Response to reviewer 3:  
Stratospheric and mesospheric HO2 observations from the Aura Microwave Limb Sounder  

16 December 2014  

We sincerely thank reviewer 3 for his/her thoughtful comments on the previous draft, we hope this new version is more suitable for publication.  
In doing the corrections of all reviewers we added the following 3 major changes:  
(1) A paragraph at the beginning of the results sections explains that the averaging kernels were applied to all comparisons:

In this section we compare the offline HO2 dataset with balloon-borne and other satellite measurements, as well as, with global climate and photochemical model simulations. In making these comparisons, i.e. when showing the absolute or percentage differences between the datasets, the MLS averaging kernels has been applied to properly compare them. Furthermore, when comparing the global climate or the photochemical model simulations, its high vertical resolution has been reduced to the MLS one using a least square fit as described by Livesey et al. (2011, Sect. 1.9). In these comparisons, no altitude extrapolation has been applied to any dataset.  

(2) the discussion about the impact of the O2 and H2O cross section was deleted, the
discussion about the mesospheric discrepancies now reads:
These discrepancies might be due to a variety of reasons, for example: (1) our understanding of middle atmospheric chemistry may not be complete, (2) there might be due to differences between recent solar spectral irradiance (SSI) satellite measurements (Snow et al., 2005; Harder, 2010) and most parameterizations. These SSI measurements display a larger variability in solar UV irradiance which can-not be reconstructed with SSI models, including the model of Lean et al. (2005), used in this SD-WACCM run (Marsh et al., 2013). These SSI measurement-model differences have been proven to affect the HOx photochemistry (Haigh et al., 2010; Merkel et al., 2011; Ermolli et al., 2013); more UV irradiance leads to an enhancement of O3 photolysis as well as H2O photodissociation, which leads to more HOx production through (Reactions R4 to R8). Further, Wang et al. (2013) showed that using a solar forcing derived from these SSI measurements the modeled OH variability agrees much better with observations. Lastly, (3) these discrepancies might be related to the WACCM representation of the mean meridional circulation which has been shown to have some deficiencies (Smith et al., 2011; Smith, 2012), suggesting that the gravity wave parametrization needs to be modified. In addition, Garcia et al. (2014) has shown that adjusting the Prandtl number, used to calculate the diffusivity due to gravity waves, significantly alters the CO2 SD-WACCM simulations improving its agreement with satellite measurements. Such adjustment should also affect the H2O and hence the HOx chemistry.

(3) the photochemical model discussions now reads:
As shown in Fig. 12, in the upper mesosphere (pressures smaller than 0.1 hPa), the Kinetics 1 simulations do not reproduce the magnitude of the measured peak, underestimating it by as much as 60%. On the other hand, Kinetics 2 shows an improvement in the modeling of this peak, reducing the underestimation to less than 40%. These discrepancies coincide with the ones discussed in the previous section strongly suggesting that they are related to the model assumptions rather than to measurement errors. As with the SD-WACCM simulations, several factors could be the reason for this discrepancy: it might be due to limitations in our current understanding of middle atmospheric chemistry and/or due to the deficiencies in the model solar spectral irradiance used, in this case Rottman (1982). Also, considering that Kinetics 2 (the run testing the HOx partitioning) represents the measured HO2 better, these simulations might suggest that, the modeling problems are related to the HOx production and loss balance rather than the HOx partitioning. In the upper stratosphere and lower mesosphere (between 1 and 0.1 hPa) for the most part the photochemical model underpredicts HO2 by around 20% concurring with the SD-WACCM simulations as well as with previous studies (Sandor et al., 1998; Khosravi et al., 2013) but contradicting the result of the study by Canty et al. (2006).

Below are our responses to the reviewers comments in red.

Review of manuscript “Stratospheric and mesospheric HO2 observations from the Aura Microwave Limb Sounder” by Millán et al.

General comments:

This work presents a new HO2 dataset derived from Aura MLS measurements using an offline retrieval algorithm. The product, retrieved from averaged radiance profiles, presents several advantages over the standard v3.3 product, as the extended altitude range, the coverage of the polar regions, and also provides nighttime values for a wide altitude region. The manuscript describes the algorithm and the characterization of the retrieved quantity and assess the different error sources.

Comparisons with balloon-borne measurements and satellite measurements are
also presented as well as a comparison with a 3D chemistry climate model and a 1-D photochemical model. It is claimed that this dataset can be useful for a better understanding of the mesospheric O3 and HOx chemistry. In particular they found that the HOx partitioning in the retrieved HO2 and OH from MLS are compatible with our current understanding of the mesospheric chemistry. However, the absolute values of mesospheric HO2 are significantly underestimated by the models. Possible reasons for this underestimation are mentioned/listed although not really discussed or addressed.

I think that this new dataset of HO2 measurements add significant extra information to the standard product (e.g Fig. 2) and hence worth to be published. The result on the HOx partitioning of MLS products is also a significant contribution from the scientific (not only methodological) point of view. The other scientific result is just to point out to a models/MLS measurements disagreement which is not addressed. Then, it is not clear for me if the paper should be published in AMT or in ACP. I suggest that the authors give some more details and discussions on the possible causes of the disagreement (see below). This would make easier its publication in ACP.

Major comments:

Page 22913. Lines 21-24. I do not understand the meaning of the "retrieval numerics" error. My first guess would be that they are the "forward model" error, but this is considered in a separate contribution. Would that be what is normally called "smoothing" error? I.e., the effects of the regularization used in the retrieval? In the sentence "It is calculated as the retrieved value from the unperturbed radiances and the "truth" model atmosphere, i.e. that used for computing the synthetic radiance.", was that retrieval done with or without adding the noise to the synthetic radiance? I think it is important to clarify this error, since it is the major uncertainty in the region of HO2 maximum, above 0.1 hPa (Fig. 5).

We modified the sentence to state: The comparison between the unperturbed noise-free radiances run, and the "true" model atmosphere estimates the errors due to the retrieval numerics, which, in other words, is a measure of error due to the retrieval formulation itself, in this case, mostly an smoothing error.

Related to this point, if they are actually the "smoothing" errors, they would be already taken into account when applying the AKs to the data to be compared and hence, the "bias" would not be as large as the 1 ppbv shown in Fig. 5 but significantly smaller. If this interpretation is correct, I would not consider this error as a "bias" and would not mix with the other systematic (bias) errors. In this case, we do not apply the kernels to the truth profile, to see the effects of the retrieval (the smoothing).

Related to this point, what is the "scatter" of the errors? What do they indicate? The standard deviation, the title of the plot was changed from Scatter to Standard Deviation

In Secs. 4.3 and 4.4 the authors mention possible reasons for the discrepancies between the WACCM and 1-D models and MLS HO2 measurements. In particular they refer to our current uncertainty on the knowledge of the solar spectral irradiance measurements and/or its model representations, and the spectral resolution of the absorption cross sections of H2O and O2. Could the authors give some more details on what use the two models for these quantities? Do they have some hints on why they think they are possible causes or is it just speculation? After careful consideration the discussion about the spectral resolution of the absorption cross sections was deleted because it was based plainly in the photochemical model representation of these values but we didnt change the resolution to corrobo-
rate the hypothesis. With respect to the solar spectral irradiance the discussion was expanded to: These discrepancies might be due to a variety of reasons, for example: (1) our understanding of middle atmospheric chemistry may not be complete, (2) there might be due to differences between recent solar spectral irradiance (SSI) satellite measurements (Snow et al., 2005; Harder, 2010) and most parameterizations. These SSI measurements display a larger variability in solar UV irradiance which cannot be reconstructed with SSI models, including the model of Lean et al. (2005), used in this SD-WACCM run (Marsh et al., 2013). These SSI measurement-model differences have been proven to affect the HOx photochemistry (Haigh et al., 2010; Merkel et al., 2011; Ermolli et al., 2013); more UV irradiance leads to an enhancement of O3 photolysis as well as H2 O photodissociation, which leads to more HOx production through (Reactions R4 to R8). Further, Wang et al. (2013) showed that using a solar forcing derived from these SSI measurements the modeled OH variability agrees much better with observations.

Other comments.
- Figures are very small and they have so many panels that are hardly readable in the printed version (I could read them only when zoomed out on the screen). Figures 5,7,8 and 12 should be pagewidth in the final publication aiding the readability.

In this sense, most of the figures are duplicated presenting the results in vmr and in number density. I cannot see any advantage of presenting additionally the number density figures. I think they could be removed and would help to make the other panels more readable.

After careful consideration we decided to leave the duplication. Even though most people are familiar with the VMR unit, in the OH and HO2 community most papers (Pickett 2006,2008, Canty 2006, and Wang 2013) use number density units.

- Fig. 1 and Page 22910 (lines 8 and 9). The text refers to 1K, 2K and 4K limb radiance precision. Is any of these that shown as the noise in Fig. 1? Why do you compare between these three precisions to say that the noise is large and averaging is needed? The signal at the top panel (band 28) for 4.6 hPa is much smaller (particularly at night) than the noise. However it looks as not affected by noise (very smoothed). Is it because the number of measurements averaged is very large? Would be useful to mention that number in the figure caption.

The text was changed to The anLi 1 K HO2 signal is relatively small compared to the individual limb radiance precision which varies from 2 K at the bands edges to 4 K at the band center (gray dotted line), hence ...

The caption does state that this is a monthly radiance average.

Page 22910. Lines 25 and ff. Just for curiosity, are the non-zero nighttime abundances positive, negative, both?
Both, that’s why we didn’t specify the sign.

You suggest to take the nighttime values as the “zero” for calculating the daytime values. However, the daytime and nighttime measurements are taken on different parts of the orbit (either ascending or descending). For other instruments the offset changes significantly along the orbit. Is that a good approach for MLS or is the uncertainty in the correction of a similar magnitude that the correction itself?
To imply that this is a valid approach for MLS we added: In addition to the MLS HO2 product, this day–night difference approach to ameliorate biases has been used successfully for the BrO and OH MLS products. (Livesey et al., 2006b; Pickett et al., 2008; Millan et al., 2012)

Sec. 3. First full par. To be safer, I would consider as the daytime scans those with SZA< 85. Would that make a significant change in the polar regions?
Before setting in the 90-100 SZA we did check other options but they did not make much difference.

Near the end of this par., lines 10-13. "interpolated radiances". Apparently the sampling in altitude of MLS is <1 km, and the vertical grid used here is 3km. Hence, it is also done some kind of "averaging" rather than "interpolation" in the radiances. Isn’t it?
Correct, we changed it to averaged

Lines 19-23. You mention here that "... for pressures between 10 and 1 hPa where the nighttime values exhibit non-zero values indicative of biases." However, Fig. 5 shows that the biases are not particularly large at those pressure levels; actually they are larger at lower pressures (higher altitudes). Shouldn’t daytime values be calculated in a similar way above around 0.1 hPa, where the bias is also large?
At 0.1 hPa the night values are expected to be non-zero hence not usable for bias correction. Below 1 hPa they are expected to be zero and hence any non-zero value is a sign of an artifact

Fig. 3. Are the results shown here for a daytime case? Please, state that, if so.
Done

Fig. 4. I would remove the number density plot and would use a log scale for the errors. Log scale Done. About the log scale for the errors, we prefer the absolute to easily emphasize how the increase with height.

The caption refers to a "This profile". Is it the solid black line?
The caption was change to: The black lines show typical HO2 profiles, daytime in solid and nighttime dashed. These profiles are a yearly average over all latitudes of the SD-WACCM model

Last par. in Sec. 3.2, lines 6-8. "For pressures smaller than 0.1 hPa, the main source of bias and scatter are retrieval numerics, which, although unsatisfactory, is understandable given the 14 km vertical resolution in this region." This suggests to me that you are talking about a "smooth" error (see above). Correct?
Correct

Fig. 5 caption. families of systematic errors - sources(?) of systematic errors
Done

As before, I suggest to remove the panels with the errors in the density. Idem for Fig. 6.
See above

Sec. 4.1 Comparisons with FIRS-2. How many FIRS-2 profiles are available for that day? Just that used? If there are more but taken at other SZA’s, and if SZA is very important, they could be corrected with a photochemical model. I think the statistics should be increased.
We carefully thought about this. There are more than one but adding the photochemical correction will also add an extra uncertainty and so, we decided against it.

BTW, in the figure caption is not mention that it is just one FIRS-2 profile.
We added in the caption: The FIRS profile corresponds to the one with the closest SZA to the MLS (daytime only) data.
"The retrieval top level differences will need to be explored further..." Given that there are so few HO2 measurements, and the importance of these measurements for the mesospheric chemistry (next sections), should not this be explored further in this work? It is important to clearly state that the models/MLS measurements comparison in the next section is not caused by a bias in MLS HO2 data.

To find the cause we will need a joint effort between the MLS and SMILES teams outside the scope of this study. We added: The retrieval top level differences will need to be explored further, to investigate if they are due to retrieval artifacts (both retrievals are more sensitive to the apriori at these levels), calibration uncertainties or sampling differences (unlike MLS, SMILES data are not regularly distributed); this will require a joint effort from the MLS and SMILES teams.

Sec. 4.3. It is known that WACCM does not reproduce very well the measured temperature and O3 fields and even the meridional circulation (e.g. Smith, 2012; Smith et al., 2011; 2013). Could these be possible reasons to explain the HO2 WACCM-MLS differences? Furthermore, Garcia et al. (2014) has found that the parameterization of the gravity waves (GW), done through the change of the Prandtl number, significantly changes the CO distribution in the upper mesosphere. This might also impact H2O and hence HO2.

We added the following discussion: Lastly, (3) these discrepancies might be related to the WACCM representation of the mean meridional circulation which has been shown to have some deficiencies (Smith et al., 2011; Smith, 2012), suggesting that the gravity wave parametrization needs to be modified. In addition, Garcia et al. (2014) has shown that adjusting the Prandtl number, used to calculate the diffusivity due to gravity waves, significantly alters the CO2 SD-WACCM simulations improving its agreement with satellite measurements. Such adjustment should also affect the H2O and hence the HOx chemistry.

In connection with this and the possible reason mentioned in the manuscript about possible inaccuracies in the representation of the absorption cross sections of H2O and O2 around the Lyman–Alpha region and the Schumann–Runge bands, Garcia et al. (2014) has found that an overestimation of the O2 cross-section in the 105–121 nm wavelength range was causing a too low CO concentration in the upper mesosphere. The large O2 cross-section assumed in the standard WACCM absorbed the UV radiation at high altitudes, preventing its penetration into lower altitudes and hence the CO production from CO2 photolysis. Although this spectral range is just at the edge of the Lyman-alpha, which affects H2O, this might be a reason for the WACCM/MLS discrepancy. With the reduced O2 cross-section, radiation will penetrate deeper, H2O will be more strongly photodissociated and hence producing more OH and more HO2.

It might worth to explore this point.

We decided to leave this discussion out of the study, because it was based plainly in the photochemical model representation of these values but we didn't change the resolution to corroborate the hypothesis.

In connection with this and the possible reason mentioned in the manuscript about possible inaccuracies in the representation of the absorption cross sections of H2O and O2 around the Lyman–Alpha region and the Schumann–Runge bands, Garcia et al. (2014) has found that an overestimation of the O2 cross-section in the 105–121 nm wavelength range was causing a too low CO concentration in the upper mesosphere. The large O2 cross-section assumed in the standard WACCM absorbed the UV radiation at high altitudes, preventing its penetration into lower altitudes and hence the CO production from CO2 photolysis. Although this spectral range is just at the edge of the Lyman-alpha, which affects H2O, this might be a reason for the WACCM/MLS discrepancy. With the reduced O2 cross-section, radiation will penetrate deeper, H2O will be more strongly photodissociated and hence producing more OH and more HO2.

It might worth to explore this point.

We decided to leave this discussion out of the study, because it was based plainly in the photochemical model representation of these values but we didn't change the resolution to corroborate the hypothesis.

Page 22917, par. at lines 17-21. Since the feature discussed in not shown in the presented figures I cannot see the reason for its discussion. I suggest to remove it.

It is there, in the number density subplots, the text was changed to: In Fig. 9 in the number density subplots, between 10 and 0.1 hPa, both the offline MLS dataset and the SD-WACCM simulations behave in a similar manner both in structure and in magnitude; however, due to the small HO2 signal in the MLS radiances, the offline MLS retrieval is noisier.

Page 22918, lines 10-15. It would be useful to mention which solar flux data is used in WACCM and how other data would change (at least qualitatively) the results. The same applies to the 1-D model described in Sec. 4.4 and it is extensive to the parameterization of the cross-sections (see major comment above).
Page 22918, lines 18-20. “For pressure levels smaller than 0.1 hPa, the lack of a clear second peak in the SD-WACCM dataset reflects the smaller mesospheric concentrations in this dataset.” To which “second peak” do you refer? I cannot see it (just see the peak at 0.02 hPa).

The text was changed to: The lack of a clear second peak at \( \lambda_{\text{Lij}} \) 0.02 hPa in the SD-WACCM dataset reflects the smaller mesospheric concentrations in this dataset.

Page 22918, line 27. "poleward" - "towards the winter pole"?

Page 22919, lines 1-2. “and overestimating by as much as 50% over the polar winter regions.” This might be right but I would not conclude that when comparing second (WACCM with AKs) and third (MLS) panels in the left column of Fig. 11. They both appear with the same light green color. BTW, in the percentage differences panels, which WACCM is being used, with or without the applied AK’s? This comment is extensive to all figures where the AKs are applied (e.g. Figs. 7, 8, 9, and 10 as well). We added a paragraph at the beggining of the result section stating that all the comparisons used the MLS averaging kernels: n this section we compare the offline HO2 dataset with balloon-borne and other satellite measurements, as well as, with global climate and photochemical model simulations. In making these comparisons, i.e. when showing the absolute or percentage differences between the datasets, the MLS averaging kernels has been applied to properly compare them. Furthermore, when comparing the global climate or the photochemical model simulations, its high vertical resolution has been reduced to the MLS one using a least square fit as described by Livesey et al. (2011, Sect. 1.9). In these comparisons, no altitude extrapolation has been applied to any dataset.
