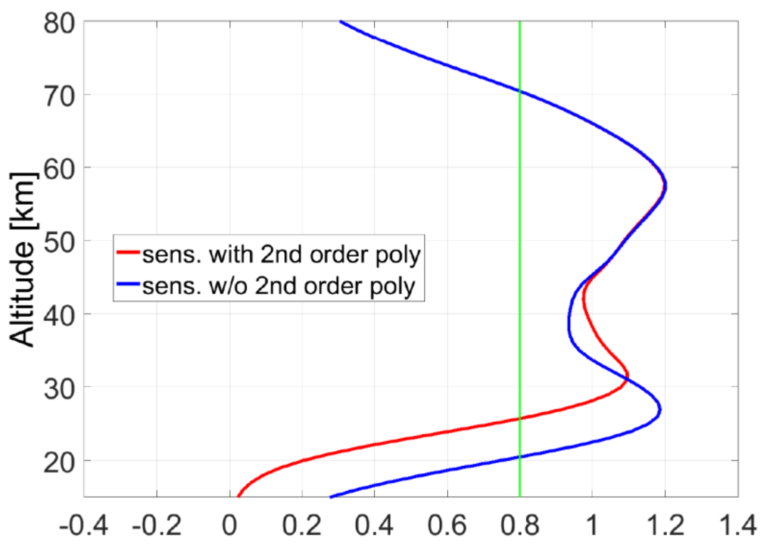
Replies to Reviewer 1

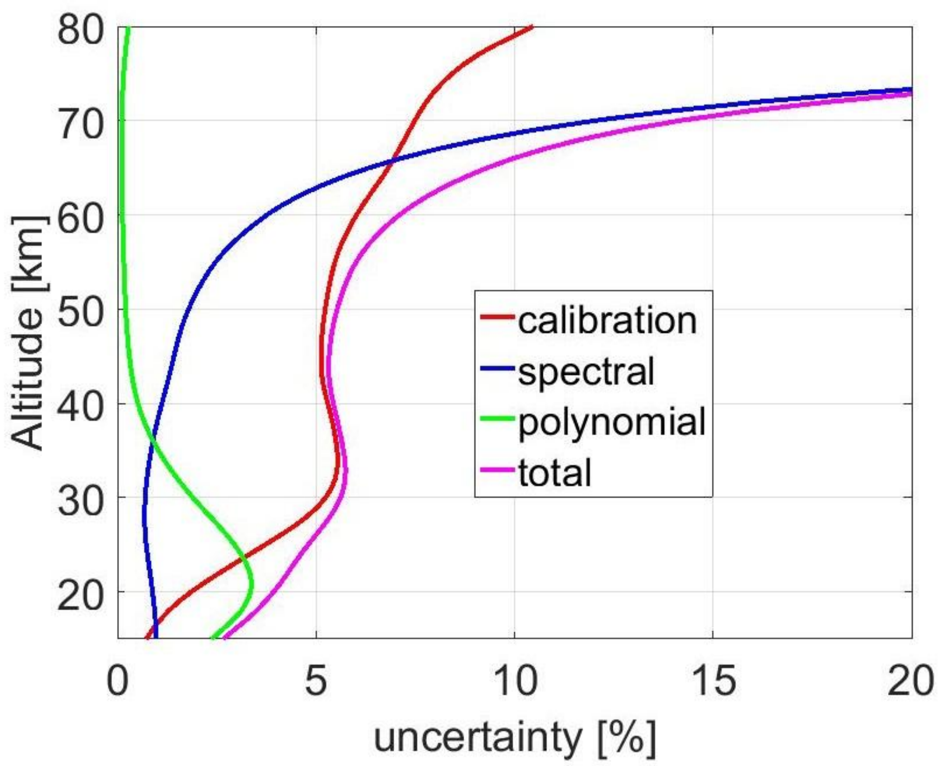
We thank both reviewers for their knowledgeable and valuable comments. Our efforts in addressing them, together with the reviewers' suggestions, led to a revised manuscript that represents a great improvement with respect to the original version. In what follows, reviewer's comments are in black and authors' replies in red.

There is a second order polynomial which is "added to the retrieval", but there is no discussion of the details of this process. If it is fit individually during each retrieval, and is properly included in the error analysis, then it will have a large effect on the lower stratospheric sensitivity. I am therefore somewhat skeptical of the accuracy of the sensitivities shown in Figure 8 at _30km and below. The other choice is to treat this as a systematic term, in which case it affects the systematic error but not the sensitivity. In either case, this can introduce an important uncertainty, and the authors need to discuss exactly what they have done.

We added to the manuscript a discussion on how the retrieval process includes a second order polynomial. This is done by fitting it in each retrieval, individually. It is not a systematic term. It does have a large effect on the sensitivity below 30 km but this is already taken into account in the sensitivity profile of figure 8 (original manuscript, now figure 9). In the figure below, we show the difference in sensitivity between employing the second order polynomial or only a first order one.

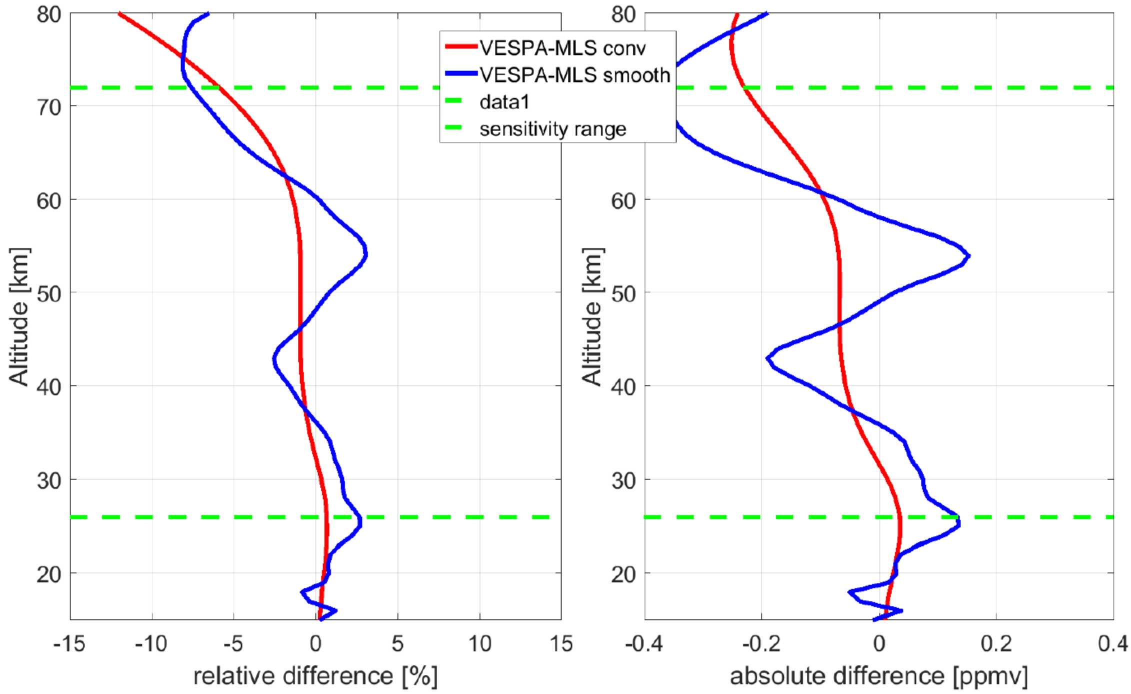


In order to address this issue raised by the reviewer, we estimated the uncertainty in the retrieved profile due to the use of the second order polynomial. The most rigorous way of doing it, in our opinion, is to take the uncertainty associated to the second order coefficient (Δa_2) in the retrieval process (say for example 20%, e.g., $a_2 = (-5 \pm 1) \times 10^{-3}$) and then perform two retrievals for the same spectrum with fixed values of the second order coefficient equal to $a_2 \pm \Delta a_2$ (in our example they would be $a_2 = -4 \times 10^{-3}$ and $a_2 = -6 \times 10^{-3}$). The resulting two vertical profiles would then provide an estimate for the uncertainty of the regular retrieved profile associated with the uncertainty in the second order coefficient. We found that the average uncertainty of the second order coefficient calculated by the optimal estimation routine over the entire dataset is 6%, with few retrievals showing more than 20%. We therefore decided to employ a fixed maximum uncertainty on the coefficient of 20% for the whole data set, rejecting from the data set those few retrievals (less than 5% of the total) that had an uncertainty larger than 20% in the determination of a_2. Displayed below is the updated figure related to the error analysis of the spectrum observed on 23 Dec, 2016, where we added the polynomial uncertainty to the other 2 sources. As expected the polynomial uncertainty has an impact mostly in the lower part of the retrieval.

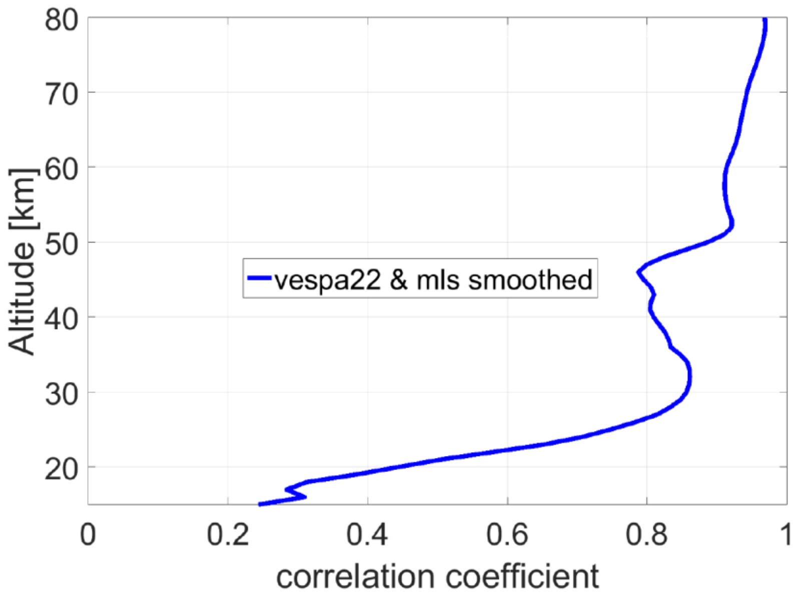


On top of it, in order to demonstrate the good quality of VESPA retrievals down to 25 km altitude, we changed the apriori used for the retrievals, which is now a fixed climatological profile up to about 48 km and a seasonal profile in the mesosphere (see also our reply to this specific issue raised by the reviewer in the following), and we introduced in the manuscript the correlation between VESPA profiles and MLS smoothed profiles. These are the original high resolution MLS profiles smoothed in the vertical with a running average of 10 km. This was done in order to decouple MLS profiles from the apriori and the averaging kernels used in VESPA retrievals (which was a short-coming of the correlation between VESPA and MLS convolved profiles shown in the originally submitted manuscript) and yet make the MLS vertical resolution somewhat similar to that characterizing VESPA profiles.

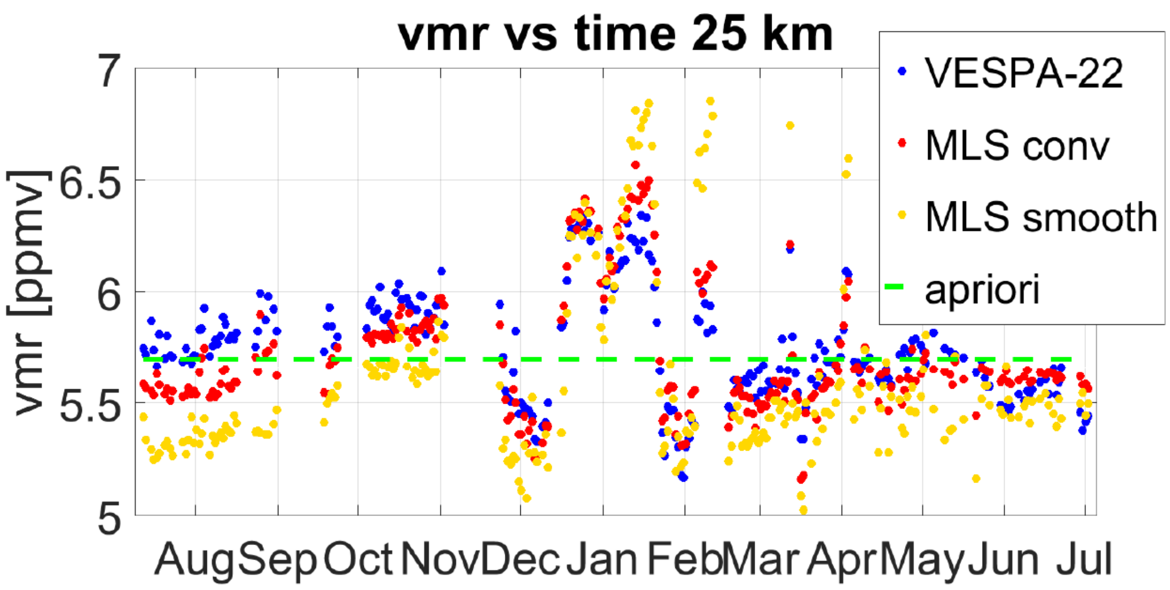
Figures down below show the relative and absolute average differences between VESPA and MLS smoothed (blue), and between VESPA and MLS convolved (red),



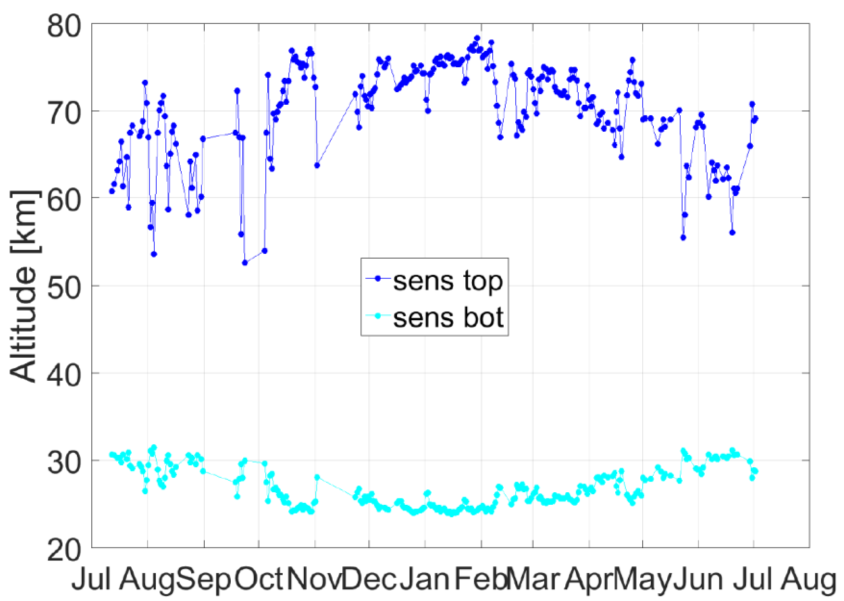
the correlation coefficient between VESPA and MLS smoothed profiles (blue),



and the time series at 25 km of VESPA22, MLS convolved and MLS smoothed.



All the figures above demonstrate in our opinion that VESPA22 retrievals are scientifically valuable in the sensitivity range indicated in the manuscript. Please note, however, that we also specified in the revised manuscript that the sensitivity range changes with seasons (see fig. 18 in the revised manuscript) and in summer it is approximately between 30 and 65 km. See figure below.



What was changed in the revised manuscript:

- Added discussion in Section 4 on how the second order polynomial is treated in the retrieval process;
- Modified figure 12 with the updated error analysis which includes the polynomial uncertainty;
- Modified figure 9 according to the new error analysis;
- Inserted the MLS smoothed data set, with its correlation with Vespa-22 profiles;

- Introduced the time series of the sensitivity interval, i.e., how the interval of accepted sensitivity changes through time. Added this time series in figure 18;
- Added two altitude levels in former figure 15 (figure 17 in the revised manuscript) to prove the good quality of VESPA-22 retrievals at the lower (25 km) and upper (75 km) limits of the sensitivity interval.

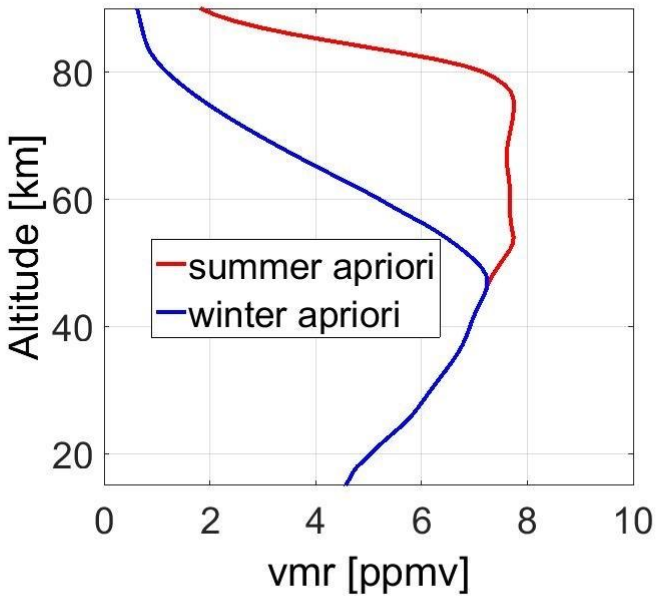
In the error analysis, many of the terms have been classified as systematic, when in truth, with the exception of the spectroscopy, almost all of them have a significant random component. As a result, the "retrieval uncertainty" shown in Figure 11, which by implication is the only random component, is, in my estimation, absurdly small. I think it would be acceptable to not specifically label the bulk of the errors as either completely systematic or random, but the present suggestion that the precision is <1% over much of the atmosphere would require a great deal of additional evidence.

We agree with the reviewer and changed the terminology accordingly. The uncertainty due to spectroscopic and calibration parameters is now indicated with "calibration uncertainty" in the revised manuscript.

The VESPA retrievals are performed using a seasonally varying a priori based upon 3 years of seasonally varying MLS data, hence the statement on page 29 that these are "independent datasets" is not true. Claims of high correlation between the measurements are therefore unsubstantiated. If the authors wish to publish claims related to correlations they should either perform their retrievals with a constant a priori (which would probably still leave them with good correlations), or compare deviations from the a priori for the two datasets (which is a tough test). Alternatively they could drop the whole discussion of correlation, along with Table 6 and Figure 14. Also, as long as the retrievals are done with a varying a priori Figure 15 should include a point indicating the a priori for each month.

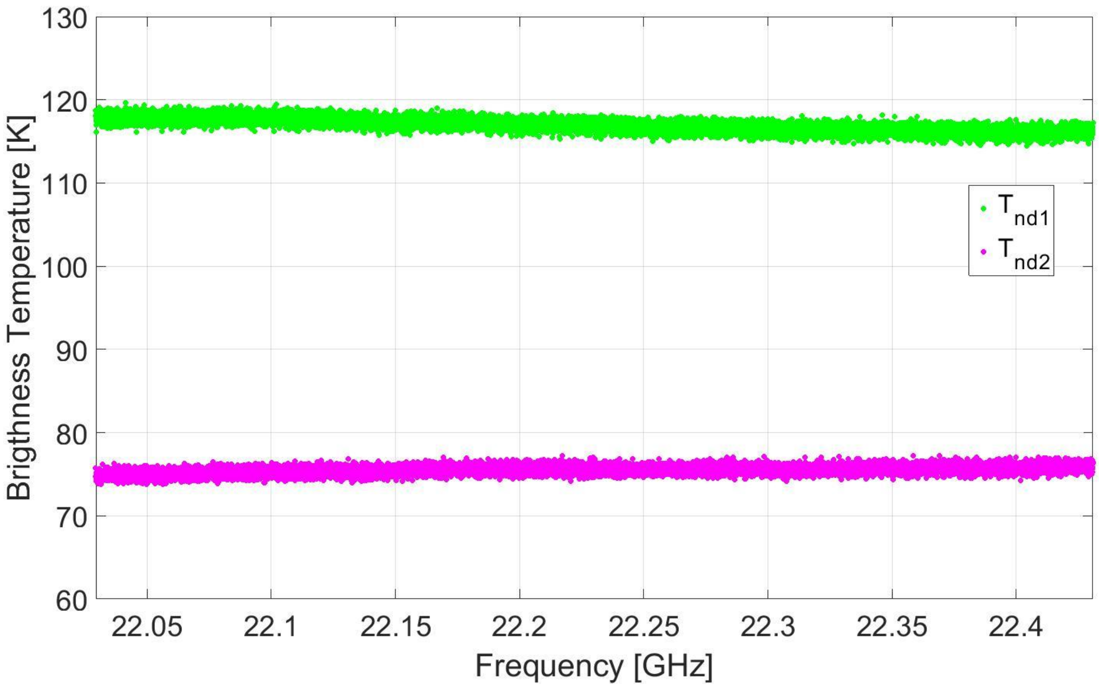
We followed the reviewer's suggestion and changed the whole analysis by using only two apriori profiles during 10 out of the 12 months of the year, identical below 48 km and diverging above, due to the large difference in Polar regions between summer and winter water vapor mesospheric profiles. We added figure 7 in the revised manuscript (and below) showing the two apriori profiles. The summer apriori (red line) is used during the period from June 1 to August 31, while the winter apriori (blue line) is used during the period from October 1 to April 31. During the month of May and September, there is a transition period in which we used a linear daily interpolation from one apriori to the other to provide continuity in the retrieved profiles timeseries.

We also added the time series of the apriori mixing ratio values in figure 17 (former figure 15) as suggested by the reviewer.



Page 10 line 7 – Where does this “constant in frequency within 1.5%” number come from? The diode data sheets? Or perhaps a calibration measurement?

The indicated 1.5% is half of the maximum difference between brightness temperature values over the spectral passband (see figure below). The noise diodes emission spectra are measured only during the LN_2 calibration. In the revised manuscript we rephrased the “constant in frequency” sentence with “The noise diode produces a signal that is measured to be quite stable in frequency. In fact, single-channel T_{nd} values are always within 1.5% of the spectral mean of the diode temperature brightness T_{nd}.”



Equation (4) – presumably ‘x’ is known for this transition. What is it?

We added a sentence to direct readers to Table 2 where all the spectroscopic parameters used are listed

Concerning all the comments below, we agree with the reviewer and changed the manuscript accordingly.

Page 6 line 20 “negligible”. Perhaps “small” would be better here. The numbers are given later in the paragraph, so everything here is okay, but for some applications a 1.5% difference might not be considered “negligible”.

The use of Tatm is a bit confusing, as this appears to be the atmospheric temperature for the troposphere, but not for other parts of the atmosphere. Would it be better to label this as Ttrop?

“Tsup” is a rather odd abbreviation for surface temperature.

Page 11 line 7: “explicit” should be “explicitly give”

Page 15 line 13 – “Selee” should be “Seele”

Page 17 line 8 – missing ‘(‘   ?????????

Figure 9 –“specttrocscopy” should be “spectroscopy”