Reply to Reviewer # 2

We would like to thank the reviewer for his/her encouraging and constructive comments which will help to significantly improve our paper. Below, the reviewer's comments are printed in ***bold italic face*** while our replies are printed in normal face.

This paper is an important contribution to the ongoing discussion about large-scale circulation changes under climate change in the stratosphere. There is a lot to like about this paper, and it offers new and previously unavailable information about age variability in the stratosphere. In some areas messages could be clearer, and I am sure many modellers would appreciate a figure showing a simple seasonal climatology of the age in conjunction with all the other 'clever' diagnostics.

We will include a Figure on the zonal means of mean age of air representing the seasonality. An example how this figure would look like is provided in the reply to reviewer # 1, Fig. 2.

I am happy for the paper to be published with some revisions (largely concerning clarity), which I have detailed below. The only slightly bigger concern I have is the contradictory message arising from figures 8 and 10 (as mentioned below). But I am sure the authors will be able to clarify this with a better discussion.

We will expand the discussion of the figures 8 and 10 in order to make clearer what has been changed from Figure 8 to Figure 10 (i.e. the handling of measurement error propagation in the fitting procedure and its impact on the error assessment of the fitted parameters), what the differences between the two figures are (mainly the reduced areas of significance), and what we conclude from their similarities (that the results of Figure 8 may be robust as well, even if the errors of the "trends" are slightly underestimated).

P28015, l12: maybe add: in chemistry-climate models prescribing increases in long-lived greenhouse gases

P28015, l15: suggest instead of indicate

P28016, l4: the tropics

P28016, l12: give time horizon; 10% can be small or big depending on the time required for the change

P28017, l1: please remind the reader which time horizon the Engel et al. study covered

P28017, l17: please untangle this sentence; or just say you provide a climatology and give details later in the section

All these suggestions will be included in the revised paper.

P28020, l2: maybe it would be good to have statement about the similarity of the systematic biases in the early observational period compared to the later, degraded observational period; looking at Figure 2 one has the impression that the degraded period is systematically lower compared to the early observational period

We thought we did (see first bullet on page 28021). Nevertheless we will clarify.

P28020, l5: calculated instead of presented

P28021, l16: this addresses my comment above (p28020, l2), but I feel this should be mentioned earlier

Ok, both will be done.

P28024, l19: please specify which standard deviation you are referring to (all binned profiles?)

P28025, l14: if you so wish introduce AoA early on and use everywhere

P28025, l19: 'and the terms under the sum are' should just read 'and the sum of'

All these corrections will be applied.

P28026, 17: I do not understand this sentence

We'll explain in more detail how the regression method works, and in which way correlated and non-correlated errors are considered in the fitting approach.

P28027, l1: I am not sure what this equation is doing here, without any further explanation – two possibilities: re-write the previous paragraph in a more descriptive way and refer to the relevant literature, or finish of as detailed as you have started explaining

We' are in favor of the second choice and give some more explanation.

equation 5P28027: I am not quite sure what you are telling me; you refer to Figure 4 in a very generic way, yet you provide a lot of specific information that is hard to find in the figure – I suggest you re-write this part and guide the reader through Figure 4 (and include only the panels required for the story).

Ok, we'll guide the reader through figure 4 and give more explanations.

P28027, l14: just 'Recall'

Will be corrected.

P28028, l1: consider reordering, start with the fact that mesospheric is depleted in SF6 and explain the difference between apparent and real age (this way you would be able to shorten the discussion on the following page, which I would strongly recommend)

We assume that the comment refers to section 4.2. We'll reorder the section and shorten it.

P28031, l1: I like this discussion, but it would be nice if in the beginning something as simple as seasonal means of age could have been shown, instead of the large amount of information about the time series analysis

We'll include a short section on the seasonal means of mean age of air with a related figure before starting the detailed analysis of the time series.

P28033, l8: I am happy to have this discussion in the paper, but it should be made clear that the 'trend' (linear increase) might be significantly influenced by non-periodic, e.g. clustered events, like the number of sudden stratospheric warmings, or ENSO warm events – I doubt these events are correctly removed by the statistical model used – if anything the model might

be 'over-fitting' some variability

The existence of such non-periodic features is exactly the reason for our investigation in Section 7: Since we cannot easily predict these features or include them in our parametric model, we understand these as model uncertainties, and we try to estimate these via the fit residuals with the method described in Section 7. Since any of these events would introduce some correlation among the fit residuals of a number of consecutive data points, we have also included an autocorrelation analysis in the method, as described also in Section 7.
We'll try to add more discussion on these non-periodic events and how they would affect the estimated "trend" into the paper. Nevertheless, we have already pointed out that we do not consider the linear increases/decreases derived in our analysis as trends in a climatological sense, but as a temporal evolution which is left over if all periodic variations are removed.

P28034, l8: I suggest to simplify the phrasing, basically you are looking at different time periods, and as pointed out above (p28033, l8) you cannot remove reliably some long-term variability, or clustered events, which you should discuss earlier

Yes, this will be done.

P28034, l17: it is important to discuss the model results, but I would suggest highlighting the observed regional differences more (you have many regions where age decreases, but not necessarily where I would have expected them) – some of the high latitude NH structure might be influenced by the regular major mid-winter warmings of the last few years http://www.geo.fu-berlin.de/met/ag/strat/produkte/northpole/index.html

We'll go in some more detail regarding the discussion of observed regional differences, also related to non-periodic events as mentioned above.

P28036, l14: I have no clue what the message of this paragraph is – something is corrected and apparently no significant change occurs; if I look at figures 8 (top) and 10(top) I see many important differences (tropical lower stratosphere, SH high latitudes) and are left with a feeling of confusion – if figure 1o is the better estimate, the discrepancy between modelling and observations is bigger than previously stated; please explain better (and shorter) and provide a stronger link to figure 8.

As pointed out above, we'll explain in more detail what the similarities and differences of Figs. 8 and 10 mean, how significance of the resulting parameters has been estimated, and how it has to be taken into account when interpreting the figures.

We decided to include the analysis described in Section 7 and Figure 10 after a discussion with colleagues who tried to convince us that we had underestimated the "real" errors in the data – by not including auto-correlations among the data, and ignoring the inability of our simple periodic model to described the "real" variations in the atmosphere – and thus obtained "trends" appearing much more significant as they really were. They predicted that our "trends" would become all insignificant (i.e. with errors so large that they cannot be distinguished from zero) if all error sources had been properly included. This statement made us to include the missing error sources (our approach is presented in Section 7) and analyze the significance of the trends again. In short, Figure 10 tells us that the "trends" become less significant (i.e. difference to zero is less significant) if the residuals between the simple fit of Fig. 8 and the observations are taken into account as additional errors representing the inability of the model to describe the real "noisy" atmosphere. Differences between the upper panels in Fig. 8 and Fig. 10 point towards the latitude/altitude regions where the "noise" in the atmosphere (i.e. non-periodic variations) has its main effect. However, we find that most differences between Fig. 8 and 10 are in regions where the trend is no longer considered significant according to Fig. 10, lowest panel,

i.e. where the uncertainty of the trend becomes large, so that the differences do not really matter. In contrast to the statement of our colleagues, there remain regions with significant “trends”. The features in the regions where the trends are still significant are quite similar in Figs. 8 and 10, pointing toward robustness of these results.

What might have confused reviewer # 2 is the fact that we have discussed our results based on Figure 7 and 8, while we later state that Fig. 10 provides the more robust results in terms of significance of the “trends”. This was done because of the following reason:
Our approach to include further, so far not considered errors not covered by the measurement errors themselves is an approximation only; e.g. the model error has been assessed from the residuals between the fits and the measured data, which is an *ad hoc* approach. While this exercise was done to demonstrate that our results remain significant and robust (at least for some regions) even if some estimate of a model error was taken into account, we would not like to base all our discussion on such an approximate approach which would force us then to reject the results for large atmospheric regions because of their insignificance.
However, in order to make this clearer, more transparent and traceable to the reader, we will include a caveat in the paper before starting the discussion of the results, that the uncertainties of the discussed parameters – linear increase/decrease, amplitudes and phases of the seasonal cycle – may be underestimated to some degree, together with the hint towards the more rigorous error estimation in Section 7.