to the potential temperature θ above a predefined level, and gradually turns to zero at the surface. Independent of the elevation of the surface, ζ is defined in a way that it is equal to zero at every surface point [for details see Pommrich et al., 2014]. The lowest model layer, representing the lower boundary layer of the model, extends from the surface to approximately 1.5 km (specifically, $0\,\mathrm{K} < \zeta < 70\,\mathrm{K}$), which roughly corresponds to the height of the planetary boundary layer.* (Lines 114 to 118 in revised manuscript)

3) Page 6: These loss reactions will, to varying degrees, depend on temperature (including T-dependent cross-sections) and the overhead column ozone. How is the OH field calculated? What about loss rates in the troposphere (e.g., tropospheric OH)? Please explain if and how these variations are taken into account, and the implications if they are not (e.g., increased uncertainty if these factors affect different tracers in different ways).

   The mechanism is in principle explained in the paper by Pommrich et al. (2014). Photolysis reactions are calculated as diurnal averages by the DISSOC photolysis code using overhead ozone from a climatology, in this case derived from HALOE observations (Grooss and Russell, 2005). OH, O(1D), Cl and Br are taken from a climatology created with the Mainz photochemical 2D model [Grooß, 1996] as diurnal average mixing ratios. This climatology had been pre-calculated for 18 latitudes, 34 pressure levels and 12 months by the CLaMS code. To clarify this we revised manuscript as follows:

   > *For the use in the present study, the additional trace gas species sulfur hexafluoride (SF_6) and HCFC-22 have been implemented into the CLaMS model similar to the method described by Pommrich et al. [2014]. The depletion mechanisms for HCFC-22 implemented in CLaMS is represented by the reactions ...* (Lines 131 to 133 in revised manuscript)

4) Pages 6 and 7: Please clarify the length of the model simulation. It must extend to 2022 for the GLORIA flights but, for example, HALOE cannot be providing the upper boundary conditions through to this date.

   The way in which the upper boundary conditions for N_2O and CH_4 were created over the entire simulation period is now specified in section 2.1 "Setup of the Chemical Lagrangian Model of the Stratosphere CLaMS":

   > *Similar to Pommrich et al. [2014], the climatology of the Halogen Occultation Experiment [Grooß and Russell, 2005] was used as the upper boundary for methane (CH_4) in this study. More specifically, the mean seasonal cycle from the Halogen Occultation Experiment climatology was used for every year of the simulation period. The Mainz photochemical 2D model [Grooß, 1996] was used for the upper boundary of nitrous oxide (N_2O).* (Lines 159 to 162 in revised manuscript)

5) Page 10, Table 2: It would be good to include some estimates of the stratospheric lifetimes of CH_4 (much longer than overall total) and N_2O (very similar to overall total) from other sources.

   The stratospheric lifetimes of N_2O and CH_4 are now added to table 2.

6) Page 16, Line 347: It is confusing to say "boreal winter season" and then point to results at high southern latitudes (which is in the summer). Please rephrase.

   The passage has been rephrased as follows:

   > *On average the uncertainties appear highest for austral summer season (1 January) at high southern latitudes, which is likely due to the collapse of the Antarctic polar vortex a few months prior, and the subsequent gradual spread of air masses that were contained in the vortex before its collapse.* (Lines 379 to 382 in revised manuscript)

7) Page 17, Figure 7 caption: Explain what definition is being used for the tropopause (T, lapse rate, PV, etc.).

   The following clarification has been added to the caption of Fig. 7 (original manuscript) and the captions of all other figures depicting the tropopause:

   > *(World Meteorological Organization (WMO) definition by lapse rate from 1957)*

Typos

1) Page 2, Line 48: "posses" → "possesses".
2) Page 4, Line 78: "the" → "they".
3) Page 6, Line 123: Use a small "s" for "sulfur".
4) Page 10, Line 232: Add a space after the closing parenthesis.
5) Page 13, Line 289: The lack of meridional gradients will be partly due to the much-decreased emissions by 2011. This should be pointed out.
6) Page 12, Line 288: "Lifetime" (remove the space).
7) Page 13, Figure 4 caption: "25 km in southern hemisphere" (not "on"). Use a small "p" in "part".
8) Page 13, Line 293: Add "section" for section 2.2.
9) Page 14, Line 314: Change "pretending" to something like "assuming".
10) Page 15, Line 340: Change "high" to "large".
11) Page 22, Line 432: "descend" → "descent".

Thank you for carefully watching out for typos in the manuscript. They have all been corrected for in the revised version.

4 Discussion

4.4 Model and method uncertainties

This study should, first and foremost, be viewed as a proof of concept for the proposed method to calculate AoA. Such a proof of concept can only be demonstrated within the perfectly self-consistent environment of a model, where the "true" values of the desired quantity are already known. Since no model is a perfect representation of the actual atmosphere, the lookup tables required for applying the proposed method should ideally be constructed directly from measurements, as laid out in Sect. 4.1, rather than from model results. Due to the lack of suitable measurement data, we had to establish the lookup tables required for applying the method to the GLORIA-B data using the available model results. We are confident that the CLaMS-based lookup tables used on the GLORIA-B data represent the actual atmospheric conditions reasonably well and find the AoA profiles derived from these lookup tables to be more plausible than the corresponding ones derived with the standard method (see Fig. 11). Our confidence in the CLaMS-based lookup tables stems from the following three reasons:

1. The reliability of the CLaMS model for the utilized trace gases has been demonstrated in previous studies (see, e.g., Laube et al. [2020] for trichlorofluoromethane (CFC-11), dichlorodifluoromethane (CFC-12), and HCFC-22, Konopka et al. [2004] for CH_4, and Pommrich et al. [2014] for N_2O).

2. There is excellent agreement between the AoA profiles reported by Schuck et al. [2024b] and those derived using the proposed method from the GLORIA-B measurements at Kiruna (see Fig. 11). The balloon-borne air samples collected by Schuck et al. [2024b] should closely resemble the air masses observed by GLORIA-B during its first flight, as both balloon flights took place in August 2021 in Kiruna, Sweden. The fact that Schuck et al. [2024b] obtained similar results for comparable air masses using a different calculation method strongly supports our expectation that the utilized CLaMS-based lookup tables reflect the conditions of the actual atmosphere reasonably well.

3. The upper boundary region, arguably the biggest source of bias in the simulation as a whole, seems to have only limited impact on the model tracer mixing ratios in the considered altitude region (i.e., below 25 km). Figure A4 in the appendix shows the temporally averaged zonal mean SF_6 distribution of the model results up to the upper boundary over two different three-month periods. Fig. 11a shows the average distribution from December 2010 to February 2011, representing winter in the Northern Hemisphere, and Fig. 11b shows the average distribution from June 2011 to August 2011, representing winter in the Southern Hemisphere. The winter periods were chosen because they represent the time when the downward transport of air from higher altitudes by the deep branch of the Brewer–Dobson circulation (BDC) is strongest in the respective hemisphere. SF_6 was selected because, in relative terms, it reaches the upper boundary in higher amounts than all other tracers (due to the absence of a stratospheric sink). Out of all model results, the influence of the upper boundary is therefore expected to be strongest for SF_6 during winter in the respective hemisphere. Figure 11 suggests that the downward transport 11, the downward transport of SF_6 depleted air from the upper boundary only seems to have an affect below 25 km at high latitudes in the winter hemisphere. This limited influence of the upper boundary on model results below 25 km could partly be related to the slow bias of the stratospheric European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5) circulation, as pointed out in sect. 3.4. This slow bias, however, effects the six long-lived trace gases in the same way as the model clock tracer and therefore can't be responsible for any biases the model-based lookup tables might have.

While these considerations give us confidence in the plausibility of the CLaMS-based lookup tables, it is important to emphasize that a thorough analysis would be necessary to fully assess the influence of the upper boundary on each of the six trace gases specifically, and to evaluate the extent to which the derived lookup tables reflect the conditions of the real atmosphere. Such an analysis would ideally involve the use of several atmospheric models and/or reanalysis datasets to better quantify potential biases and uncertainties. Additionally, the influence of uncertainties in the kinetic model parameters, such as photochemical data and model-calculated loss rates, on AoA cannot be ruled out. A thorough analysis of the influence of these uncertainties on AoA would also be needed. However, conducting such extensive analyses is beyond the scope of this study.