Revised Response to Referee 2

We are thankful to the referee for the important discussions and the suggestions toward improving the manuscript, in particular regarding the theoretical part of the forward model (section 5.4.1). We have tried to answer the questions and have given our responses to all comments.

Keeping in view the referee's genuine concerns regarding the use of superscripts and subscripts in section 5.4.1., and after discussions with the coauthors, further changes and corrections are made to this section. This implies that here the authors' response to referee 2 is revised for this particular section

Page 1032, line 6 & 7:
"high accuracy" is not a given outcome of due to the "self calibrating" nature of occultation. The "self calibration" needs to be of high accuracy.

Rephrased the sentence as: "SCIAMACHY's measurements in the occultation mode are self calibrating and have high accuracy".

Page 1032, line 14:
What is the size of the SCIAMACHY FOV relative to the moon?

"The SCIAMACHY's FOV in lunar occultation mode is 0.045° (2.5km) in vertical direction and 1.8° in horizontal direction. The apparent diameter of the Moon is between 0.49 and 0.57° in the horizontal direction, which is within the instrument's FOV". These sentences are added before line 14.

Page 1032, 17 & 18:
The sentence about the angular rate of the moonrise is not clear.

Below 17.2 km, due to the refraction in the Earth's atmosphere, the Moon is no longer visible at the predicted position. We have replaced the sentence with the following simpler form as, "Below this altitude, the lunar signal is rapidly decreasing"

Page 1032:
Sentences on line 18-20 should be moved to the end of the paragraph. It disrupts the flow of the discussion.

The sentences on line 18-20 are moved as suggested.

Page 1032:
How does the switch between the scanner following predicted movement and using the MFD impact the occultation quality? Does this affect the calibration?

In principle, the switch between the scanner following predicted movement and using the MFD impacts the occultation quality, but it is in the order of 5 millidegree in our case and does not have a significant effect on calibration.

Page 1033, line 26:
Reword to "…fact that the latitude of the sub-satellite point changes…"

The sentence is reworded as suggested.

Page 1034, line 9:
What causes the light to be "scattered"? Presumably it is the Earth's atmosphere, but this should be clarified.

The clarification is inserted as "The smaller the SZA value, the higher is the influence of the sunlight scattered due to the Earth's atmosphere…".

Section 4, Page 1034 & 1035:
Why is an a priori used in the retrieval? One of the hallmarks of solar/lunar occultation is the minimal dependence upon assumptions. What is ill-posed about this particular problem that causes you to use an a priori?

Optimal Estimation (OE) is a standard method in obtaining profile information in remote sensing (Rodgers, 2000). In our case, the a priori has almost no influence on the retrieval (see measurement response in fig. 11). One main advantage of OE is to get such diagnostic information as the measurement response. We agree that other retrieval methods without a priori information are also possible here.

The problem is ill-posed because the solution is not unique. Since we have spatial under-sampling, i.e. the measurements are performed on a 3.3 km grid but the profiles are retrieved on a 1km grid, so other nonphysical solutions are possible (oscillations). The shape of the a priori and the regularization avoids this.

Section 5.1:
Why are two different version (6.3 & 7.0) of data used?

In our study we used the SCIAMACHY lunar occultation measurements for the period 2002-2010. The data version 6.3 contains the measurements from 2002 until 2009. The version 7.0 covers 2002-2010. For the period 2002-2009, there are no differences among both versions, regarding the calibrations and the wavelength region used.

Section 5.1:
What is level 1b data? Is it radiance or transmission?

The level1b raw spectra are the radiances. Following additions are implemented:

Section5.1:- "The l1b data contain the measured raw spectra and all information necessary for calibration".

Section5.3:- "The ratio of the lunar spectrum to the reference is the transmission spectrum used".

Section 5.2:
Define 'differential optical depth'.

Since the differential optical depths are discussed in detail in section 5.4.1, a reference to this section is inserted as "The attributes of the differential optical depth spectra and their calculation are described in sect. 5.4.1".

Section 5.4.1:
Page 1038, line 15: Change prior to priori.

Done.

Page 1038, line 15: Does the Sa also come from the US Standard Atmosphere climatology (1976)?

We have no climatology value for the covariance of the a priori. We assume 100% a priori covariance. Because the solution is dominated by the measurement, this rough assumption is feasible. We have added this in the text after eq. 4 as "In this study, 100%, a priori covariance is assumed".

Page 1038, line 23: Need to state that the 'increased' linearity is due to the fact that the fundamental physics is the transmission of the light.

Suggestion implemented.

Page 1039, line 6: What is the 'retrieval parameter index?'

The retrieval parameter index 'k' stands for the number of parameters retrieved, which is '1' in our case since we only retrieve the water vapor profiles. We have removed this index at this point in eq. 8. Moreover, the subscript of c (the coefficient of the polynomial) is corrected as 'I' representing the polynomial index and not 'j'.

Page 1039, line 10: What is superscript "l" in the numerator? Shouldn't the 'ref' be a subscript since that is the script used for altitude?

In eq. 9, the superscript 'l' refers to the lunar spectra at different altitudes other than the one at 120 km which is used the reference spectrum, specified here with the superscript 'ref'. The spectra recorded at different altitudes are normalized using the reference spectra to obtain the transmission differential spectra. We have revised and corrected the relevant indices and their specifications as follows:

Correction at page 9, lines 18-20:
"….the lunar spectrum $I_j(h_j, \lambda)$ or simply $I_j(\lambda)$ was extracted for 13 tangent heights, between 17—50 km, selecting the 14th at 120 km which was the reference spectrum $I_{ref}(\lambda)$. The ratio of the lunar spectrum to the reference is the transmission spectrum used."

Thus in eq. 9, 'l' is removed (we noticed that it was also not specified in the text) and 'j' and 'ref' are placed as subscripts for the spectra at the given tangent heights 'j' and the reference tangent height respectively. The radiances in eq. 9 should be with tilde, it is now corrected.

Page 1039 and elsewhere try nonlinear instead of non linear

Done.

Page 1039, Eq. 10: Need a subscript on I in the numerator.

Inserted the subscript ‘j’.

Page 1040, line 10: What W are you referring to? Eq. 12, “the above equation” has W hat (or is it W tilde).

It is W tilde. Corrected this typo in eq. 11 and 12 and in the text.

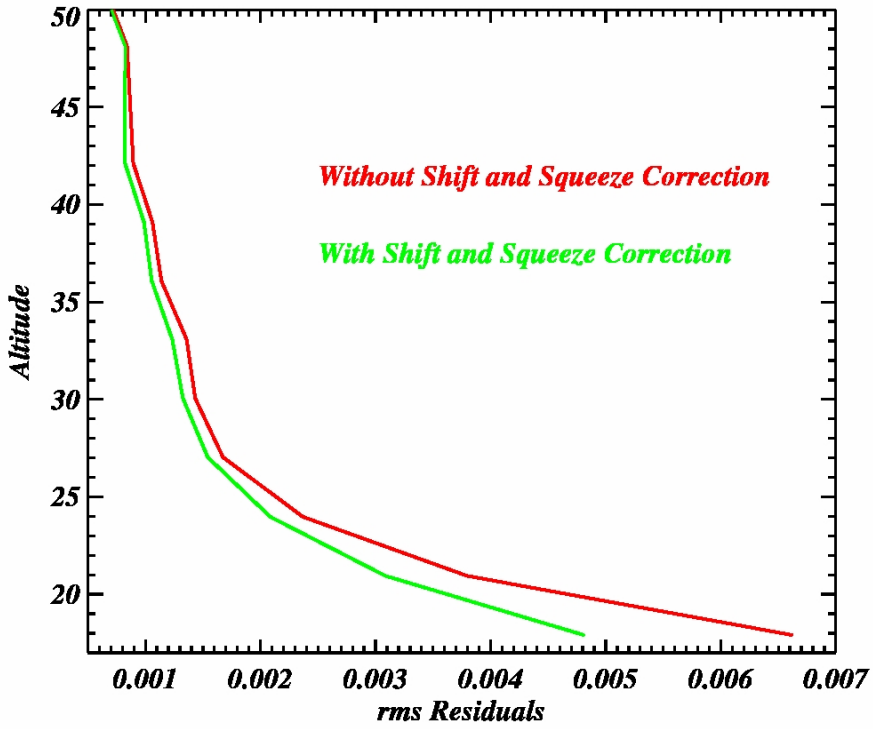
Changes in eq. 10:
The weighting functions for a given atmospheric parameter ‘k’, at a certain tangent height ‘j’ are represented now simply as $W_j^k(\lambda)$ in accordance with the specification for the retrieved parameter α^k.

Changes in eq. 11 & 12:
The indices ‘j’ and ‘k’ are adopted as in the previous equations. Eq. 12 is corrected following the suggestion from coauthors.

We are grateful to referee for a careful reading. After the simplification of indices as adopted now and omitting multiple useage of same indices, in sect. 5.4.1, the author hopes that the section is not confusing anymore.

Page 1040: Is Doppler shift computed explicitly? Why isn’t a wavelength shift computed for the event using the Reference spectra and then applied to all the other spectra within the event? How do the shift and squeeze parameters vary with altitude?

All possible spectral misalignments are handled by the shift and squeeze algorithm including the Doppler shift. We cannot use the reference spectrum only, to compute the wavelength shift as it would not contain any spectral features of water vapor. We obtain smallest residuals with individual shift and squeeze since the water vapor absorption signature is different for different altitudes. The shift and squeeze correction improves the residuals especially for the lower stratospheric heights, and the response to shift and squeeze decreases with increasing heights. The figure shown below demonstrates the effect of shift and squeeze on the residuals. In the figure, the rms of the residuals, with and without shift and squeeze application, are plotted for the whole altitude range of the retrieval.



Page 1041:
So, really how much of a time difference is there between the ESFT and LBL? Since you ran multiple cases you should have reliable estimates of how much time is saved.

In case of lunar occultation, where the applied radiative transfer is rather simple, on the average, one LBL computation takes about 4 hours while, one ESFT profile takes around 1-2 minutes. The following sentence is included in the section. "ESFT is preferred over LBL since it provides a good compromise between efficiency and accuracy".

Page 1043, Section 5.4.4:
What resolution is addressed here? SCIAMACHY has finite resolution in both the spectral direction (most likely the resolution discussed here) and the spatial dimension, which I don't think is accounted for in the retrievals.

Yes, the spectral resolution is addressed here. Specification included in the text.

Page 1044:
Was the ESFT vs LBL fit done for all events during 2008?

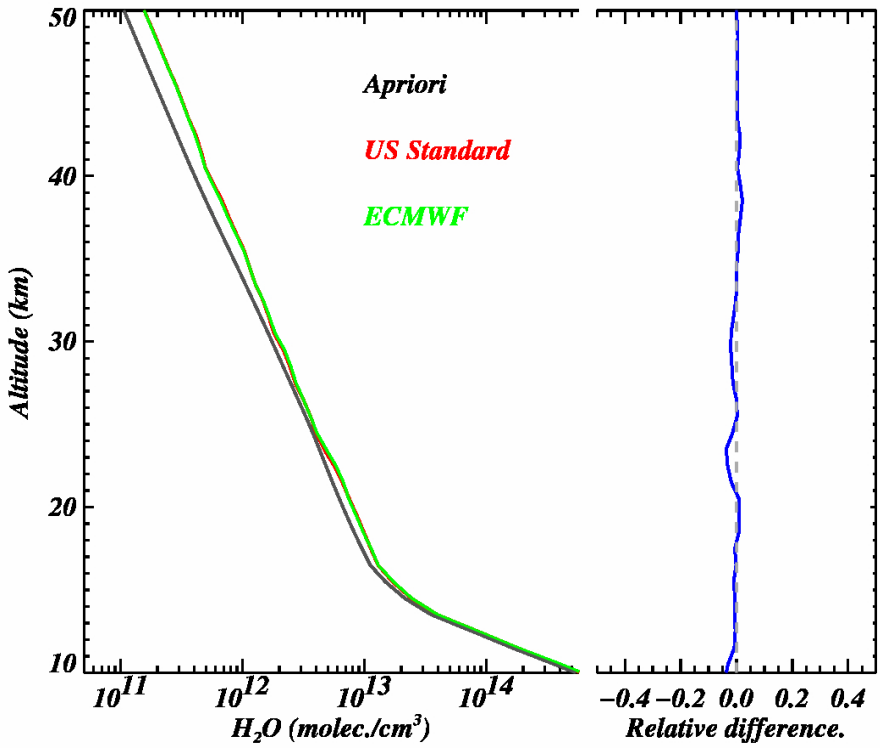
Yes, the ESFT-LBL fit was done for all the events during 2008.

Page 1045, line 12:
Again, what is the FOV of SCIAMACHY, 3.3km is spacing of the samples.

The question is already answered above in response to the comment on Page 1032, line 14**.**

Page 1045, section 7:
Does it make any difference if the SCIAMACHY concentrations are converted to vmr using the temperature and pressure profile that is assumed in the forward model?

We do not convert the SCIAMACHY concentrations to vmr. SCIAMACHY does not measure temperature (T) and pressure (P). The SCIMACHY vmr would include T and P from the climatology source. The retrieval is actually found to be rather insensitive to T and P source. The figure shown below explains this observation, where for a single measurement, the number density profiles retrieved using the US Standard climatology (in red) and the ECMWF source (in green) are plotted on the left. The right hand side shows the relative difference between these profiles using the investigated T and P sources.



Section 8:
Do the comparisons change depending on the: lunar phase, latitude of event or solar zenith angle?

The measurements used in the study are selected implementing the quality criteria mentioned in sect. 3 i.e. selecting the measurements with moon phase ≥ 0.75 and solar zenith angle, SZA ≥ 96°. Therefore the measurements with bad lunar phase or SZA are already excluded. Furthermore no dependence of the comparisons on the moon phase, latitude or the SZA was observed during the course of study.

Page 1050, Conclusions, line 15:
Delete 'very' since it can mean different quantitative values to different people.

Suggestion implemented.

Page 1050, line 16:
The figure 10 does not show zero difference between ESFT and LBL, so you can't claim that there is no bias. There still exist a 3% difference that may be condition dependent.

We agree to referee's comment. In the sentence at line 16 "….. increasing the number of coefficients and thus preventing any systematic biases in the resulting ESFT profiles." is changed as "increasing the number of coefficients and thus reducing any systematic biases in the resulting ESFT profiles below 3% on the average."

Page 1050, line 20:
Again, drop the use of "very good". What you've done is attempt to make the case that SCIAMACHY Lunar Occultation Water Vapor profiles are comparable to other measurements of stratospheric water vapor and thus you believe they are scientifically useful.

"very good" is replaced by "scientifically useful".

Fig. 8:
What is the other line plotted in each figure?
The other grey line is the a priori profile. It is now mentioned in the caption of fig. 8.

Fig. 9:
What is the definition of Theoretical error? Is this precision or accuracy? I think you mean precision, since the OE method only produces that quantity. Why do you plot results below 17 km, when you only use data between 17 and 50 km?

The theoretical errors represent the precision. Already mentioned at the end of section 4.

Yes, it is better to have plot corresponding to the retrieval altitude range only. The plot is modified and the results below 17km are excluded. The same is done for the number densities plots in fig. 8. Thus fig. 7 and 8 are changed in the revised manuscript.

Fig. 12:
I don't understand how the retrieval is so different from the a priori at 50 km, when Fig. 11 says only half the answer is from the measurement and half from the a priori.

The regularization applied in the retrieval does not allow for jumps (non physical). When there is an influence from the a priori, the profile smoothly goes to the a priori. The applied regularization is the main explanation for this behavior.