Responses to Referee #1:

We thank referee's helpful and constructive comments and review. The referee's comments are listed in *italics*, and our responses in black with revised texts in **bold black**. Please noted that figure numbers are different from those in the original manuscripts.

I. General impression:

As a follow-up to initial validation work by Liu et al. in 2010, this work reports on the comparative validation with respect to AURA MLS measurements of 10 years of SAO OMI nadir ozone profile data. It thus nicely complements a recent validation exercise of the same OMI data with respect to ozonesonde measurements (Huang et al., 2017). The impact of the occurrence of a serious OMI row anomaly in January 2009 is well addressed, and the comparative analysis is insightfully adopted accordingly. The only major thing missing seems to be a clear motivation for the comparison grid that has been used (see details below). Additionally, it is believed that the clarity of the presentation of the results could be improved by slightly adopting some of the figures, and possibly by including a summary table.

II. Specific comments:

In abstract/introduction and throughout the text, please mention the validated SAO algorithm/product version, as is done for the MLS data, for traceability and for comparison with the results presented in Liu et al. (2010).

We have added the version number (v0.93) for the current product. Note that the product of Liu et al. (2010) does not have a version as it is a research product that is not produced routinely, but with very limited spatiotemporal coverage.

Introduction, page 2, lines 35-36: The mentioning of a "suggestion that the possible affection of OMI measurements at shorter wavelengths in the UV-1 channel may have been affected by the RA at all cross-track positions" lacks any notion on how the affection could take place. This seems important however for the succeeding validation motivation. Please provide an appropriate indication of the 'affection' source.

We have added some notion and a reference as follows:

"…been affected **by blockage and solar radiation effects** of the RA… **(Schenkeveld et al., 2017).**"

Section 2: The major motivation for the use of MLS data for comparison is its location on the same platform. This instrument however measures microwave thermal emission, whereas OMI is

a UV/VIS spectrometer. Please provide a brief sketch of how this wavelength mismatch might affect the comparison results.

OMI and MLS have very different vertical sensitivities on the ozone profiles due to the different wavelengths of these sensors deployed for retrievals and mostly due to different observation modes (nadir for OMI vs. limb for MLS). These observation differences result in the different vertical resolutions and grids. Therefore, we also include the comparison between OMI and MLS with applying OMI averaging kernels due to much higher vertical resolution in the MLS data to account for differences in vertical sensitivities.

We have added **"…due to the use of observation modes (nadir for OMI and limb for MLS) and different spectral regions …" in P6 L4 of previous manuscripts.**

Section 3, page 6, lines 1-3: The apparent mismatch between the 0.35 hPa to TOA (line 2) and 0.22 to 261 hPa (line 3) ranges is only explained in lines 23-28 on the same page. Please combine these paragraphs for clarity.

We have combined these paragraphs.

Section 3, page 6, lines 4-22: Although in agreement with Liu et al. (2010), the OMI and MLS profile matching and comparison approaches (still) raise some questions, especially regarding the choice of the MLS grid as the comparison grid: (1) Why opt for the MLS grid as the comparison grid, when OMI is the instrument to be validated? (2) The complementary validation paper using ozonesonde data (Huang et al., 2017) makes use of the OMI grid as the comparison grid, so comparison of results in both papers would be simplified by using the same grid (using the sonde grids would indeed not make sense). (3) Opting for the finer-resolution grid seems to contradict the OMI DFS, which is of the order of 6-7. This means that all grids with a higher resolution than six to seven layers introduce information that is not in the measurement (i.e. coming from the prior). Why then make the grid even finer than the OMI retrieval grid? (4) Taking the MLS grid as the comparison grid increases the number of operations that has to be performed on the data, hence also the uncertainties: - OMI grid comparison: MLS data to OMI grid, AK-smoothing, Comparison - MLS grid comparison: MLS data to OMI grid, AK-smoothing, MLS data to MLS grid, OMI data to MLS grid, OMI prior to MLS grid, Comparison (5) Returning to the MLS grid after smoothing might indeed make the calculation of relative differences using the MLS data as a reference less stable. The authors' choice for the OMI prior profile as reference in the relative difference denominator is legitimate but somewhat unfortunate. Therefore, (a) please provide a strong motivation for the use of the MLS grid as the comparison grid, taking into account the consequences thereof as described above, and (b) possibly add a plot that clarifies the different vertical grids and comparison ranges (as e.g. in Hubert et al. (2016) Fig. 1). An indication of the approximate OMI unit-DFS layers in this plot would be a very nice surplus.

(a)

We have tested the validation by using OMI grid as shown in the figure below. As has been expected, the mean biases and standard deviations with OMI grids are slightly better than those with MLS grids, but both results with OMI and MLS grids have very similar morphology. The clearly larger values with the MLS grid near the top and bottom are due to smaller MLS layers and broader rangers at both ends than cannot be compared with the use of OMI grid. There are a number of reasons to adopt the use of MLS grid in Liu et al. (2010) in addition to that retrievals can be validated at slightly broader altitude ranges.

(1) We want to show that OMI retrievals, despite coarser vertical resolutions (but smaller retrieval errors), compare quite well with MLS data that have much better vertical resolution) even at MLS vertical grid and without applying OMI averaging kernels. This is not to show that OMI have enough vertical resolution comparable to MLS, but show that OMI retrievals, constrained by climatological a priori, agree quite well with MLS data at MLS grid. It should be noted that the a priori already agrees with MLS data to the 10% level in the middle stratosphere. Our retrieval uses this tight a priori constrain to minimize retrieval error although it decreases the DFS/resolution compared to use of a larger a priori constraint.

(2) Using the MLS grid facilitates the comparison of stratospheric ozone columns down to MLS pressure levels like 100 hPa or 215 hPa as MLS stratospheric ozone columns have been used to derive tropospheric ozone column in combination with OMI total ozone. Using the OMI grid, we cannot directly integrate the ozone down to 100 hPa and 215 hPa from the convolved MLS profiles unless we interpolate it back to the MLS grid or use some portion of the original MLS profile that cannot be convolved due to OMI/MLS altitude grid mismatch.

So in this study, we continue the use of OMI grid for contrast with the results in Liu et al. (2010). Also, we would like to cross-validate MLS stratospheric ozone columns down to 261 hPa and MLS ozone from 215 hPa to 261 hPa because the MLS official documents suggest that MLS ozone at 261 hPa still requires further evaluation for scientific use. Using the MLS grid will facilitate these additional comparisons as well.

We added the following sentence to P6 L5 of the original manuscript: "**In contrast to typical profile comparison using the vertical grid of the product with coarser vertical resolution, we conduct the comparison using the finer MLS vertical grid following Liu et al. (2010a) to demonstrate that OMI retrievals compare quite well with MLS data even at the MLS grid and facilitate the cross-validation of ozone from 215 hPa to 261 hPa and stratospheric ozone columns integrated down to several MLS pressure levels.**"
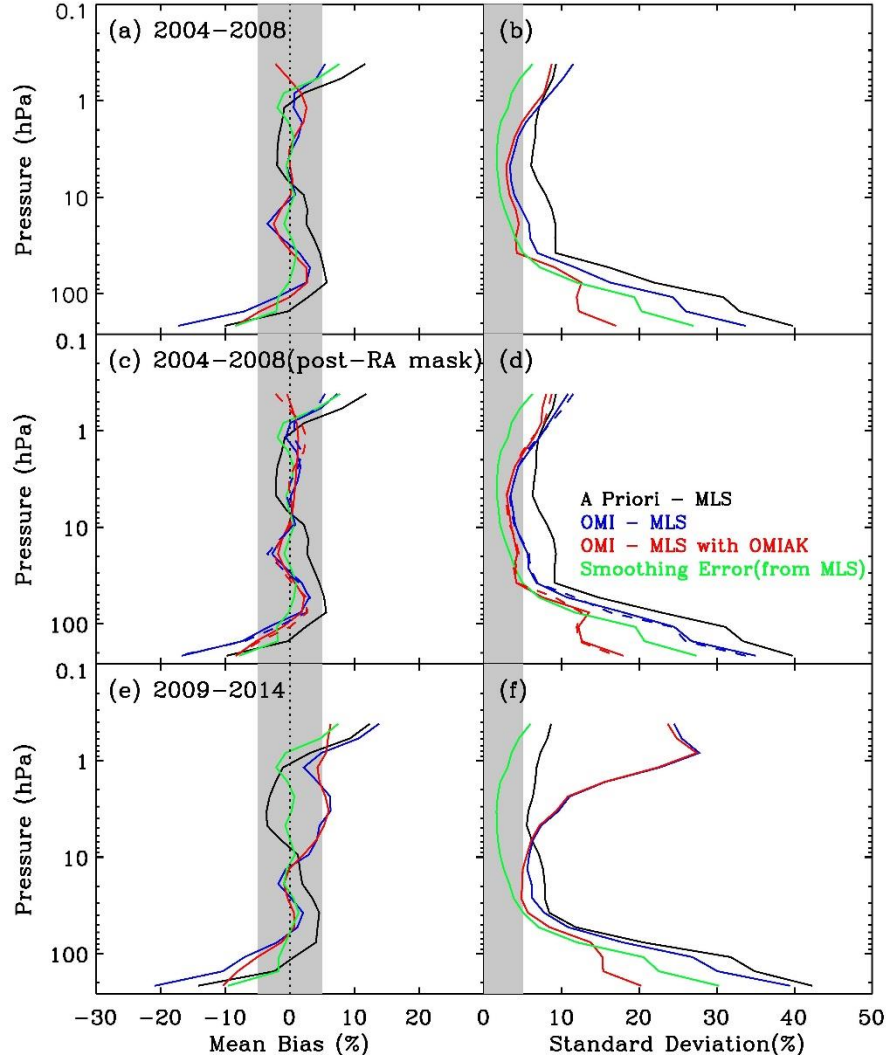
Figure: Similar to Figure 4, but with OMI grids.

(b) We have added Figure 2 to clarify the vertical grids and comparison ranges. Because this paper focuses on the validation, we have added a reference Liu et al. (2010) for readers who are interested in DFL profiles. We have revised it as "More details on the PROFOZ product and ozone profile retrieval algorithm, **including the unit Degree of Freedom for Signal (DFS) profiles,** can be found in Huang et al. (2017) and Liu et al. (2010)."
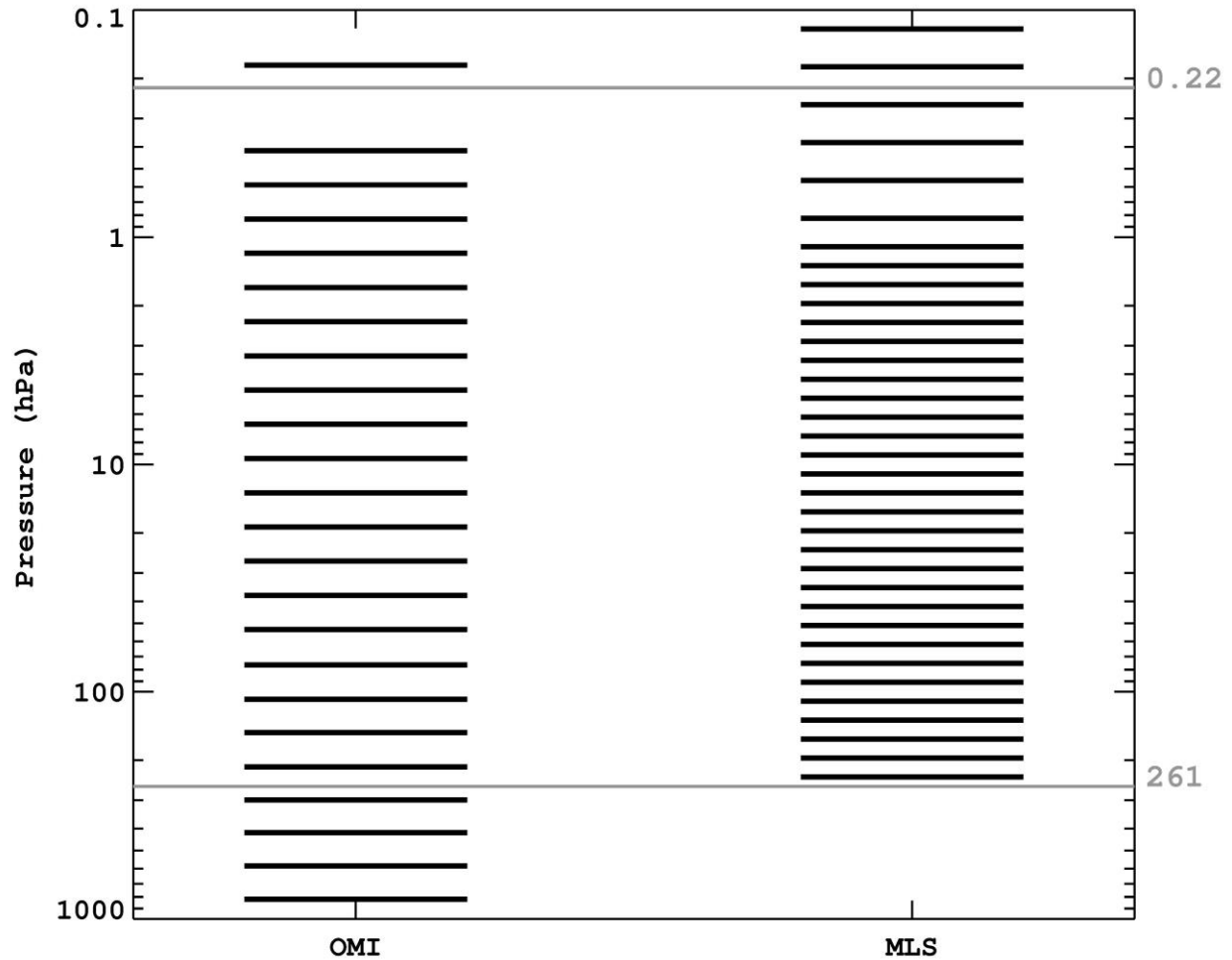
"Figure 2. Overview of the vertical range of OMI and MLS. The solid black bars represent vertical grids of OMI and MLS, respectively, while the validation range of this study is marked by grey lines."

Section 4.1, Fig. 3(e): Please provide an indication of why the OMI bias exceeds the prior bias within the 1-10 hPa range.

This is due to the same reason mentioned shortly afterwards that RA affects radiance in the UV-1 channel. We have added as follows "**In addition, the retrieval biases are larger those in the a priori biases within 1-10 hPa**" before explaining the degradation of the results.

Section 5, page 12, lines 20-21: "The impact of using different coincidence criteria on the ozone profile comparison is negligible" is indeed correct for the different criteria considered in this work, but to what extent (spatio-temporal distance) is this true? Please briefly clarify.

We didn't further test to find out the spatiotemporal extent to which different coincidence criteria can be considered. As we discussed in the first paragraph of Sect. 3, the OMI cross-track position collocated with MLS data varies from position 20 in the tropics to position 15 at high latitudes. After the occurrence of RA, the cross-track positions vary from position 22 in the tropics to position 12 at high latitudes. For most cases, the distance shift between two different coincident criteria is approximately 100-200 km; for some extreme cases, the maximum distance shift is ~300-400 km. The variability of stratospheric ozone is relatively small on average within these distances. However, if the RA affected pixels spread in the future, this assumption would not be necessarily true. Ozone variability and longer distance will increase the standard deviations of the differences. In addition, there are cross-track dependent biases in our product especially for those extreme off-nadir positions as shown in Huang et al. (2017); collocating OMI retrievals at these extreme off-nadir positions with MLS data will further increase the standard deviations of the OMI/MLS differences.

Section 5, page 12, lines 24-26: "The much worse comparison at higher altitudes indicates that RA significantly affects the UV-1 channel of the OMI measurements, although they are not flagged as RA pixels." Based on these results, have you considered introducing an additional L1/L2 flagging/filtering criterion? If so, it would be helpful to mention this.

This validation study will caution the users about using the ozone profile retrievals at high altitudes during the post-RA period. We will suggest to the instrument calibration team to see if they can add this additional flagging in the level 1b data based on instrument calibration. However, we are not considering it by ourselves because of the following reasons. A good flagging criterion should be preferably based on radiances in the level 1b data or instrument calibration process. The UV1 channel is mainly used for ozone profile retrievals, not used in most of the OMI products. In addition, we did not validate retrievals all cross-track positions except for those positions collocated with MLS data. Furthermore, despite of larger biases and standard deviations at higher altitudes, retrievals for pressure > 10 hPa and stratospheric ozone columns are not significantly degraded and can still be useful.

III. Technical corrections:

Introduction, page 2, line 13: Move link to references and provide access details to capture future changes?

Done.

Section 2.1, page 3, lines 31-32: Please use consistent spelling for "cross track positions" or "cross-track positions" throughout the text.

Done. We use "cross-track positions" in this paper.

Section 3, page 6 line 6: Please provide reference to "a procedure provided by the MLS team"

Done. We added Liu et al. (2010a).

Section 5, page 12, line 30: "straylight" instead of "staylight"

Done.

Presentation of results: (1) It might be helpful for the reader to include a comparison result overview table next to the plots. Rows: bias, spread for pre-RA, post-RA for preRA period, post-RA Columns: profiles, 100 hPa SOC, 215 hPa SOC, 261 hPa SOC And possibly: Do these numbers comply with the OMI instrument targets?

The comparison results are all included in Figure 4 and Figure 6, the information would be repeated if we had a table. Generally, all our results comply with the OMI instrument targets. We have added in Sect. 5 as
**"From the OMI scientific requirement document (Levelt et al., 2000), the scientific requirement for OMI ozone is to achieve 10% accuracy in the stratosphere and 30% in the troposphere. The validation of stratospheric ozone profiles with MLS data indicates that our retrieval performance generally meets OMI's scientific requirement in most of the stratosphere especially during the pre-RA period (Levelt, 2000) except for high latitudes/SZAs."** in the 3rd paragraph of Sect. 5.

and

"**Overall, our validation indicates that our retrieval performance in the stratosphere generally meets OMI's scientific requirement for the ozone profile product to have 10% accuracy in the stratosphere (Levelt, 2000).**" in the 5th paragraph of Sect. 5.

(2) Figure 1: Horizontal lines could indicate the ideal colocation windows and enforced post-RA windows for clarity. By use of interpolated colors, the discreteness of the cross-track position is flawed: E.g. "red" values appear through interpolation, but are not in the data.

We added horizontal lines and discrete colors as follows:
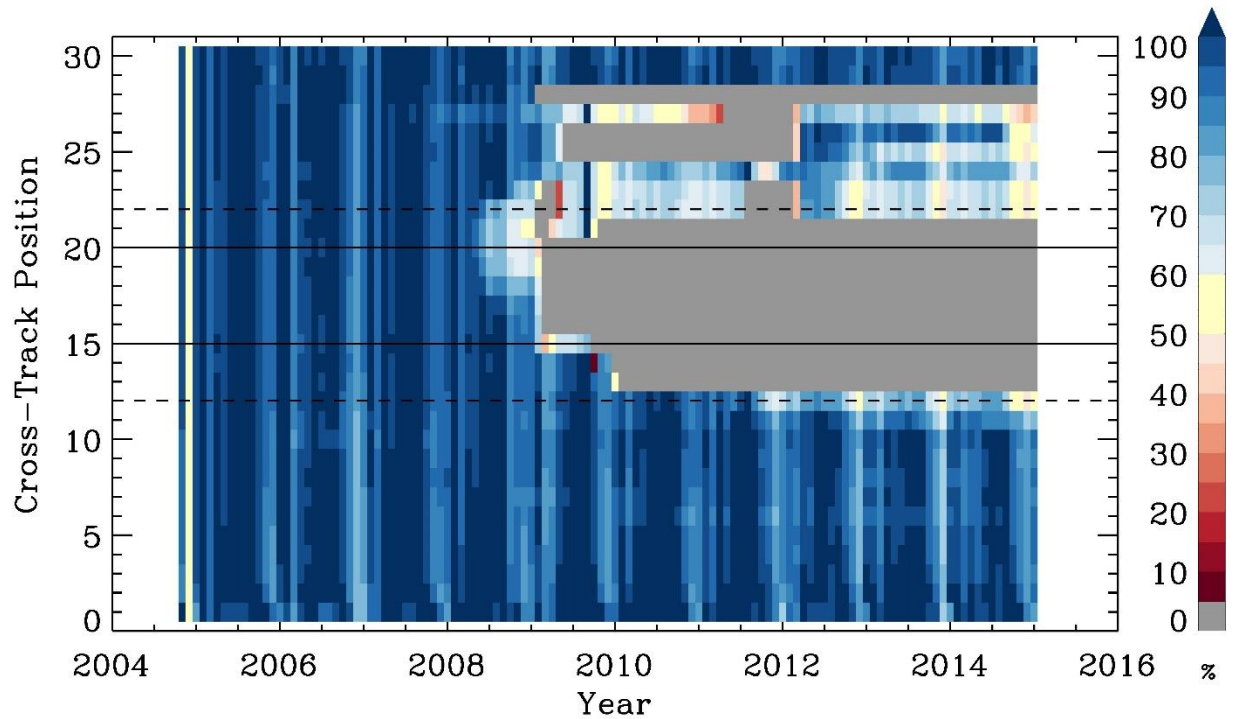
We have also added in the caption as "… **The solid lines represent the range of OMI-MLS collocations that vary from position 20 in the tropics to 15 at high latitudes, while the dashed lines represent the enforced post-RA OMI-MLS coincidence criteria (the closest OMI pixel from position ≤ 12 or ≥ 22)**"

(3) Figure 2: Vertical lines or gray areas that mark the periods that have been left out of the analysis would be helpful for interpretation of the later plots.

We have added gray areas to mark the periods that have been left out as follows:
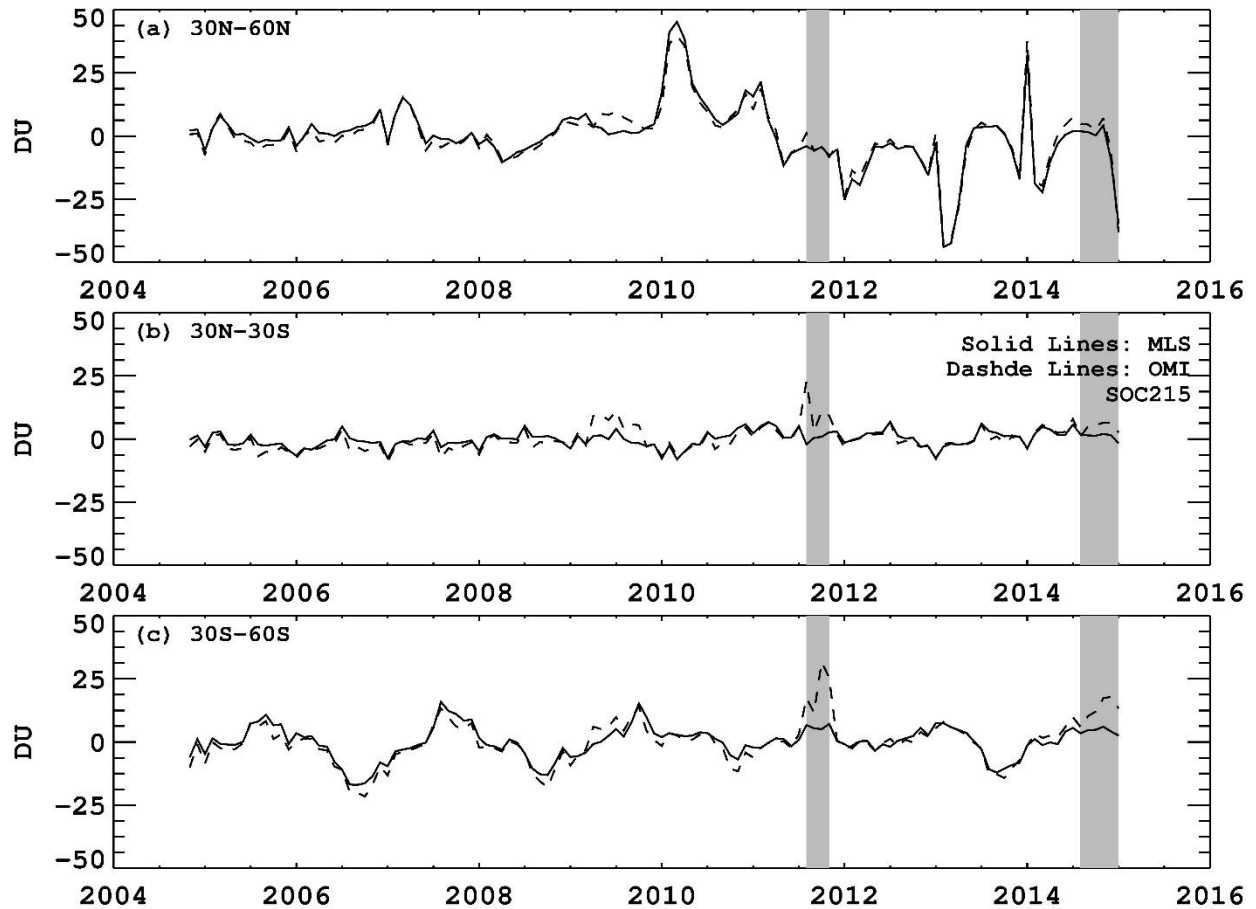
(4) Figures 7 and 8: Standard deviations are often plotted with respect to the mean biases, but are plotted with the zero-line as a reference here. Please mention this in the figure captions.

We have added this in the captions of Figure 8 and 9 as:

"… Note that standard deviations are plotted with the zero-line as a reference."

(5) Figure 9: The gray areas are not specified in the figure caption.

We have added **"… The grey areas indicate the range of -3% and 3% in each panel."**

(6) Figure 10: Please in the text or in the figure caption provide an indication of the OOM of the P-values that are now indicated as "0.00".

Done.

Huang, G., Liu, X., Chance, K., Yang, K., Bhartia, P. K., Cai, Z., Allaart, M., Ancellet, G., Calpini, B., Coetzee, G. J. R., Cuevas-Agulló, E., Cupeiro, M., De Backer, H., Dubey, M. K., Fuelberg, H. E., Fujiwara, M., Godin-Beekmann, S., Hall, T. J., Johnson, B., Joseph, E., Kivi, R., Kois, B., Komala, N., König-Langlo, G., Laneve, G., Leblanc, T., Marchand, M., Minschwaner, K. R., Morris, G., Newchurch, M. J., Ogino, S. Y., Ohkawara, N., Piters, A. J. M., Posny, F., Querel, R., Scheele, R., Schmidlin, F. J., Schnell, R. C., Schrems, O., Selkirk, H., Shiotani, M., Skrivánková, P., Stübi, R., Taha, G., Tarasick, D. W., Thompson, A. M., Thouret, V., Tully, M. B., Van Malderen, R., Vömel, H., von der Gathen, P., Witte, J. C., and Yela, M.: Validation of 10-year SAO OMI Ozone Profile (PROFOZ) product using ozonesonde observations, Atmos. Meas. Tech., 10, 2455-2475, doi: 10.5194/amt-10-2455-2017, 2017.

Levelt, P. F.: Science requirements document for OMI-EOS, KNMI, 2000.

Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., and Kurosu, T. P.: Ozone profile retrievals from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 10, 2521-2537, doi: 10.5194/acp-10-2521-2010, 2010.

Responses to Referee #2:

We thank referee's helpful and constructive comments and review. The referee's comments are listed in *italics*, and our responses in black with revised texts in **bold black**. Please note that figure number has been revised according to Referee #1's comments and suggestions.

This paper is well thought out and contains a thorough analysis of the differences between OMI and MLS profile retrievals. The writing can be somewhat "dense" at times and this reviewer suggests that some of the long, highly complex sentences be split into two to make the reading less difficult. Other than a few minor changes listed below, this manuscript is recommended for publication.

Well, all my underlines and color have disappeared. I hope that you can follow my changes.....

Minor changes: Page 1: line 17-19 remove the words "larger" and "smaller" Larger than what?

This sentence shows the comparison results of the post-RA period with relative to those during the pre-RA period as we show the comparison results during the pre-RA period in the previous two sentences. To avoid confusion, we have changed the sentence to "**Compared to the retrievals during the pre-RA period, OMI retrievals during the post-RA period** …"

Line 20,23 & 25 comparisons.

Done.

Line 25 significant bias

Done.

in the Line 28-9 The sentence about 261hPa MLS ozone sounds very "arrogant" as if MLS is being validated with OMI and not the other way around. Suggestion- just state that they agree well at this pressure and leave out the interpretation as you have done on Page 5.

We deleted this sentence, because we already have the agreement statement in the abstract.

Page 2: line 11 change 'in' to 'at' Line 29 ozonesondes measure Line 32 remove "but also....validated"

Done.

Page 3 line 22 comparisons

Done.

Page 4, line 21 remove "in the stratosphere".

Done.

Line 21-22: Please either explain how the cross track position changes as a function of latitude here or refer to section 3.

We have added "The **threshold of** cross-track position **…(More details will be discussed in Sect. 3).**"

Page 6, line 1 change is to are (data is plural) Line 2 to the top of Line 3 change "for avoiding" to "to avoid" Line 20: OMI a priori is used in the calculation Line 24: The ozone column Line 27: remove "SOC comparison" and add an 's' to "are for comparisons"

Done.

Page 7, Line 2,3 Remove "only, missing to......than 7hPa"

Done.

Line 13:mask which is consistent

Done.

Line 32: (red lines), the OMI

Done.

Page 8 line 3: change to: "these comparisons are similar to the OMI/MLS comparisons shown in 2006 in Lui et al although both OMI and MLS versions are now different and this study is done....."

Done.

Line 15: change "Such worse comparison" to "The larger differences"

Done.

Line 24-5: please reference cross track biases or include a plot.

We have added (Huang et al., 2017) as a reference.

Line 33: This supports the theory that....

Done.

Page 9: line 11: remove "especially"

Done.

Page 10: Line 30: as in the introduction, please reduce the conclusion to "the two agree at 261 level"

We have removed the sentence and added to the previous sentence **"indicating the scientific use of MLS ozone at 261 hPa in applications such as the OMI/MLS TOR method"**

Page 11: line 1 add "see Figure 2" after "trend analysis" Line 7 & 8 change "of" to "at"

Done.

Line 10- you need stronger words to dissuade people from using the data in the upper strat from 30-90N

We have revised it as "**… where there are significantly large trends of up to -5%/year…**"

Line 24: ...not suitable for trend studies.... This is a conclusion and should be either moved or repeated in the conclusion section.

We have rephrased it in the conclusion section as **"These significant bias trends indicate that the current ozone profile product is not suitable for trend studeis, especially during the post-RA period."**

Page 12: line 10: profiles

Done.

Page 13: remove "original"

Done.

Figure 4 & 5: Why is there such a positive bias in 70-90 South above 1 and below 100 hPa plot in figure 5 (SZA) but not figure 4?? Shouldn't it "smear out" and be a red streak in the latitude plot like in the north high latitudes (Fig 4)? Please explain.

The large positive mean biases with 85°-90° solar zenith angles in the southern hemisphere are due to their large solar zenith angles. In Figure 4, mean biases at southern high latitude regions include not only large mean biases due to the large solar zenith angles during southern winter, but also smaller and/or negative mean biases with smaller solar zenith angles during southern summer. Naturally, large positive biases with large solar zenith angles during southern hemisphere winter are smeared out by more frequent smaller or negative biases with smaller solar zenith angles during other seasons.   Also the figure shows very similar patterns between south and north high latitudes.

Figure 6 is very "cluttered" with text. Please remove N= for the lower two plots as it is redundant information.
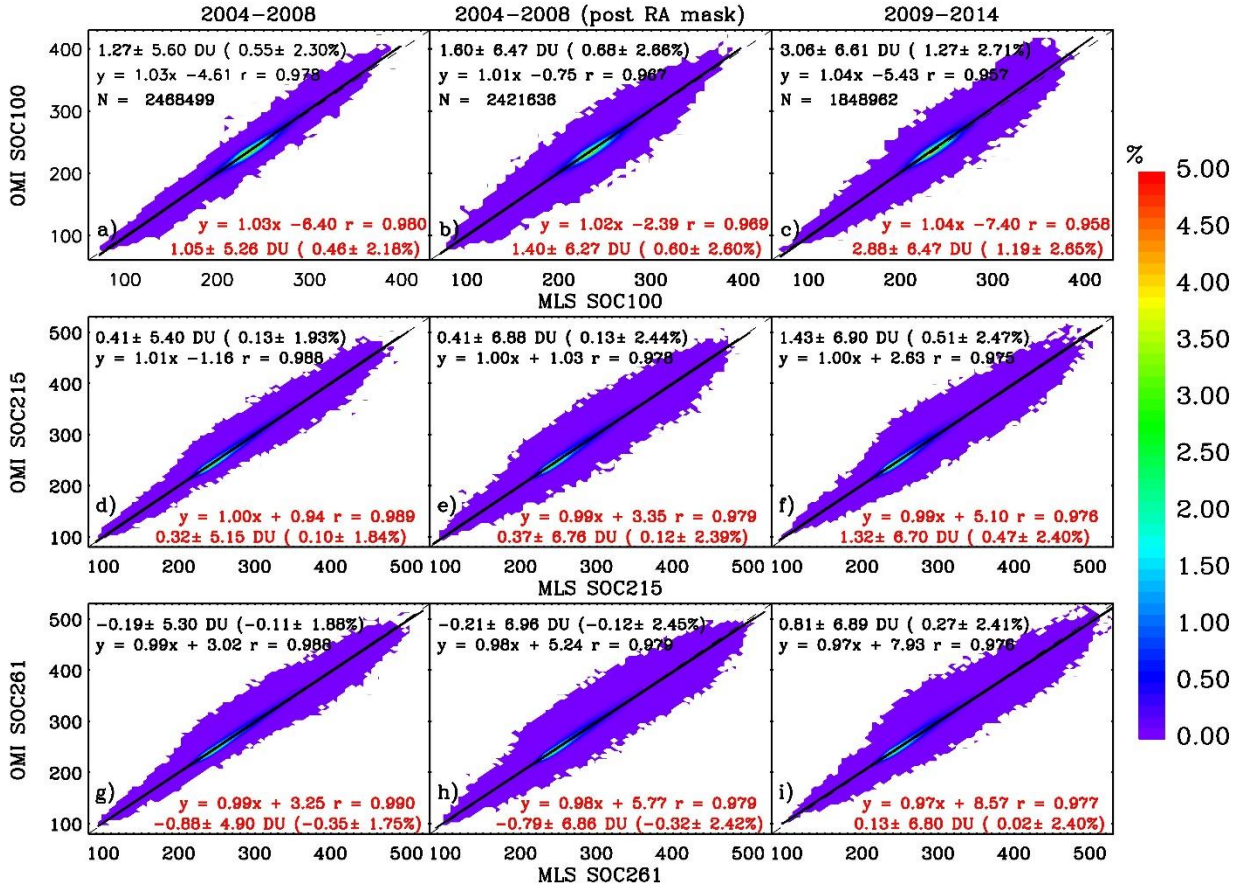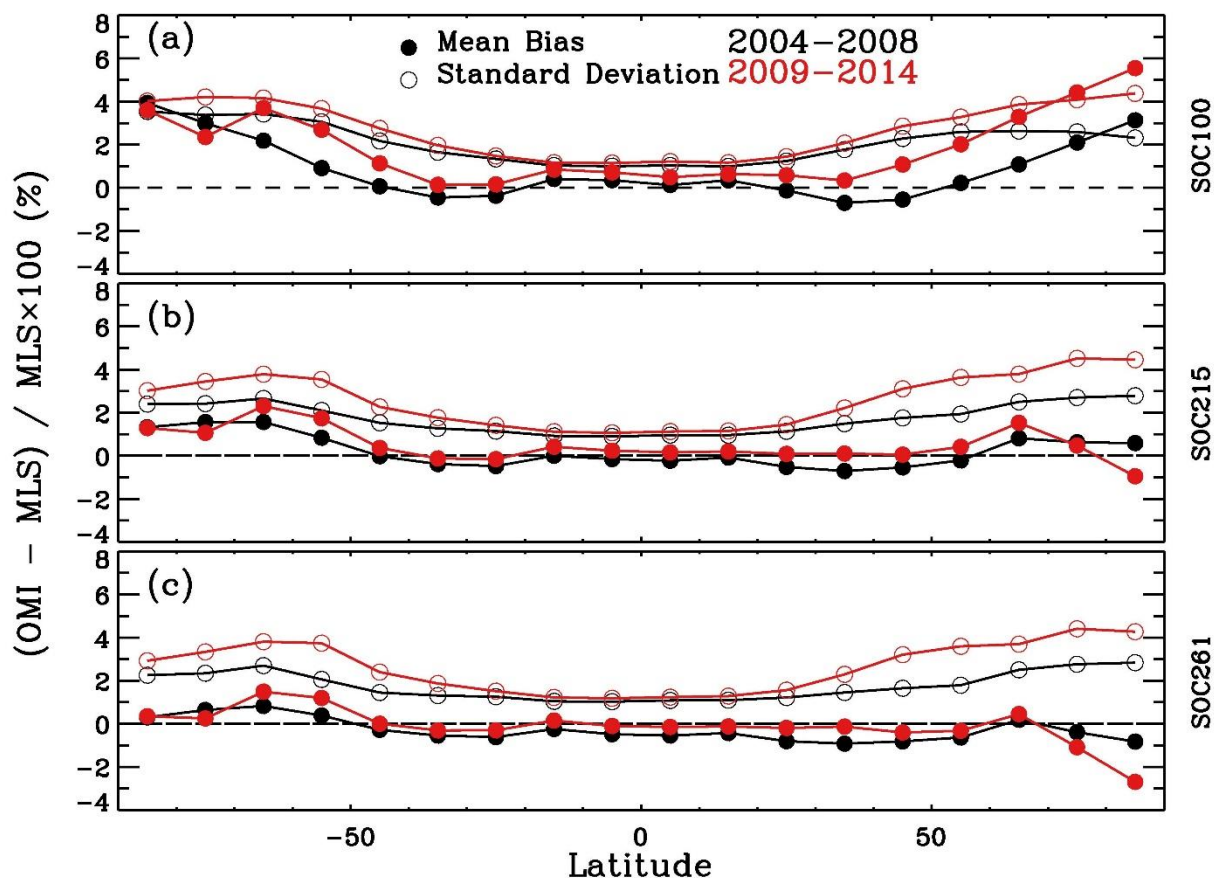
Done.

Figure 7: Please scale the middle plot to the same absolute scale as the other two (-4 to 6 or -3 to 7 would be fine)

Done.

(a)

Mean Bias 2004−2008
Standard Deviation 2009−2014

SOC100

(b)

SOC215

(c)

SOC261

(OMI − MLS) / MLS×100 (%)

Latitude

−50   0   50

Huang, G., Liu, X., Chance, K., Yang, K., Bhartia, P. K., Cai, Z., Allaart, M., Ancellet, G., Calpini, B., Coetzee, G. J. R., Cuevas-Agulló, E., Cupeiro, M., De Backer, H., Dubey, M. K., Fuelberg, H. E., Fujiwara, M., Godin-Beekmann, S., Hall, T. J., Johnson, B., Joseph, E., Kivi, R., Kois, B., Komala, N., König-Langlo, G., Laneve, G., Leblanc, T., Marchand, M., Minschwaner, K. R., Morris, G., Newchurch, M. J., Ogino, S. Y., Ohkawara, N., Piters, A. J. M., Posny, F., Querel, R., Scheele, R., Schmidlin, F. J., Schnell, R. C., Schrems, O., Selkirk, H., Shiotani, M., Skrivánková, P., Stübi, R., Taha, G., Tarasick, D. W., Thompson, A. M., Thouret, V., Tully, M. B., Van Malderen, R., Vömel, H., von der Gathen, P., Witte, J. C., and Yela, M.: Validation of 10-year SAO OMI Ozone Profile (PROFOZ) product using ozonesonde observations, Atmos. Meas. Tech., 10, 2455-2475, doi: 10.5194/amt-10-2455-2017, 2017.

Validation of 10-year SAO OMI Ozone Profile (PROFOZ) Product Using Aura MLS Measurements

Guanyu Huang[1], Xiong Liu[1], Kelly Chance[1], Kai Yang[2], Zhaonan Cai[1]

1. Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
2. Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA

Correspondence to: Guanyu Huang (guanyu.huang@cfa.harvard.edu)

Abstract. We validate the Ozone Monitoring Instrument (OMI) ozone profile (PROFOZ v0.93) product including ozone profiles between 0.22-261 hPa and Stratospheric Ozone Columns (SOCs) down to 100, 215, and 261 hPa from October 2004 through December 2014 retrieved by the Smithsonian Astrophysical Observatory (SAO) algorithm against the latest Microwave Limb Sound (MLS) v4.2x data. We also evaluate the effects of OMI row anomaly (RA) on the retrieval by dividing the data set into before and after the occurrence of serious RA, i.e., pre-RA (2004-2008) and post-RA (2009-2014). During the pre-RA period, OMI ozone profiles agree very well with MLS data. The global mean biases (MBs) are within 3% between 0.22-100 hPa and negative 3-9% for lower layers, and the standard deviations (SDs) are 3.5-5% from 1-40 hPa, 6-10% for upper layers and 5-20% for lower layers, after applying OMI averaging kernels to MLS data. OMI shows latitude and solar zenith angle (SZA) dependent biases, but MBs and SDs are mostly within 10% except for low/high altitudes of high latitudes/SZAs. Compared to the retrievals during the pre-RA period, ~~During the post RA period,~~ OMI retrievals during the post-RA period degrade slightly between 5-261 hPa with MBs and SDs typically larger by 2-5%, and degrade much more, with larger MBs by up to 8% and SDs by up to 15% for pressure less than ~5 hPa, where the MBs are larger by 10-15% south of 40°N due to the blockage effect of RA and smaller by 15-20% north of 40°N due to the solar contamination effect of RA. The much worse comparisons at high altitudes indicates the UV-1 channel of pixels that are not flagged as RA is still affected by the RA. During the pre-RA period, OMI SOCs show very good agreement with MLS data with global mean MBs within 0.6% and SDs of 1.9% for SOCs down to 215 and 261 hPa and of 2.30% for SOC down to 100 hPa. Despite clearly worse ozone profile comparisons during the post-RA period, OMI SOCs only slightly degrade, with SDs larger by 0.4-0.6% mostly due to looser spatial coincidence criterion as a result of missing data from RA and MBs larger by 0.4-0.7%. The retrieval comparisons indicates significant bias trends, especially during the post-RA period. The spatiotemporal variation of the retrieval performance suggests the need to improve OMI's radiometric calibration to maintain the long-term stability and spatial consistency of the PROFOZ product. ~~The good comparison with SOC down to 261 hPa supports that MLS ozone at 261 hPa, recommended for further evaluation by the MLS team, is suitable for scientific use.~~

1 Introduction

The Dutch-Finnish built Ozone Monitoring Instrument (OMI) on board the NASA Earth Observing System (EOS) Aura satellite has been making useful measurements of trace gases including ozone and aerosols since October 2004. As for ozone, there are two independent operational total ozone algorithms (Bhartia and Wellemeyer, 2002; Veefkind et al., 2006) and two ozone profile algorithms. Of the two ozone profile algorithms, one is the operational algorithm developed at KNMI (van Oss et al., 2001) producing the OMO3PR product, and the other one is a research algorithm developed at Smithsonian Astrophysical Observatory (SAO) by Liu et al. (2010b) producing the PROFOZ product. Both KNMI and SAO algorithms use the optimal estimation technique to retrieve ozone profiles from the spectral range 270-330 nm, but they have significantly different implementation details (e.g. radiometric calibration, radiative transfer model simulation, a priori constraint, retrieval grids, and additional retrieval parameters) (Liu et al., 2010b). The PROFOZ product has been produced at~~in~~ the OMI operational Science Investigator-led Processing System (SIPS) for the entire OMI data record. This product is publicly available at the Aura Validation Data Center (AVDC) (NASA, 2013) ~~(http://avdc.gsfc.nasa.gov/index.php?site=2045907950)~~. In the prequel of this paper, we performed a comprehensive and global assessment of the long-term quality of PROFOZ product (surface to ~ 7 hPa) using ozonesonde observations and showed good agreement with sondes especially in the tropics and mid-latitudes, with mean biases less than 6% and standard deviations of 5-10% for pressure less than 50 hPa and up to 18 (27%) for altitudes below in the tropics (mid-latitudes), despite time-dependent biases especially after the occurrence of serious OMI Row Anomaly (RA) in January 2009 and some biases depending on latitude, season, solar zenith angle (SZA) and cross-track positions (Huang et al., 2017). In this study, we complement the ozonesonde validation of this product with stratospheric ozone data measured by Microwave Limb Sounder (MLS) also aboard the Aura satellite.

Although we used many ozonesondes globally in our previous OMI validation study, there were a few limitations. First, the number of ozonesonde observations is limited and their geographical and temporal samplings are uneven. Only ~10,500 out of ~27,000 ozonesonde profiles are eventually used after OMI/ozonesonde data screening and cloud filtering. Most of them are in the northern mid-latitudes and the tropics, with much fewer observations in the southern middle and high latitudes. Second, the accuracy of ozonesonde observations depends on data processing technique, sensor solution, and instrument type, and the availability of correction factors (Smit et al., 2007; Thompson et al., 2007). Consequently, station-to-station biases may add uncertainties to the OMI ozone validation results (Huang et al., 2017; Worden et al., 2007). Third, ozonesonde**s** measure~~s~~ only from the surface up to ~7 hPa and ozonesondes' burst pressures can vary from sonde to sonde. We used sondes that burst at 7-12 hPa and validated Stratospheric Ozone Columns (SOCs) integrated from tropopause to the OMI layer below sonde burst altitude. Therefore, the top part of OMI ozone profile was not evaluated, ~~but also SOCs were not entirely validated~~ due to the missing ozone information above burst altitudes, which on average consists of ~14% of SOC, ranging from ~6-33% (Huang et al., 2017). This is also why our ozonesonde validation paper focused on ozone evaluation in the troposphere. Furthermore, it has been suggested that OMI measurements at shorter wavelengths in the UV-1 channel may have been affected by blockage and solar radiation effects of the RA at all cross-track positions

2

including those not flagged as RA pixels (~~Personal communication with S. Marchenko~~(Schenkeveld et al., 2017)~~)~~. As the shorter part of radiance spectra in the UV-1 channel mainly contributes to ozone retrieval in the middle and upper stratosphere, it is necessary to evaluate ozone at this altitude range to understand the RA impacts on the UV-1 channel. Therefore, using ozonesondes only cannot fully validate the retrieval quality of the PROFOZ product, understand the impacts of RA, and assess the radiometric calibration over the entire UV channel.

MLS has been measuring stratospheric ozone since its launch in 2004. This stratospheric ozone product has been shown to have high accuracy and long-term stability by using multiple observations (Froidevaux et al., 2008; Hubert et al., 2016; Jiang et al., 2007). MLS v2.2 and v3.3 ozone data down to 215 hPa have been used to evaluate earlier versions of our OMI ozone profile retrievals as well as SOCs down to 100 hPa and 215 hPa (Bak et al., 2013; Bak et al., 2016; Liu et al., 2010a) for limited time periods (1 year or less) and were demonstrated to be an excellent source to validate OMI stratospheric ozone profiles due to MLS's close collocation with OMI, finer vertical resolution and high quality. In this study, we applied the same methodology to validate stratospheric ozone profiles and SOCs of 10-year PROFOZ product (v0.93) with the recently released MLS v4.2x data, quantifying OMI stratospheric ozone biases spatiotemporally, assessing its long-term performance, and the impacts of RA on the retrievals. In this new version, the pressure range of useful data has been extended from 215 hPa down to 261 hPa although the ozone at 261 hPa is still marked as "requires further evaluation" (Livesey et al., 2015). As MLS stratospheric ozone profiles and SOCs have been used to derive tropospheric ozone column (TOCs) using the Tropospheric Ozone Residual (TOR) method (Jing et al., 2006; Schoeberl et al., 2007; Yang et al., 2007; Ziemke et al., 2006; Ziemke et al., 2014), this extension of useful data range to lower altitudes of the critical Upper Troposphere and Lower Stratosphere (UTLS) region has potential to improve the derived TOCs using the various OMI/MLS TOR approaches. Therefore, we extend the ozone profile comparison down to 261 hPa and compare SOC down to 261 hPa from top of the atmosphere. Such comparisons will provide a quantitative evaluation of the quality of MLS data at this lower level relative to those at higher levels (e.g., 215 hPa). This paper is organized as follows: Section 2 describes OMI and MLS data; the validation methodology is introduced in Sect. 3; Section 4 presents results, analysis, and discussions; Section 5 concludes this study.

2 Data

2.1 OMI and OMI Ozone Profile Retrievals

OMI is a nadir-viewing pushbroom UV/visible spectrometer aboard the NASA EOS Aura satellite that was launched into a sun-synchronous orbit in July 2004. It measures backscattered radiances in three channels covering the 270-500 nm wavelength range (UV-1: 270-310 nm, UV-2: 310-365 nm, visible: 350-500 nm) at spectral resolutions of 0.42-0.63 nm (Levelt et al., 2006) with daily global coverage. There are 60 cross-track positions for UV2 and visible channels, and 30 cross-track positions for UV1 channel due to the weaker signals. The nadir spatial resolution of 13 km × 24 km (along × across track) for UV2 and visible channels, and 13 km × 48 km for UV1 channels. Our PROFOZ product is processed by coadding 4 UV-1 pixels along the track, resulting in product nadir spatial resolution of 52 km × 48 km. More details on the PROFOZ product and ozone profile retrieval algorithm,

3

including the retrieval Averaging Kernels (AKs) and profiles of Degree of Freedom for Signal (DFS), can be found in Huang et al. (2017) and Liu et al. (2010b).

As aforementioned, certain cross-track positions in OMI data have been affected by the RA since June 2007 (Kroon et al., 2011). The RA has spread to other rows and shifted with time since January 2009, and it became more serious that some of these positions are not recommended for scientific studies. Figure 1 shows the monthly mean number of OMI ozone profile retrievals relative to that in January 2005 as a function of UV-1 cross-track position and time (from October 2004 to December 2014), indicating the spatial and temporal distribution of RA affected pixels. The OMI pixels between 13 and 21 cross-track positions are mostly not processed due to RA flagging after 2009. Positions 23-27 were partially (a portion of an orbit) affected by RA except that positions 25-27 were mostly affected by RA during 2009-2011 and positions 22-23 were mostly affected by RA during summer and fall 2011. The effects of RA on the lack of retrievals at certain positions are significant to our OMI/MLS coincidences and will be further discussed in Sect. 3.

To screen out OMI profiles for validation, we applied the same criteria for the retrieval fitting quality as Huang et al. (2017) i.e., applying time-dependent thresholds and selecting retrievals with Root Mean Squire (RMS) of the fitting residuals smaller than the sum of monthly mean RMS and its 2σ. However, the other criteria regarding cloudiness, SZA, and cross-track position are different from Huang et al. (2017). The threshold of effective cloud fraction is removed due to its limited impacts on stratospheric ozone retrievals (Liu et al., 2010a). The SZA threshold is changed from less than 75° to less than 88°, same as the SZA criterion used by Liu et al. (2010a). At SZA larger than 75°, although the retrieval sensitivity to tropospheric ozone is significantly reduced to almost 0, reduction in stratospheric ozone retrieval sensitivity due to reduced signal is offset by increased vertical sensitivity to stratospheric ozone as a result of longer path length in the stratosphere. The threshold of cross-track position is also removed because MLS collocates with OMI at an almost fixed cross-track position although the position varies with latitude (More details will be discussed in Sect. 3).

2.2 MLS and MLS ozone retrievals

MLS, co-located with OMI aboard the NASA EOS-AURA satellite, is a forward-looking microwave limb sounder that measures thermal emission at millimeter and sub-millimeter wavelengths to observe vertical profiles of atmospheric trace gases, temperature, pressure, and other constituents. It takes measurements along-track, and performs 240 limb scans per orbit, providing ~3500 profiles daily during both daytime and nighttime (Waters et al., 2006). Measurements are taken 7 min ahead of OMI for the same locations during daytime orbital tracks.

The MLS v2.2 products have been validated to be highly accurate using multiple correlative measurements and have been widely used in many studies (Froidevaux et al., 2008; Jiang et al., 2007; Livesey et al., 2008). In MLS v3.3/3.4 versions, stratospheric ozone profiles are generally very similar except that the profile is reported on a finer vertical grid and the bottom pressure level with scientifically useful value increases from 215 hPa to 261 hPa. Ozone in the upper troposphere shows smaller biases under clear-sky conditions but more vertical oscillations under thick cloudy conditions. The vertical resolution increases from ~3 km to 2.5-3 km at most stratospheric altitudes. The across-track resolution is kept ~6 km, but the along-track resolution has been updated from ~200 km in v2.2 to

300-450 km in v3.3/3.4 versions, depending on altitude (Livesey et al., 2015). The latest MLS version 4.2x ozone product, released in February 2015, is used in this paper for the validation of our PROFOZ product. MLS v4.2x ozone profiles are also generally very similar to previous versions. One of the major improvements of MLS v4.2x is the handling of contamination from cloud signals in trace gas retrievals that results in significant reduction in the number of spurious MLS profiles in cloudy regions and a more user-friendly and efficient screening of cloud-contaminated measurements. Furthermore, the MLS ozone products have been improved through additional retrieval phases and reduction in interferences from other species (Livesey et al., 2015).

The recommended altitude range of MLS v4.2x data is 0.02–261 hPa, but ozone at the 261 hPa level still requires further evaluation. We use MLS 0.22–261 hPa to compare with our OMI PROFOZ product. The comparison at 261 hPa serves as a cross-evaluation of both OMI and MLS data. According to the data screening criteria of the MLS v4.2 data document (Livesey et al., 2015), we use only profiles with even values of the "Status" field, "Quality" fields greater than 1.0 and "Convergence" fields less than 1.03. The vertical and horizontal resolutions are almost the same as those in v3.3/3.4, mentioned above in this section. The precision is estimated to be ~7-30 % (0.1 - 0.2 ppmv) at 0.2 -1 hPa, 2-3% (0.06 -0.15 ppmv) at 2-46 hPa, and 0.02-0.04 ppmv (3-100% due to the high ozone variation) at 68 – 215 hPa. The accuracy is estimated to be 0.05-0.3 ppmv (~3-10%) in most of the stratosphere (0.2-68 hPa), approximately 0.02-0.05 ppmv + 5-20% in the 100-215 hPa region. The precision of the ozone column (for a single profile) down to 100-215 hPa is approximately 2% or less and the accuracy is estimated to be 4% (Livesey et al., 2015).

3   **Methodology**

The ideal OMI-MLS coincident criterion is that the center of an MLS footprint lies within a collocated OMI footprint, same as in Liu et al. (2010a). Then the spatial difference arises from unequal horizontal resolutions: 52 km × 48 km for OMI vs. 300-500 km × 6 km for MLS, and the time difference is 7 minutes. The OMI cross-track position collocated with MLS data varies with latitude, ranging from UV-1 position 20 in the tropics to position 15 at high latitudes (Liu et al., 2010a). Before the occurrence of serious RA (2004-2008, pre-RA period), we can always find the ideal OMI-MLS collocation. However, after the serious occurrence of RA (2009-2014, post-RA period), retrievals for some of these positions are not available as shown in Fig. 1. Consequently, we must downgrade our OMI-MLS coincident criterion to nearest OMI footprint from the center of an MLS footprint. In the tropics, the collocated OMI position changes to position 14 during 2009-2010 and to position 12 after 2010; at high latitudes, the collocated OMI position changes to position 22 after 2009 except for position 24 during summer and fall 2011. These different OMI-MLS coincidence criteria between pre-RA and post-RA periods, and the cross-track dependent retrieval biases as shown in Huang et al. (2017) might influence our evaluation of the OMI long-term stability. Similar to Huang et al. (2017), we will conduct the comparison for both pre-RA and post-RA periods, respectively to evaluate the impacts of RA on the retrievals. To find out whether our validation is affected by different coincident criteria, we will also apply the same post-RA coincidence criterion (position 12 in the tropics and 22 at high latitudes) to pre-RA measurements by masking out positions 13-21 (i.e., pre-PA period with post-RA

5

mask). The comparison differences between using ideal coincidence and post-RA coincidence for the pre-RA period will provide the impacts of coincidence criteria on the comparison.

~~Although MLS v4.2x ozone data in the vertical range 0.02-261 hPa is recommended for scientific use, the top layer in our OMI retrievals is a broad layer from 0.35 hPa to top of the atmosphere. We mainly use MLS data to validate our retrievals from 0.22 to 261 hPa for avoiding the large interpolation errors in this broad layer.~~

MLS and OMI ozone profiles have different vertical grids and resolutions due to the use of observation modes (nadir for OMI and limb for MLS) and different spectral regions. In contrast to typical profile comparison using the vertical grid of the product with coarser vertical resolution, we conduct the comparison using the finer MLS vertical grid following Liu et al. (2010a) to demonstrate that OMI retrievals compare quite well with MLS data even at the MLS grid and facilitate the cross-validation of ozone from 215 hPa to 261 hPa and stratospheric ozone columns integrated down to several MLS pressure levels. we follow the approach used in Liu et al. (2010a) to account for these vertical differences. MLS ozone profiles in volume mixing ratios are integrated to partial ozone columns using a procedure provided by the MLS team (Liu et al., 2010a). To compare OMI retrievals with original MLS data directly, OMI partial ozone columns are integrated to the MLS vertical grids. To account for the different resolutions, MLS partial ozone columns are first interpolated to OMI vertical grids, and then degraded to the OMI vertical resolution by assuming MLS data to be the truth (X_{MLS}), and simulating the expected retrieval (X'_{MLS}) from our OMI algorithm using OMI ~~averaging kernels (~~AKs):

$$X'_{MLS} = X_a + A(X_{MLS} - X_a) \tag{1}$$

where X_a is the a priori ozone profile used in OMI retrievals and A is the AK matrix. The differences between X'_{MLS} and X_{MLS} are the estimated OMI smoothing errors with relative to MLS data, although we note that errors in MLS data could affect the estimates. The convolved MLS data are then interpolated back to the original MLS grids and compared with OMI retrievals interpolated to the MLS grids. In addition to the comparisons between OMI and convolved MLS data, we also compare OMI with original MLS data. The comparison with original MLS data indicates how well our retrievals can represent actual ozone data. The comparison with convolved MLS data removes the OMI smoothing error-related component of the differences between MLS profiles and OMI a priori, which is significant in the UT/LS region, and therefore allows us to better identify other sources of OMI/MLS errors. The normalized difference is defined as (OMI-MLS)/OMI a priori $\times$ 100%. OMI a priori is used in the calculation instead of MLS values, because the relative differences with respect to MLS data could be unrealistically large due to the small MLS values in the tropical upper troposphere and lower stratosphere (Liu et al., 2010a).

Although MLS v4.2x ozone data in the vertical range 0.02-261 hPa are recommended for scientific use, the top layer in our OMI retrievals is a broad layer from 0.35 hPa to the top of the atmosphere. We mainly use MLS data to validate our retrievals from 0.22 to 261 hPa to avoid the large interpolation errors in this broad layer as shown in Figure 2. For validation of OMI SOCs, original/convolved MLS SOCs are integrated from original/convolved MLS ozone profiles between 0.22 and corresponding pressure levels (100 hPa, 215 hPa or 261 hPa). The ozone column from 0.22 hPa to the top of the atmosphere, which is included in the OMI SOCs, is generally less than 0.1-0.2 DU, and therefore negligible (Liu et al., 2010a). We compare OMI SOCs with both original and convolved MLS SOCs

6

although most of the ~~SOC comparison~~ results are for comparisons with original MLS data as the smoothing errors in SOCs are relatively small.

Although we have applied RA flagging in the OMI level 1b data, we notice unusually large retrieval anomalies during several time periods. Figure 3 ~~Figure 2~~ shows the deseasonalized time series of monthly mean OMI and MLS SOC215 in 60°N-30°N, 30°N -30°S and 30°S-60°S, respectively. It clearly shows large positive OMI retrieval anomalies of up to 20 DU (10 DU) in the tropics and southern mid-latitudes during July - October 2011 (July - December 2014), and smaller but still noticeable positive anomalies at all latitudes during March - October 2009. The reason for such sporadic large retrieval anomalies in cross-track positions not flagged as RA pixels is still not clear, but likely associated with the impacts of RA. And it is recommended not to use the data over these periods. Consequently, we exclude the use of OMI PROFOZ products during these three time periods in this validation. However, these anomalies are not readily detected in our OMI/ozonesonde comparison (Huang et al., 2017). This is probably because the number of OMI-ozonesonde pairs is much smaller than that of OMI-MLS pairs during these time periods, and the OMI/sonde SOCs are integrated from tropopause up to ~7 hPa. ~~only, missing to include ozone at pressure less than 7 hPa.~~

For both profile and SOC comparisons, we will first show global comparison during pre-RA and post-RA periods, respectively. The pre-RA comparison will be done using both ideal collocation criterion and post-RA mask. Then the comparisons will be done as a function of latitude, and SZAs in the southern and northern hemispheres, in which results with the ideal coincidence criterion is applied for the pre-RA period.

To evaluate the long-term performance of our ozone profile retrievals, we analyze the monthly mean biases (MBs) of the OMI/MLS differences as a function of time for five different latitude bands: northern high latitudes (90°N-60°N), northern mid-latitudes (60°N-30°N), tropics (30°N-30°S), southern mid-latitudes (30°S-60°S), and southern high latitudes (60°S-90°S). We then derive linear regression trends for the entire period (2004-2014), the pre-RA (2004-2008), and post-RA (2009-2014) periods, respectively. In this evaluation of long-term performance, the pre-RA comparison is done using the post-RA mask, which is consistent with the comparison during the post-RA period.

4   Results and Discussions

4.1 Comparison of OMI and MLS Ozone Profiles

Figure 4Figure 3 (a) and (b) shows OMI/MLS ozone profile comparison over the globe during the pre-RA period (2004-2008) with ideal coincidence criterion. When compared with the original MLS profiles (blue lines), OMI agrees well with MLS, with global mean biases (MBs) within 6% from 1-150 hPa, positive biases of up to 12% for upper layers and negative biases of up to 25% for the bottom layers. The corresponding standard deviations (SDs) are within 4-7% from 1-50 hPa, increasing to ~13% for the top layers and to 27%-42% between 100-261 hPa. These results exhibit significant improvements over the OMI a priori as shown in black with 1-5% smaller SDs from ~1-261 hPa, while the larger SDs from pressure less than 1 hPa than those of OMI a priori/MLS comparison indicate that the combined errors from OMI and MLS exceed the natural variability of ozone.

The smoothing errors (green lines), estimated by assuming MLS data as the truth, are generally within 3% from 0.5-215 hPa, increasing to -12% at the 215-261 hPa layer and to 12% at pressure less than 0.5 hPa. Their SDs are 2-10% from 0.2-70 hPa, increasing up to 20% and 35% for the bottom two layers. The smoothing errors dominate OMI/MLS variances over the pressure range of 5-261 hPa and the larger OMI/MLS variance at pressure less than 1 hPa are likely due to the large combined OMI/MLS errors. The smoothing errors are generally consistent with OMI retrieval estimates of smoothing errors except for pressure above 50 hPa, where MLS derived smoothing errors are significantly larger. Differences due to different spatiotemporal footprints and OMI/MLS systematic errors may contribute to larger MLS-derived smoothing errors.

After removing smoothing errors through convolving MLS profiles with OMI AKs (red lines), the OMI retrievals show good agreement with MLS data (red lines) to within 3% from 0.22-100 hPa. For pressure greater than 100 hPa, OMI has negative biases of 3-9% compared to MLS data. These negative biases are not entirely from OMI because MLS ozone has been found to have positive biases in this altitude range relative to other correlative measurements (Froidevaux et al., 2008; Liu et al., 2010a). The SDs are 3.5-5% from 1.5-40 hPa, 6-10% for upper layers and 5-20% for lower layers.

The comparison is generally similar to the OMI/MLS comparison in 2006 in Liu et al. (2010a) although both OMI and MLS versions are different and the comparison in this study These comparisons are similar to the OMI/MLS comparisons shown in 2006 in Liu et al. (2010a), although both OMI and MLS versions are now different and this study is done at a finer vertical grid (due to the finer grid in MLS v4.2x data than in MLS v2 data). The comparison in the bottom layer (215-261 hPa) has larger negative MBs and larger SDs compared to that in the layer above (150-215 hPa). This is probably due to a combination of larger retrieval errors in both OMI and MLS data and larger ozone variability.

Figure 4Figure 3 (c) and (d) show similar comparison during the pre-RA period but with post-RA mask coincidence criterion, with the OMI/MLS comparison in Figure 4 Figs. 3 (a) and (b) also shown as dashed lines. The results are very similar, larger by less than 0.5% in both MBs and SDs. This indicates that the choice of OMI-MLS coincident criteria has negligible effects on the ozone profile comparison.

Figure 4Figure 3 (e) and (f) show similar comparison during the post-RA period (2009-2014). The OMI retrievals clearly exhibit degrading data quality indicated by larger MBs and SDs than those during the pre-RA period. The largest changes occur in the upper layers at pressure less than ~105 hPa; MBs increase by 2-5% and SDs increase by up to ~15%. In addition, the retrieval biases are larger than those in the a priori biases within 1-10 hPa. The larger differences Such worse comparison at higher altitudes, where the retrieval sensitivity mainly comes from shorter wavelengths (< 300 nm) in the UV-1 channel, indicates that RA significantly affects the UV-1 channel, although those OMI pixels are not flagged as RA pixels. At lower altitudes, OMI still shows good comparison with MLS data although the MBs and SDs can be larger by 2-5%.

Figure 5Figure 4 shows the MBs and SDs of the differences between OMI and MLS (convolved with OMI AKs) ozone profiles as functions of latitude during the pre-RA (with ideal coincidence) and post-RA periods, respectively. During the pre-RA period, the MBs are generally within 10% except for positive biases of up to ~20% from 40-80 hPa, and negative biases of up to ~20% from below at high latitudes, and negative biases of up to ~15%

8

in the upper stratosphere of northern high latitudes. Clearly, the MBs show large oscillations at high latitudes likely due to both larger retrieval errors and ozone variability. In addition, OMI still shows some systematic cross-track position dependent biases (Huang et al., 2017). As the OMI/MLS coincident position varies with latitude, cross-track dependent biases in our OMI retrievals can be up to 5% in the UTLS and within a few percent at higher altitudes. These errors will also contribute to the latitude dependence of OMI/MLS comparison. The patterns of SDs are similar to that in Figure 4Figure 3 (b), but typically with larger values at high latitudes. In 1-40 hPa, the SDs increase from 2-4% in the tropics to 4-8% at high latitudes. At pressure less than 1 hPa or greater than 50 hPa, the SDs increase from 5-10% in the tropics to ~30% at high latitudes. The patterns in MBs and SDs are quite symmetric between two hemispheres. During the post-RA period, the most significant changes in MBs occur at pressure less than ~5 hPa, where MBs are larger by 10-15% south of 40°N and smaller by 15-20% north of 40°N so that the patterns, especially in MBs, are not symmetric anymore. This supports the fact that the RA has different effects during different portions of the orbits. According to the analysis of RA behavior in Schenkeveld et al. (2017) , in the northern part of the orbits, the effect is dominated by the solar contamination that increases radiance values and thus causes negative ozone biases; in the rest of the orbit, the effect is dominated by the blockage effect that reduces radiance values and thus causes positive ozone biases. The SDs during the post-RA period are typically larger by 2-5% but by >10% at high altitudes than those during the pre-RA period.

Figure 6Figure 5 shows comparisons similar to those in Figure 5Figure 4 except as a function of SZA in southern and northern hemispheres, respectively. During the pre-RA period, OMI has good agreement with MLS in 1-100 hPa for SZA less than 80°, with MBs of <10% and SDs of 3-15% in both hemispheres, but with larger MBs and SDs at larger SZAs. As SZA correlates with latitude, OMI cross-track dependency will contribute to the SZA dependence of the comparison. The patterns in MBs and SDs are again quite symmetric between the two hemispheres despite different ozone fields except that the MBs are more negative in the northern hemisphere. The symmetry suggests that these biases are likely caused by SZA-dependent errors such as errors due to OMI staylight errors and radiative transfer calculations errors. During the post-RA period, in addition to the larger MBs and SDs as shown in Figure 5Figure 4, the patterns especially in the MBs are not symmetric anymore as a result of the blockage effect of the RA in the southern hemisphere, and lower latitude (smaller SZA) of the northern hemisphere and the solar contamination effect of the RA in the higher latitude (larger SZA) of the northern hemisphere.

4.2 Comparisons of OMI and MLS Stratospheric Ozone Columns (SOCs)

Figure 7Figure 6 shows a scatter density plot of OMI and MLS SOCs (SOC100, SOC215, SOC261) without applying OMI AKs during the pre-RA (2004-2008 with ideal coincidence criterion and post-RA mask) and post-RA periods, respectively. The corresponding comparison statistics are shown in black legends. All OMI SOCs show excellent agreement with MLS data even during the post-RA period. The correlation coefficients are typically within 0.96-0.99 and the linear regression slopes are within 0.04 to 1. MBs are typically within 1.6 DU (0.7%) except for 3.1 DU (1.3%) in SOC100 during the post-RA period, and SDs are within 7.0 DU (2.71%).

OMI SOCs show the best comparison with MLS SOCs during 2004-2008 with ideal coincidence as clearly seen by the smallest spread of the scatters, the smallest SDs (5.3-5.6 DU or 1.9-2.3%) and the highest correlation

coefficients (0.98-0.99). The MBs are within 1.3 DU (0.6%). With the lower boundary from 100 hPa to 215 hPa to 261 hPa, MBs decrease from 1.3 DU to 0.4 DU to -0.2 DU due to negative biases in the bottom three layers as shown in Fig. 3, SDs decrease from 5.6 (2.3%) to 5.4 DU (1.9%) to 5.3 DU (1.9%). The SOC215 comparison is significantly better than that during 2006 shown in Liu et al. (2010a), where the MBs are -1.8±7.7 DU. This might

5   reflect the improvement of both OMI and MLS over previous versions. Using the post-RA mask for the pre-RA period, the comparison only slightly degrades, with MBs larger by up to 0.3 DU, SDs larger by 0.9-1.7 DU (0.4-0.6%) and correlation smaller by 0.01. During the post-RA period, the MBs systematically increase by 1-1.4 DU over those during the pre-RA period with post-RA mask, but the SDs and correlation do not change much. Despite of the degradation in ozone profile comparison especially at pressures less than ~5 hPa during the post-RA period,

10   the overall SOC comparisons do not degrade much except for the systematic increase of MBs by ~1 DU. Also, the comparison for SOC261 is very similar to that for the SOC215, indicating that ozone in 215-261 hPa and SOC down to 261 hPa can be used for scientific use, e.g., to derive tropospheric ozone columns from OMI/MLS using the TOR method.

      The comparisons with AKs applied to MLS data are not shown but the comparison statistics are shown in red

15   legends in Figure 7Figure 6. As expected, applying OMI AKs improves the comparison with smaller SDs and better correlation. But the reduction in SDs is very small, typically within 0.3 DU (0.1%), and the increase in correlation is within 0.001. The values of MBs are smaller, typically by less than 0.2 DU for SOC100 and SOC215, and by 0.7 DU for SOC261.

      Figure 8Figure 7 shows the SOC comparisons as a function of latitudes. Generally, SOC100, SOC215, and

20   SOC261 comparison results have similar patterns: small MBs (within 1%) and SDs (within 2%) in the lower latitudes, and some latitude dependence in higher latitudes. The SOC100 MBs show significant latitudinal dependence, typically increasing with latitude. For example, SOC100 MBs during 2004-2008 are within 0.5% between 50°S-50°N areas and increase up to 4% with the increase of latitude in both hemispheres. Such latitudinal dependence decreases with increasing pressure of the lower boundary. SDs typically increase with latitude for all

25   three SOCs, from within 2% in the tropics to up to ~4% at higher latitudes, mainly due to a combination of larger OMI retrieval errors and larger ozone variability. During the post-RA period, the MBs are typically more positive and SDs are always larger especially at higher latitude mostly due to the use of looser coincidence criterion; such contrast between pre-RA and post-RA periods are consistent with that in Figure 7Figure 6.

      Figure 9Figure 8 shows similar SOC comparisons as functions of SZAs for both southern and northern

30   hemispheres. During the pre-RA period, the patterns of MBs and SDs are relatively symmetric between the two hemispheres. For SOC100, MBs are within 0.5% and do not show much SZA dependence at SZA less than 45°, then generally increase with SZA to up to 2.5%. For SOC215 and SOC261, MBs are within 0.5% and show little SZA dependence until SZA > 80°, then MBs dramatically increase with SZA to up to ~4%. The SDs are within 2% gradually increase with SZA when MBs does not change much, and then increase more with the increase of SZAs to

35   up to ~5%. During the post-RA periods, the MBs and SDs generally show more SZA dependence, SDs are typically larger especially at high SZAs, and patterns are not symmetric anymore between the two hemispheres. The patterns

10

in the southern hemisphere are still similar to that during the pre-RA period. In the northern hemisphere, MBs are reduced by up to 2% at SZAs of 65°-80° due to the solar contamination effect of the RA.

Compared to the results of Liu et al. (2010a), the SOC215 comparison during the pre-RA period is more symmetric between the two hemispheres; the SDs show stronger latitude/SZA dependence, and the MBs typically show less latitude/SZA dependence and are closer to zero. The SOC261 comparisons are very similar to the SOC215 comparisons during both pre-RA and post-RA periods, indicating the scientific use of MLS ozone at 261 hPa in applications such as the OMI/MLS TOR method. ~~This implies that MLS ozone at 261 hPa is good for scientific use and SOC261 can be used to improve the TOC derivation using the OMI/MLS TOR method by covering a broader range of the stratosphere in the middle and high latitudes where the tropopause pressure is ~ 261 hPa or larger.~~

4.3 Evaluation of Long-term Performance

Figure 10~~Figure 9~~ shows the linear trends in the relative OMI/MLS biases during 2004-2014, pre-RA (2004-2008) and post-RA (2009-2014) periods for the five latitude ranges. It should be noted that the three time periods (March – October 2009, July-October 2011, and July-December 2014) with very large OMI/MLS biases are excluded in our trend analysis (gray areas in Figure 3). The solid circles indicate significant trends with P values less than 0.05. We can see significant bias trends at most altitudes during 2009-2014 or 2004-2014 for all latitude ranges. The absolute trends are typically larger during 2009-2014 than during 2004-2014 as the latter includes the pre-RA period when the retrievals are more stable. The trends are stronger at northern middle and high latitudes and become smaller at other latitude ranges, indicating that the dominant solar contamination effect of the RA occurring in the northern portion of the orbit might have a stronger temporal variation. Most of the significant trends are within 3%/year, but there are trends with absolute values of greater than 5% at high altitudes (pressure less than 3 hPa) of 90°N-30°N and at lower altitudes (pressure greater than 125 hPa) at~~of~~ 90°N-60°N. The OMI retrievals are more stable during the pre-RA period. Significant trends do occur at pressure less than 8 hPa especially in the tropics, but the values are typically less than 1%/year except for the upper stratosphere where there are significantly large trends of up to -5%/year at 90°N-30°N.

Figure 11~~Figure 10~~ shows monthly mean OMI/MLS SOC215 biases as a function of time for the five latitude ranges. Significantly large MBs occur in the tropics and both mid-latitude regions during March–October 2009, and in the tropics and southern mid-latitude during July–October 2011 and July–December 2014 as have been shown in Figure 3~~Figure 2~~. As mentioned in Sect. 3, such large biases during the three periods are excluded in the trend analysis. Otherwise, they can drive the trend calculation. The impacts of RA are clearly seen from the much larger temporal variation in the tropics and mid-latitudes even after excluding the above three periods. During the post-RA periods, there are significant linear trends of 0.65-1.16 DU/year except for the tropics, where the SOC215 MBs show large temporal variation, but the overall trend is small, and except for southern high latitudes, where the linear trend is large, but the P value of 0.14 is less significant. For the entire 2004-2014 period, there are significant linear trends of 0.26-0.40 DU/year in the tropics and mid-latitudes. During the pre-RA period, there are statistically

significant trends of 0.21-0.34 DU/year in the tropics and northern mid-latitudes, but the trends are much smaller than those during the post RA periods.

The significant trends of OMI/MLS ozone biases at different layers as well as in SOC suggest that the current ozone profile product is not suitable for trend studies, especially during the post-RA period. The retrieval is more
5   stable during the pre-RA period. The stronger temporal variation of the retrieval performance during the post-RA period is likely associated with the RA evolution. To maintain the long-term stability of our ozone profile product, we need to perform soft calibration similar to Liu et al. (2010b), especially during the post-RA period if the radiometric calibration improvement cannot be done in the level 0-1b processing. As the effects of the RA vary along the orbit (e.g., dominated by the solar contamination effect during the northern portion of the orbit and the
10   blockage effect during the rest of the orbit), the empirical correction should be derived as a function of both latitude and time.

5   Summary and Conclusion

To complement our validation of the 10-year OMI PROFOZ product using ozonesonde observations in a prequel of this paper, we evaluated this product including ozone profiles from 0.22–261 hPa and Stratospheric
15   Ozone Columns (SOCs) down to pressure levels 100, 215, and 261 hPa (i.e., SOC100, SOC215, SOC261) using MLS v4.2x product during the same period (October 2004-December 2014). To investigate the impacts of Row Anomaly (RA) on the retrievals, we contrasted the comparison before and after the occurrence of major OMI RA in January 2009, i.e., 2004–2008 (i.e., pre-RA period) and 2009-2014 (i.e., post-RA period). We applied ideal OMI-MLS coincident criterion (i.e., MLS footprint center lies in the footprint of an OMI footprint) in the pre-RA period
20   and the nearest coincident criterion where retrievals for the ideally collocated OMI pixels are not available due to the RA impacts. To show the impacts of coincident criteria on the comparison, we also conducted the comparison for the pre-RA period using the post-RA coincidence criterion by masking pre-RA retrievals at cross-track positions 13–21 (i.e., pre-RA with post-RA mask). There are unreasonably large OMI-MLS biases during March – October 2009, July-October 2011 and July-December 2014, which was very likely caused by the RA. Therefore, we
25   excluded OMI data during these periods. To better understand retrieval errors and sensitivity, we compared the retrieved ozone profiles and a priori profiles at individual layers with MLS data before and after being degraded to the OMI vertical resolution with OMI averaging kernels (AKs), and characterized the OMI smoothing errors using MLS data. To show the spatial consistency of the retrievals, we performed the comparison as a function of latitude and solar zenith angle (SZA). Finally, we analyzed the monthly variation of the mean biases (MBs) of ozone profiles
30   and SOC215 to examine the long-term stability of our OMI PROFOZ product.

During the pre-RA period, OMI ozone profiles agree well with the original MLS data. The global MBs are within 6% from 1-150 hPa, and up to 12% for the upper layers and up to -25% for the bottom layers. The standard deviations (SDs) range from 4-7% in 1-50 hPa, up to ~13% for pressure <1 hPa and up to 42% for pressure >50 hPa. The large SDs at lower altitudes are mainly due to OMI smoothing errors. After removing smoothing errors by
35   convolving MLS data with OMI AKs, the MBs are within 3% between 0.22 hPa - 100 hPa, negative 3-9% for lower layers, and the SDs are within 3.5-5% between 1-40 hPa, 6-10% for upper layers and 5-20% for lower layers. The

12

impact of using different coincident criteria on the ozone profile comparison is negligible; using the post-RA mask increases the SDs/MBs by less than 0.5%. During the post-RA period, OMI retrievals become slightly worse between 5-261 hPa with global MBs and SDs typically larger by 2-5% and are much worse for pressure less than ~5 hPa with larger MBs by up to 8% and SDs by up to 15%. The much worse comparison at higher altitudes indicates that RA significantly affects the UV-1 channel of the OMI measurements, although they are not flagged as RA pixels.

OMI ozone profiles show latitude and SZA dependent biases with respect to MLS data. During the pre-RA period, the patterns in MBs and SDs are quite symmetric between the two hemispheres despite different ozone fields, which suggests that these biases were likely caused by SZA-dependent errors such as errors due to OMI straylight errors and radiative transfer calculations errors. MBs are generally within 10%, but show larger oscillations at high latitudes/SZAs with positive biases of up to 20% from 40-80 hPa and negative biases of up to 20% from below. SDs increase from 2-4% at lower latitudes/SZAs to 4-8% at high latitudes/SZAs in 1-40 hPa, and increases from 5-10% to ~30% at pressure less than 1 hPa or greater than 50 hPa. During the post-RA period, the different effects of RA caused asymmetry in the patterns of MBs and SDs between the two hemispheres at pressure less than ~5 hPa, where the MBs are larger by 10-15% south of 40°N due to the blockage effect of RA and smaller by 15-20% north of 40°N due to the solar contamination effect of RA, and the SDs are larger by 10%. For lower altitudes (5-261 hPa), the MBs and SDs are typically larger by 2-5% than those during the pre-RA period. From the OMI scientific requirement document (Levelt et al., 2000), the scientific requirement for OMI ozone is to achieve 10% accuracy in the stratosphere and 30% in the troposphere. The validation of stratospheric ozone profiles with MLS data indicates that our retrieval performance generally meets OMI's scientific requirement in most of the stratosphere especially during the pre-RA period (Levelt, 2000) except for high latitudes/SZAs.

All OMI SOCs show very good agreement with original MLS data during both pre-RA and post-RA periods. During the pre-RA period, the global mean MBs are within 0.6%, and the SDs are 1.9% for SOC215 and SOC261 and 2.30% for SOC100. Using the post-RA mask only slightly degrades the pre-RA comparison, with MBs larger by up to 0.3 DU, SDs larger by 0.9-1.7 DU (0.4-0.6%). During the post-RA period, OMI SOCs also slightly degrade, with MBs larger by 1-1.4 DU (0.4-0.7%) and SDs comparable to those during the pre-RA period with post-RA mask. Applying the OMI AKs to MLS data only slightly improves the comparison due to the small smoothing errors in SOCs, reducing the SDs by less than 0.3 DU (0.1%). Similar to the ozone profiles, OMI SOCs show latitude/SZA dependent biases. During the pre-RA period, MBs (SDs) are within 0.5% (2%) in the lower latitudes/SZAs and can increase up to 4% at high latitudes and up to ~6% at SZAs larger than 85°. The MBs in SOC215 and SOC261 show less latitudinal dependence and show little SZA dependence until SZA > 80° compared to SOC100. The patterns in MBs and SDs are quite symmetric especially for SOC215 and SOC261 as a function of SZA. During the post-RA period, the MBs and SDs generally show more latitude/SZA dependence. The MBs are typically more positive, and the SDs are larger especially at high latitudes/SZAs. The patterns in MBs and SDs are less symmetric due to the various effects of RA along the orbit (i.e., blockage effect south of ~40°N and solar contamination effect north of ~40°N). The SOC261 comparison is very similar to the SOC215 comparison during both pre-RA and post-RA

13

periods, implying the MLS ozone at 261 hPa and MLS SOC261 are good for scientific use and can be used to improve the tropospheric ozone column derivation using various OMI/MLS tropospheric ozone residual methods.

Finally, we evaluated long-term data stability of ozone profile and SOC215 with respect to MLS data and calculate linear regression trends of monthly MBs in five different latitude ranges during pre-RA, post-RA and 2004-2014 periods, respectively. This validation shows significant bias trends or more temporal variation during the post-RA period; significant trends also occur during the pre-RA period, for example for ozone at pressure less than 8 hPa in the tropics and SOC215 in the tropics and northern mid-latitudes, but the magnitudes of the trends or the temporal variation are much smaller compared to those during the post-RA period. Overrall, our validation indicates that our retrieval performance genearlly meets OMI's scientific requirement for ozone profile retrieval to have 10% accuracy in the stratosphere(Levelt, 2000) especially during the pre-RA period. Hower, these significant bias trends suggest that the current ozone profile product is not suitable for trend studeis, especially during the post-RA period. The spatiotemporal variation of the retrieval performance suggests the need to improve OMI's radiometric calibration as a function of both latitude and time especially during the post-RA period to maintain the long-term stability and spatial consistency of our ozone profile product.

References

Bak, J., Kim, J. H., Liu, X., Chance, K., and Kim, J.: Evaluation of ozone profile and tropospheric ozone retrievals from GEMS and OMI spectra, Atmos. Meas. Tech., 6, 239-249, 2013.

5   Bak, J., Liu, X., Kim, J. H., Deland, M. T., and Chance, K.: Improvement of OMI ozone profile retrievals by simultaneously fitting polar mesospheric clouds, Atmos. Meas. Tech., 9, 4521-4531, doi: 10.5194/amt-9-4521-2016, 2016.

Bhartia, P. K., and Wellemeyer, C. G.: TOMS-V8 total ozone algorithm, in: OMI Algorithm Theoretical Basis Document, edited by: Bhartia, P. K., Greenbelt, 2002.

10 Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B., Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W.

15 V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements, J. Geophys. Res.-Atmos., 113, doi: 10.1029/2007JD008771, 2008.

Huang, G., Liu, X., Chance, K., Yang, K., Bhartia, P. K., Cai, Z., Allaart, M., Ancellet, G., Calpini, B., Coetzee, G. J. R., Cuevas-Agulló, E., Cupeiro, M., De Backer, H., Dubey, M. K.,

20 Fuelberg, H. E., Fujiwara, M., Godin-Beekmann, S., Hall, T. J., Johnson, B., Joseph, E., Kivi, R., Kois, B., Komala, N., König-Langlo, G., Laneve, G., Leblanc, T., Marchand, M., Minschwaner, K. R., Morris, G., Newchurch, M. J., Ogino, S. Y., Ohkawara, N., Piters, A. J. M., Posny, F., Querel, R., Scheele, R., Schmidlin, F. J., Schnell, R. C., Schrems, O., Selkirk, H., Shiotani, M., Skrivánková, P., Stübi, R., Taha, G., Tarasick, D. W., Thompson, A. M., Thouret, V., Tully, M.

25 B., Van Malderen, R., Vömel, H., von der Gathen, P., Witte, J. C., and Yela, M.: Validation of 10-year SAO OMI Ozone Profile (PROFOZ) product using ozonesonde observations, Atmos. Meas. Tech., 10, 2455-2475, doi: 10.5194/amt-10-2455-2017, 2017.

Hubert, D., Lambert, J. C., Verhoelst, T., Granville, J., Keppens, A., Baray, J. L., Bourassa, A. E., Cortesi, U., Degenstein, D. A., Froidevaux, L., Godin-Beekmann, S., Hoppel, K. W.,

30 Johnson, B. J., Kyrölä, E., Leblanc, T., Lichtenberg, G., Marchand, M., McElroy, C. T., Murtagh, D., Nakane, H., Portafaix, T., Querel, R., Russell Iii, J. M., Salvador, J., Smit, H. G. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Stübi, R., Swart, D. P. J., Taha, G., Tarasick, D. W., Thompson, A. M., Urban, J., van Gijsel, J. A. E., Van Malderen, R., von der Gathen, P., Walker, K. A., Wolfram, E., and Zawodny, J. M.: Ground-based assessment of the bias and long-

35 term stability of 14 limb and occultation ozone profile data records, Atmos. Meas. Tech., 9, 2497-2534, doi: 10.5194/amt-9-2497-2016, 2016.

Jiang, Y. B., Froidevaux, L., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Bojkov, B., Leblanc, T., McDermid, I. S., Godin-Beekmann, S., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W.,

40 Perun, V. S., Schwartz, M. J., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Allaart, M., Andersen, S. B., Bodeker, G., Calpini, B., Claude, H., Coetzee, G., Davies, J., De Backer, H., Dier, H., Fujiwara, M., Johnson, B., Kelder, H., Leme, N. P., König-Langlo, G., Kyro, E., Laneve, G., Fook, L. S., Merrill, J., Morris, G., Newchurch, M., Oltmans, S., Parrondos, M. C., Posny, F., Schmidlin, F., Skrivankova, P., Stubi, R., Tarasick, D., Thompson,

45 A., Thouret, V., Viatte, P., Vömel, H., von Der Gathen, P., Yela, M., and Zablocki, G.:

Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, J. Geophys. Res.-Atmos., 112, doi: 10.1029/2007JD008776, 2007.

Jing, P., Cunnold, D., Choi, Y., and Wang, Y.: Summertime tropospheric ozone columns from Aura OMI/MLS measurements versus regional model results over the United States, Geophysical Research Letters, 33, doi: 10.1029/2006GL026473, 2006.

Kroon, M., de Haan, J. F., Veefkind, J. P., Froidevaux, L., Wang, R., Kivi, R., and Hakkarainen, J. J.: Validation of operational ozone profiles from the Ozone Monitoring Instrument, J. Geophys. Res., 116, D18305, doi: 10.1029/2010jd015100, 2011.

Levelt, P. F.: Science requirements document for OMI-EOS, KNMI, 2000.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The Ozone Monitoring Instrument, IEEE T. Geosci. Remote., 44, 1093-1101, 2006.

Liu, X., Bhartia, P. K., Chance, K., Froidevaux, L., Spurr, R. J. D., and Kurosu, T. P.: Validation of Ozone Monitoring Instrument (OMI) ozone profiles and stratospheric ozone columns with Microwave Limb Sounder (MLS) measurements, Atmos. Chem. Phys., 10, 2539-2549, doi: 10.5194/acp-10-2539-2010, 2010a.

Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., and Kurosu, T. P.: Ozone profile retrievals from the Ozone Monitoring Instrument, Atmos. Chem. Phys., 10, 2521-2537, doi: 10.5194/acp-10-2521-2010, 2010b.

Livesey, N. J., Filipiak, M. J., Froidevaux, L., Read, W. G., Lambert, A., Santee, M. L., Jiang, J. H., Pumphrey, H. C., Waters, J. W., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fuller, R. A., Jarnot, R. F., Jiang, Y. B., Knosp, B. W., Li, Q. B., Perun, V. S., Schwartz, M. J., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Avery, M., Browell, E. V., Cammas, J. P., Christensen, L. E., Diskin, G. S., Gao, R. S., Jost, H. J., Loewenstein, M., Lopez, J. D., Nedelec, P., Osterman, G. B., Sachse, G. W., and Webster, C. R.: Validation of Aura Microwave Limb Sounder O3 and CO observations in the upper troposphere and lower stratosphere, J. Geophys. Res., 113, doi: 10.1029/2007jd008805, 2008.

Livesey, N. J., William G. Read, Paul A. Wagner, Lucien Froidevaux, Alyn Lambert, Gloria L. Manney, Luis F. Millan Valle, Hugh C. Pumphrey, Michelle L. Santee, Michael J. Schwartz, Shuhui Wang, Ryan A. Fuller, Robert F. Jarnot, Brian W. Knosp, and Martinez, E.: Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) Version 4.2x Level 2 data quality and description document, in, edited by: Laboratory, J. P., 2015.

NASA: The Aura Validation Data Center doi: https://avdc.gsfc.nasa.gov/index.php?site=2045907950, 2013.

Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C., Veefkind, J. P., and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, Atmos. Meas. Tech., 10, 1957-1986, doi: 10.5194/amt-10-1957-2017, 2017.

Schoeberl, M. R., Ziemke, J. R., Bojkov, B., Livesey, N., Duncan, B., Strahan, S., Froidevaux, L., Kulawik, S., Bhartia, P. K., Chandra, S., Levelt, P. F., Witte, J. C., Thompson, A. M., Cuevas, E., Redondas, A., Tarasick, D. W., Davies, J., Bodeker, G., Hansen, G., Johnson, B. J., Oltmans, S. J., Vömel, H., Allaart, M., Kelder, H., Newchurch, M., Godin-Beekmann, S., Ancellet, G., Claude, H., Andersen, S. B., Kyrö, E., Parrondos, M., Yela, M., Zablocki, G., Moore, D., Dier, H., von der Gathen, P., Viatte, P., Stübi, R., Calpini, B., Skrivankova, P., Dorokhov, V., de Backer, H., Schmidlin, F. J., Coetzee, G., Fujiwara, M., Thouret, V., Posny, F., Morris, G., Merrill, J., Leong, C. P., Koenig-Langlo, G., and Joseph, E.: A trajectory-based

estimate of the tropospheric ozone column using the residual method, J. Geophys. Res., 112, doi: 10.1029/2007jd008773, 2007.

Smit, H. G. J., Straeter, W., Johnson, B. J., Oltmans, S. J., Davies, J., Tarasick, D. W., Hoegger, B., Stubi, R., Schmidlin, F. J., Northam, T., Thompson, A. M., Witte, J. C., Boyd, I., and Posny, F.: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), J. Geophys. Res., 112, 19306, 2007.

Thompson, A. M., Witte, J. C., Smit, H. G. J., Oltmans, S. J., Johnson, B. J., Kirchhoff, V. W. J. H., and Schmidlin, F. J.: Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998–2004 tropical ozone climatology: 3. Instrumentation, station-to-station variability, and evaluation with simulated flight profiles, J. Geophys. Res., 112, doi: 10.1029/2005jd007042, 2007.

van Oss, R. F., Voors, R. H. M., and Spurr, R. J. D.: Ozone profile algorithm, in: OMI Algorithm Theoretical Basis Document, Volume II: OMI ozone products, edited by: Bhartia, P. K., Greenbelt, MD, 51-73, 2001.

Veefkind, J. P., de Haan, J. F., Brinksma, E. J., Kroon, M., and Levelt, P. F.: Total Ozone From the Ozone Monitoring Instrument (OMI) Using the DOAS Technique, IEEE T. Geosci. Remote., 44, 1239-1244, 2006.

Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Gun-Shing, C., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Yibo, J., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D. M., Snyder, W. V., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth observing system microwave limb sounder (EOS MLS) on the aura Satellite, IEEE T. Geosci. Remote., 44, 1075-1092, doi: 10.1109/TGRS.2006.873771, 2006.

Worden, H. M., Logan, J. A., Worden, J. R., Beer, R., Bowman, K., Clough, S. A., Eldering, A., Fisher, B. M., Gunson, M. R., Herman, R. L., Kulawik, S. S., Lampel, M. C., Luo, M., Megretskaia, I. A., Osterman, G. B., and Shephard, M. W.: Comparisons of Tropospheric Emission Spectrometer (TES) ozone profiles to ozonesondes: Methods and initial results, J. Geophys. Res., 112, doi: 10.1029/2006jd007258, 2007.

Yang, Q., Cunnold, D. M., Wang, H. J., Froidevaux, L., Claude, H., Merrill, J., Newchurch, M., and Oltmans, S. J.: Midlatitude tropospheric ozone columns derived from the Aura Ozone Monitoring Instrument and Microwave Limb Sounder measurements, J. Geophys. Res.: Atmos., 112, D20305, doi: doi: 10.1029/2007JD008528, 2007.

Ziemke, J. R., Chandra, S., Duncan, B. N., Froidevaux, L., Bhartia, P. K., Levelt, P. F., and Waters, J. W.: Tropospheric ozone determined from Aura OMI and MLS: Evaluation of measurements and comparison with the Global Modeling Initiative's Chemical Transport Model, J. Geophys. Res.-Atmos., 111, D19303, doi: 10.1029/2006JD007089, 2006.

Ziemke, J. R., Olsen, M. A., Witte, J. C., Douglass, A. R., Strahan, S. E., Wargan, K., Liu, X., Schoeberl, M. R., Yang, K., Kaplan, T. B., Pawson, S., Duncan, B. N., Newman, P. A., Bhartia, P. K., and Heney, M. K.: Assessment and applications of NASA ozone data products derived from Aura OMI/MLS satellite measurements in context of the GMI chemical transport model, J. Geophys. Res., 119, 5671-5699, doi: 10.1002/2013jd020914, 2014.
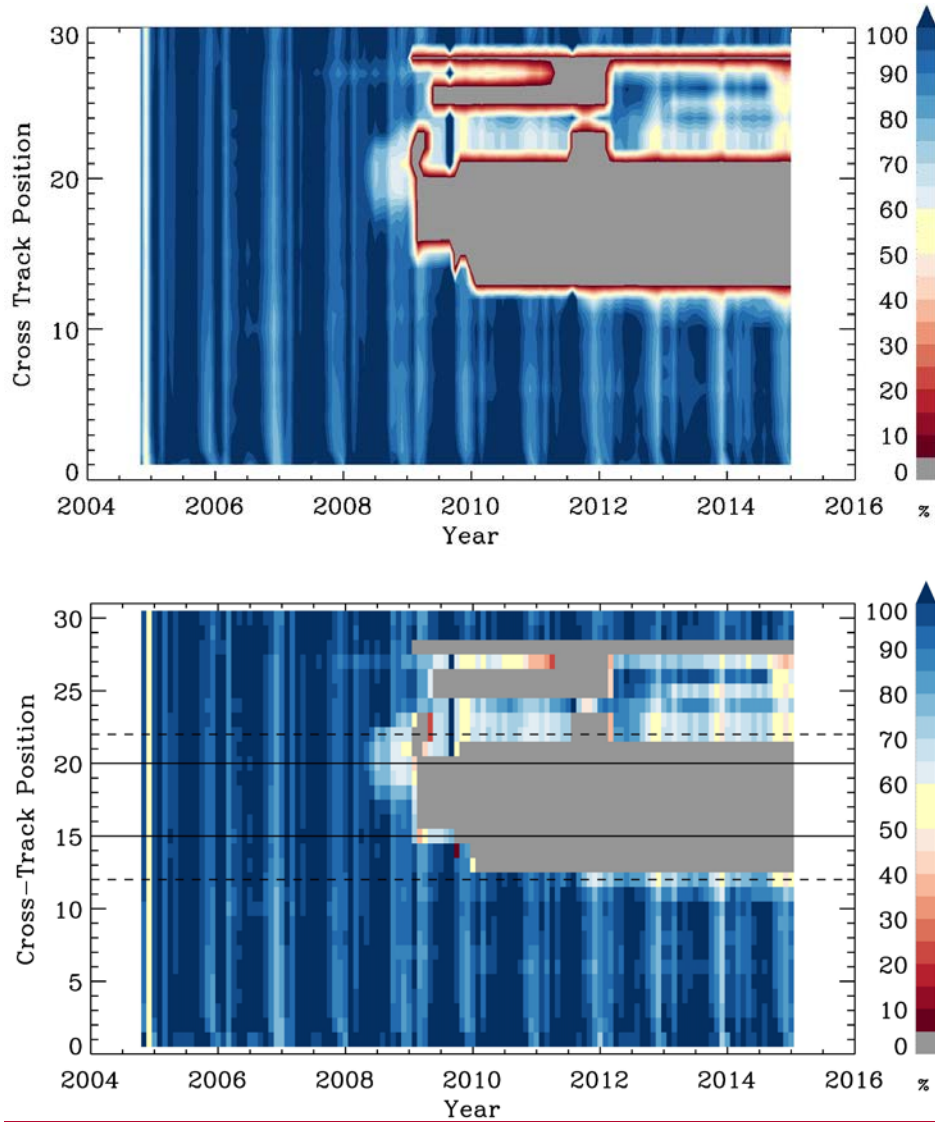
Figures and Figure Captions

Figure 1. The monthly mean number of OMI retrievals at each cross-track position from October 2004 through December 2014, relative to the corresponding monthly mean number of OMI retrievals in January 2005, i.e at the beginning of OMI operation. **The solid lines represent the range of OMI-MLS collocations** that vary from position 20 in the tropics to 15 at high latitudes, **while the dashed lines represent the** enforced post-RA OMI-MLS coincidence criteria (the closest OMI pixel from position $\leq$ 12 or $\geq$ 22).
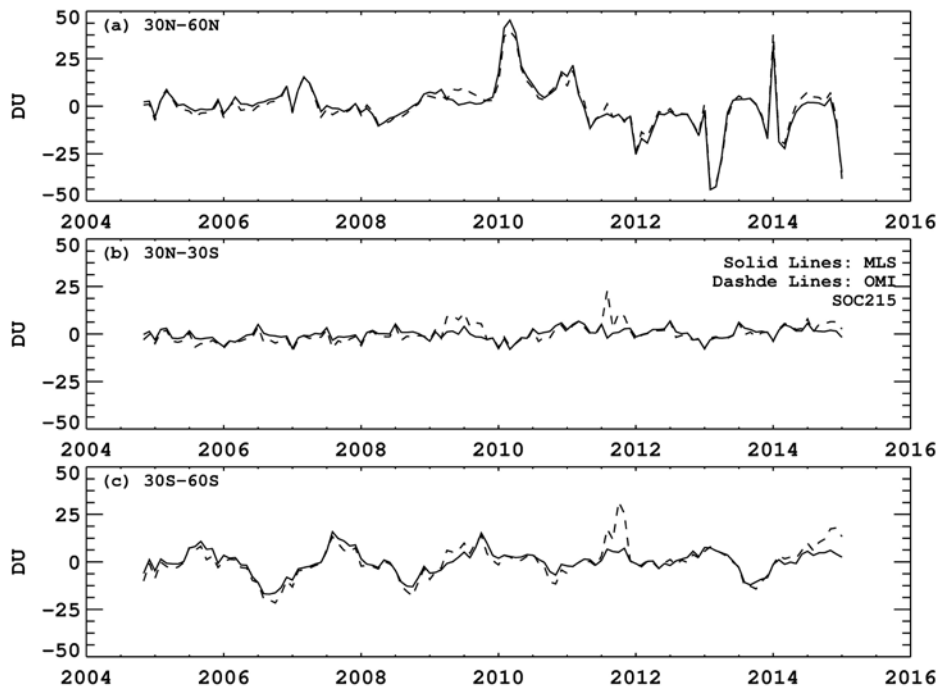
18

Figure 2. Overview of the vertical range of OMI and MLS. The solid black bars represent vertical grids of OMI and MLS, respectively, while the validation range of this study is marked by grey lines.
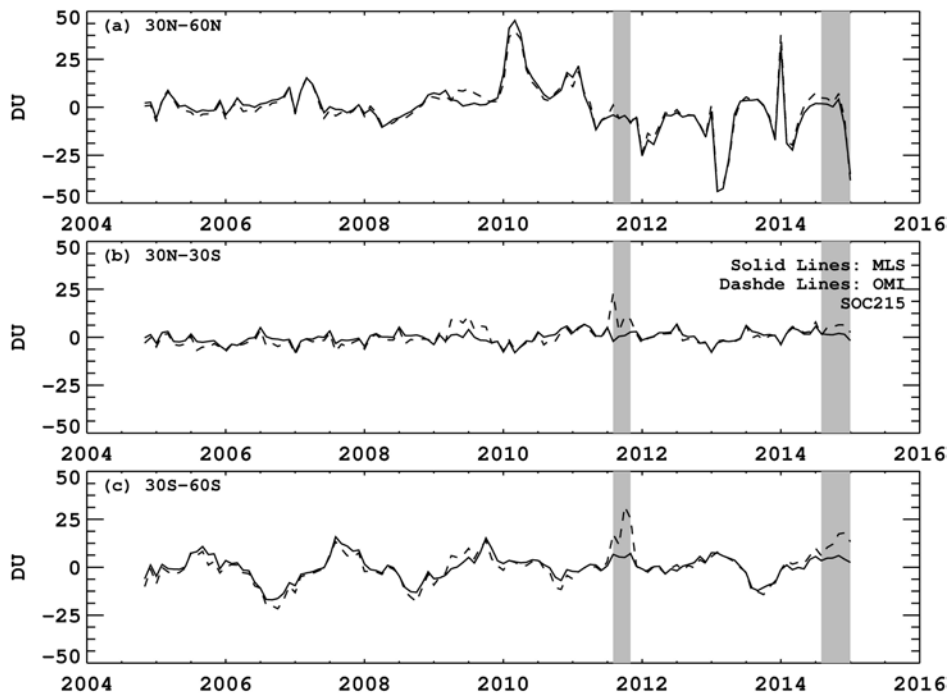
5

(a) 30N-60N

(b) 30N-30S

Solid Lines: MLS
Dashde Lines: OMI
SOC215

(c) 30S-60S

20
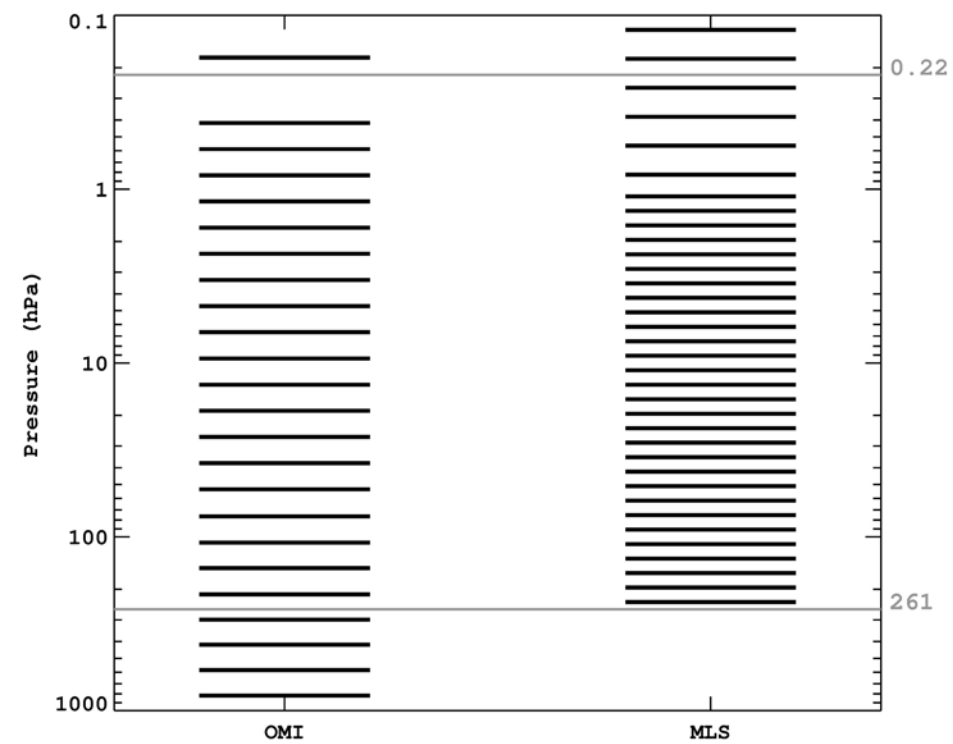
Figure 3. Deseasonalized time series of monthly mean MLS (solid lines) and OMI (dashed lines) Stratospheric Ozone Columns down to 215 hPa (SOC215s) in (a) 30°N-60°N, (b) 30°N-30°S and (c) 60°S-30°S, respectively. The periods that have been omitted in our analysis are marked as gray areas.
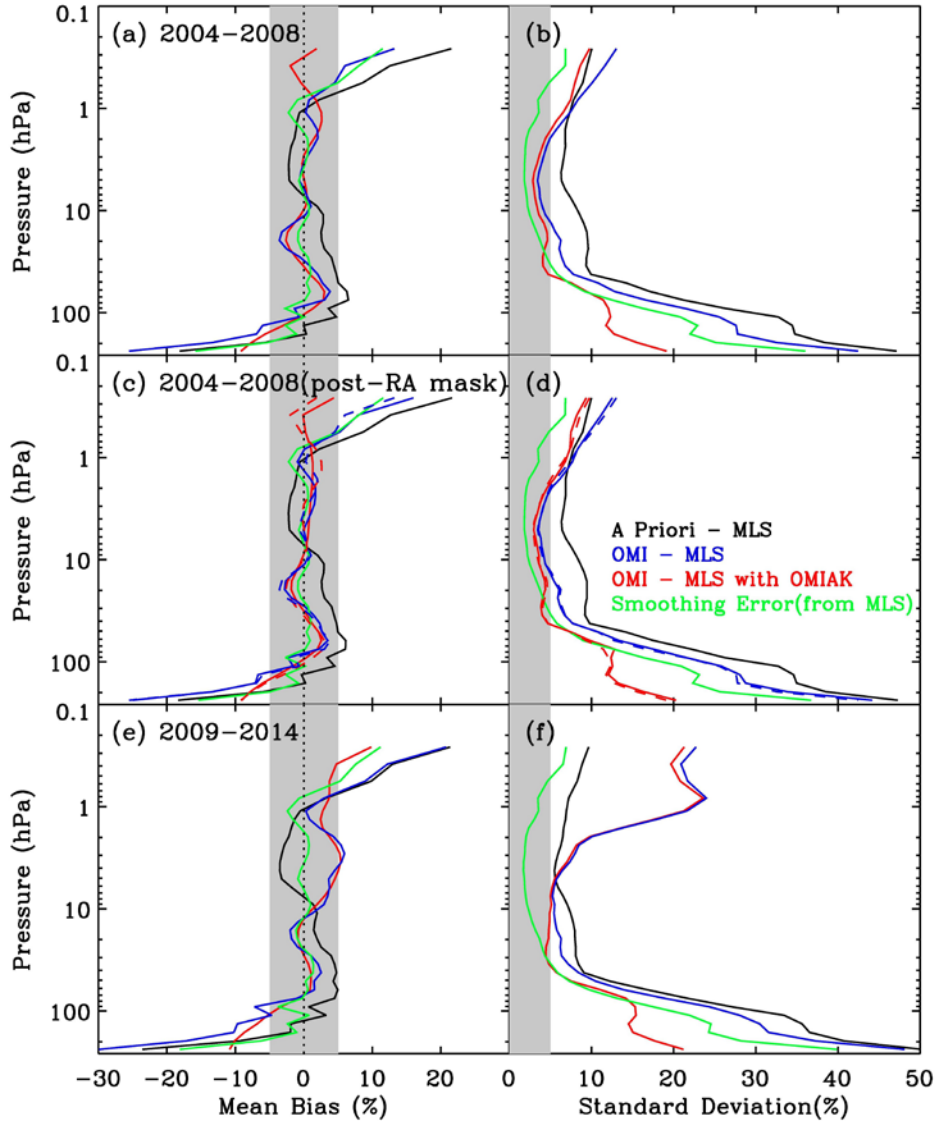
5

Figure 4. (left) Global mean biases at each MLS layer and (right) corresponding standard deviations of the differences between OMI and MLS profiles from 0.22 -261 hPa in (top) 2004-2008, (middle) 2004-2008 with post-RA mask and (bottom) 2009-2014, respectively. The black/blue lines compare a priori/OMI retrievals with original MLS profiles. The red lines compare OMI retrievals with MLS profiles after applying OMI Averaging Kernels (AKs). The green lines represent OMI smoothing errors estimated by assuming MLS data as the truth. For contrast, the comparisons of OMI and MLS data based on the ideal collocation criteria in 2004-2008 are plotted in blue and red dashed lines in the middle panels.
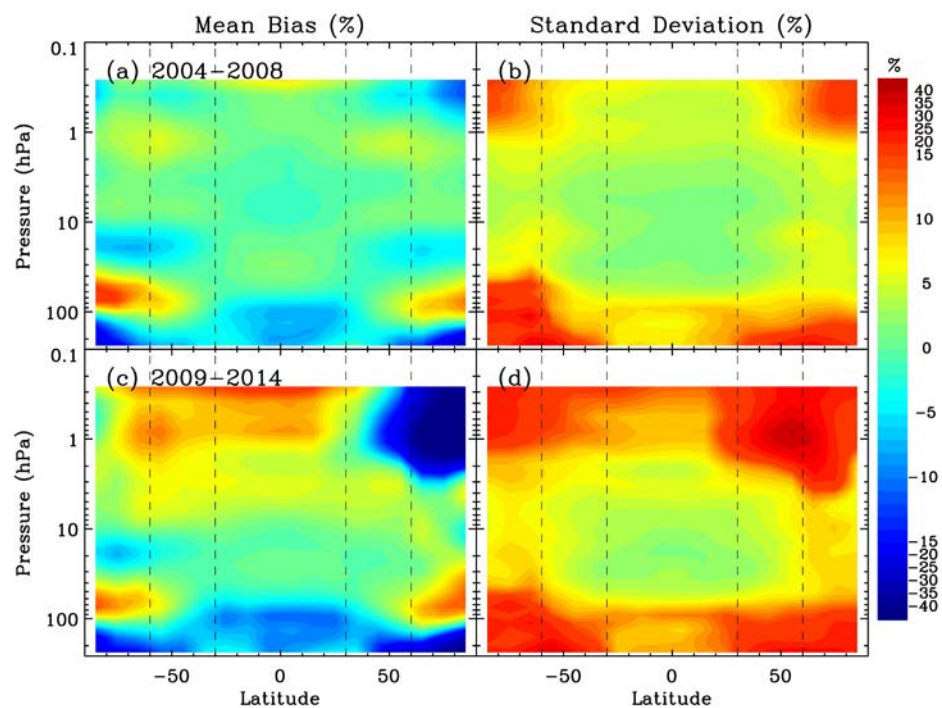
22

Figure 5. (left) Mean OMI/MLS ozone profile biases and (right) corresponding standard deviations as a function of latitudes in (top) 2004-2008 and (bottom) 2009-2014, respectively. OMI averaging kernels are applied to MLS data.
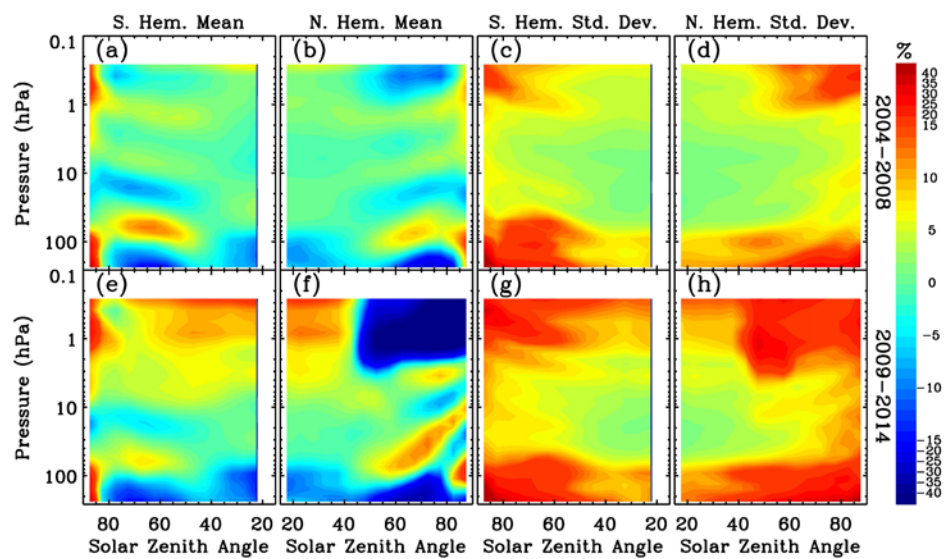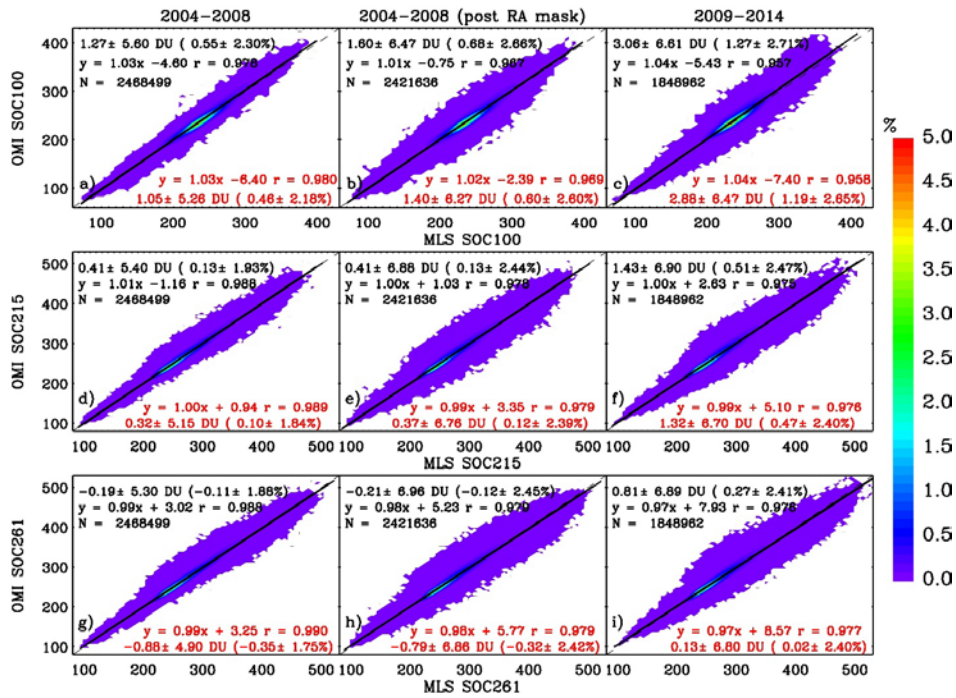
Figure 6. Similar to Figure 4, but as a function of solar zenith angles for South and North Hemispheres, respectively.

2004–2008   2004–2008 (post RA mask)   2009–2014

a) 1.27± 5.60 DU (0.55± 2.30%) y = 1.03x −4.60 r = 0.980 N = 2468499 y = 1.03x −6.40 r = 0.980 1.05± 5.26 DU (0.46± 2.18%)

b) 1.60± 6.47 DU (0.68± 2.66%) y = 1.01x −0.75 r = 0.967 N = 2421636 y = 1.02x −2.39 r = 0.969 1.40± 6.27 DU (0.80± 2.60%)

c) 3.06± 6.61 DU (1.27± 2.71%) y = 1.04x −5.43 r = 0.957 N = 1848962 y = 1.04x −7.40 r = 0.958 2.88± 6.47 DU (1.19± 2.65%)

d) 0.41± 5.40 DU (0.13± 1.93%) y = 1.01x −1.16 r = 0.988 N = 2468499 y = 1.00x + 0.94 r = 0.989 0.32± 5.15 DU (0.10± 1.84%)

e) 0.41± 6.88 DU (0.13± 2.44%) y = 1.00x + 1.03 r = 0.978 N = 2421636 y = 0.99x + 3.35 r = 0.979 0.37± 6.76 DU (0.12± 2.39%)

f) 1.43± 6.90 DU (0.51± 2.47%) y = 1.00x + 2.63 r = 0.975 N = 1848962 y = 0.99x + 5.10 r = 0.976 1.32± 6.70 DU (0.47± 2.40%)

g) −0.19± 5.30 DU (−0.11± 1.86%) y = 0.99x + 3.02 r = 0.988 N = 2468499 y = 0.99x + 3.25 r = 0.990 −0.88± 4.90 DU (−0.35± 1.75%)

h) −0.21± 6.96 DU (−0.12± 2.45%) y = 0.98x + 5.23 r = 0.979 N = 2421636 y = 0.98x + 5.77 r = 0.979 −0.79± 6.86 DU (−0.32± 2.42%)

i) 0.81± 6.89 DU (0.27± 2.41%) y = 0.97x + 7.93 r = 0.976 N = 1848962 y = 0.97x + 8.57 r = 0.977 0.13± 6.80 DU (0.02± 2.40%)

OMI SOC100   OMI SOC215   OMI SOC261

MLS SOC100   MLS SOC215   MLS SOC261

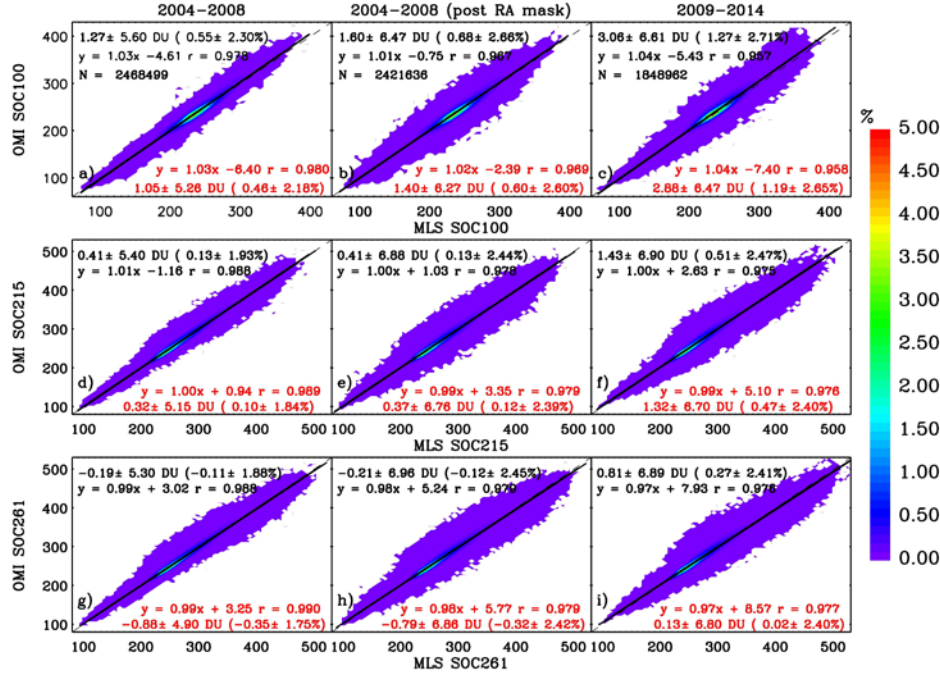% 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

Figure 7. Scattering density plots of OMI and MLS stratospheric ozone columns above 100, 215 and 261 hPa (i.e., SOC100, SOC 215 and SOC261), respectively, from the top to bottom panels, during 2004-2008 (left), 2004-2008 with post-RA mask (middle column), and 2009-2014 (right). OMI averaging kernels (AKs) are not applied to MLS data. The comparison statistics of OMI-MLS SOCs are shown in black including mean biases, standard deviations, relative mean biases and standard deviations, correlation coefficients, linear regressions and the number of OMI-MLS pairs. Comparison statistics with OMI AKs applied to MLS data are also provided in red.
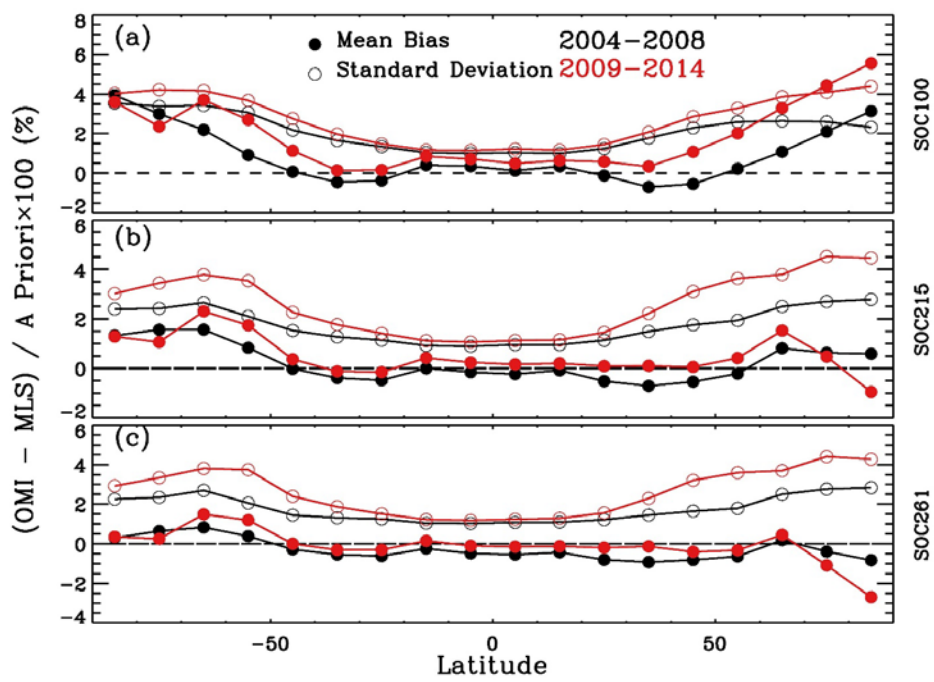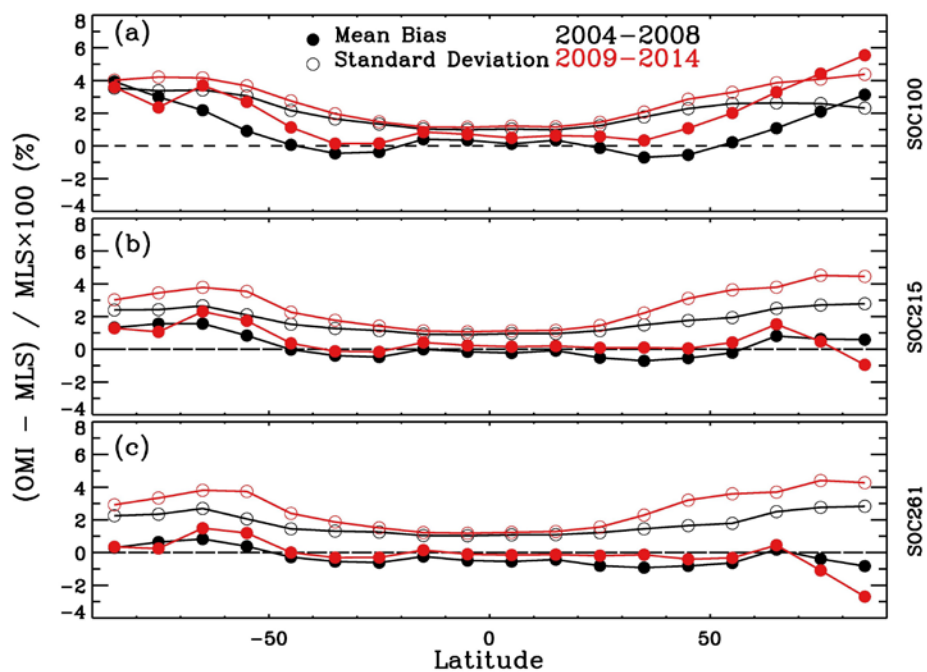
Figure 8. Mean biases (solid circles) and standard deviations (open circles) of the differences between OMI and MLS stratospheric ozone columns (SOCs) above 100 hPa (top), 215 hPa (middle) and 261 hPa (bottom) as a function of latitude, during 2004-2008 (black) and 2009-2014 (red). Please noted that standard deviations are plotted with the zero-line as a reference.
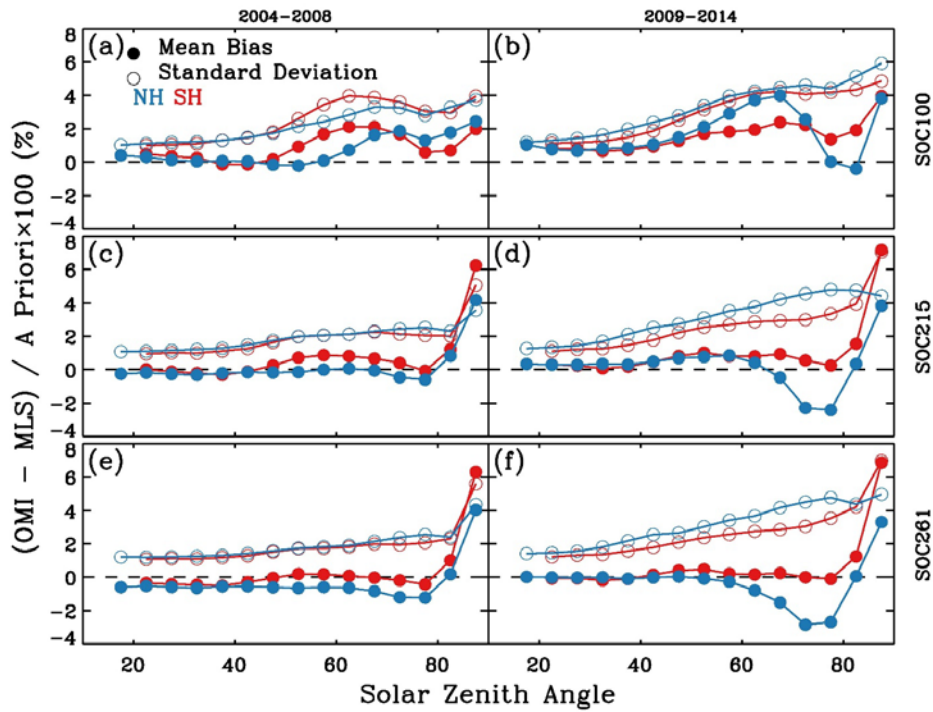
5

Figure 9. Mean biases (solid circles) and standard deviations (open circles) of the differences between OMI and MLS stratospheric ozone columns (SOCs) above 100 hPa (top), 215 hPa (middle) and 261 hPa (bottom) as a function of solar zenith angle for Northern Hemisphere (blue) and Southern Hemisphere (red) during 2004-2008 (left) and 2009-2014 (right). Please noted that standard deviations are plotted with the zero-line as reference.
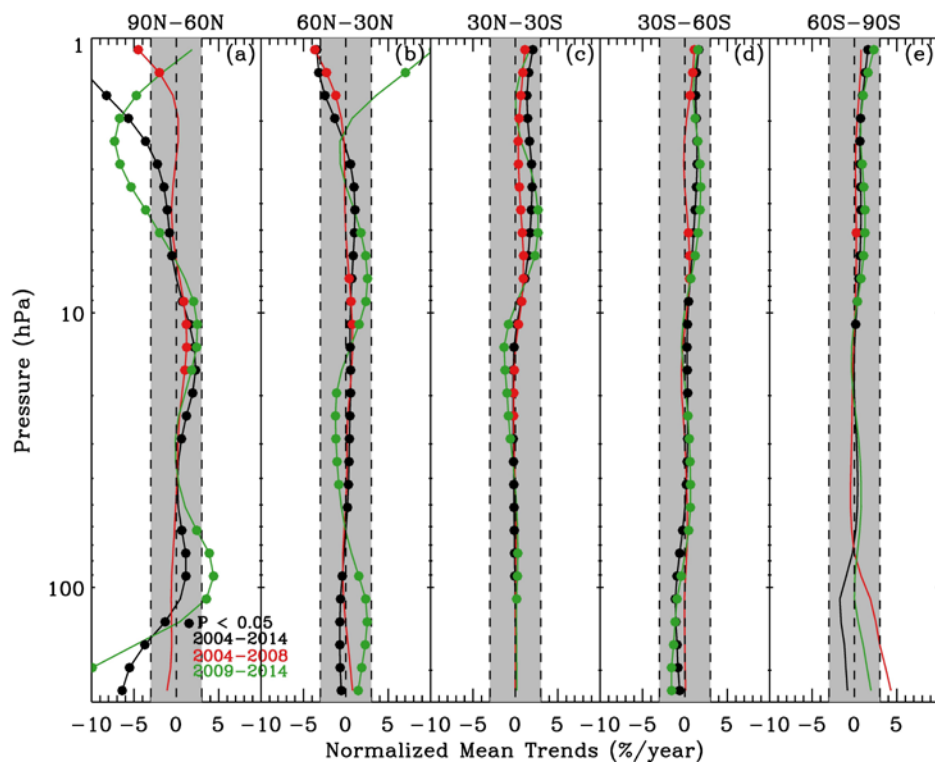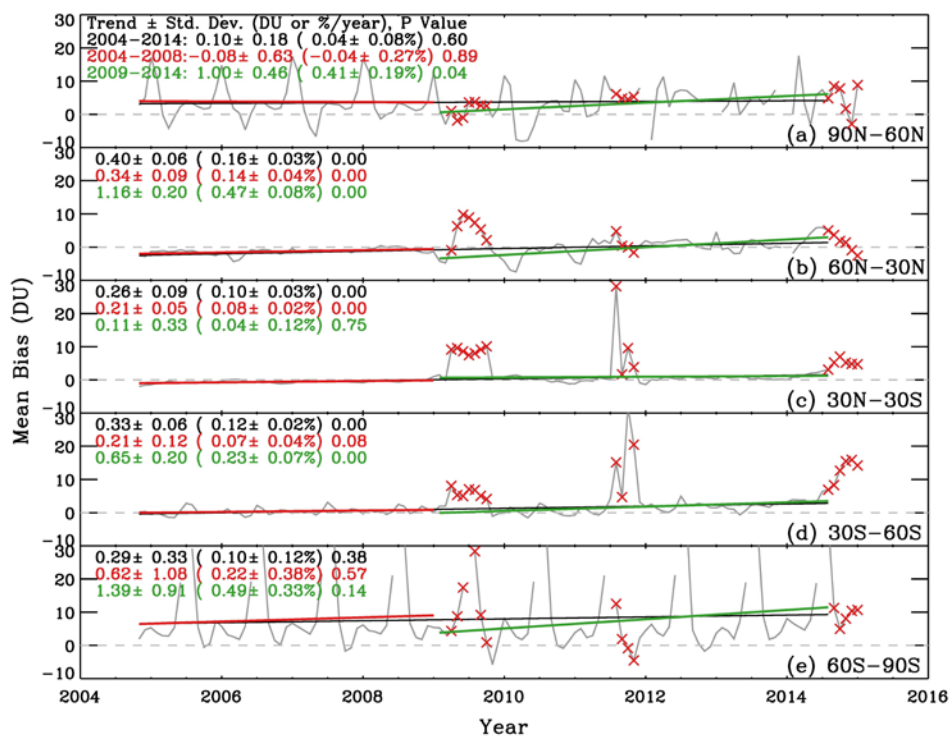
Figure 10. Linear regression trends of relative OMI-MLS biases as a function of altitude for five latitudes bands during three time periods: 2004-2014 (black), 2004-2008 (red) and 2009-2014 (green). Solid circles indicate significant trends (with P value less than 0.05). The grey areas indicate the range of -3% and 3% in each panel.
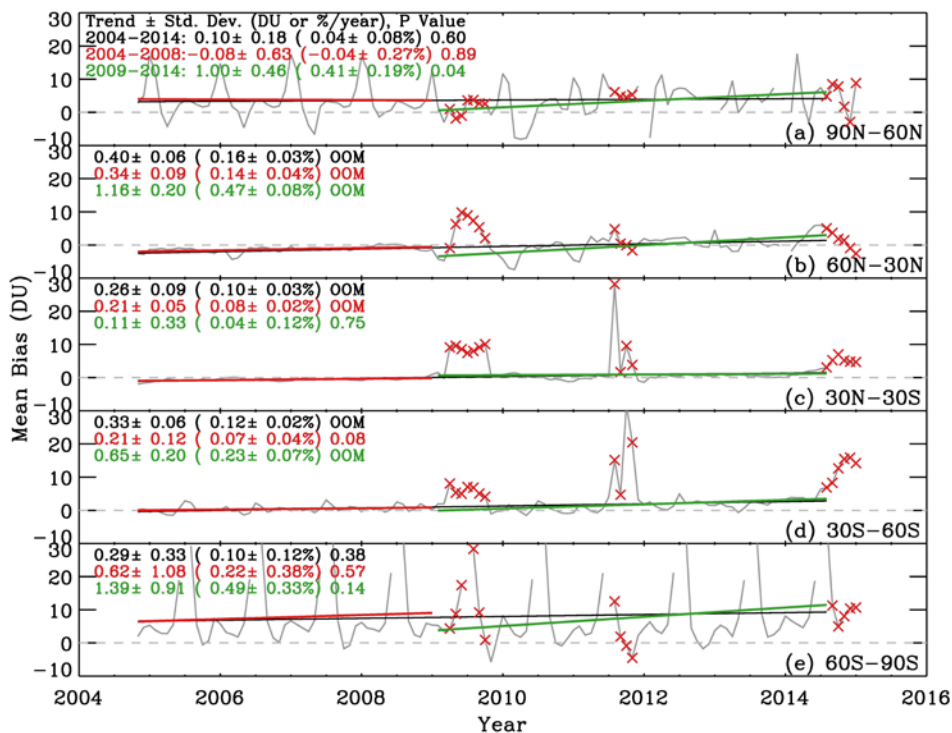
Figure 11. Monthly mean biases of OMI-MLS stratospheric ozone columns with pressure less than 215 hPa are plotted in solid gray lines for five different latitude ranges. <u>Extreme small p values are marked as OOM (out of memory).</u> The corresponding linear trends in 2004-2014, 2004-2008 and 2009-2014 are shown in black, red and green lines, respectively. The red cross signs indicate the periods when OMI has unusually large SOC errors likely due to row anomaly or other unknown issues; the values during these periods are not included in our trend calculations. The legends show linear trends in both DU/year and %/year and the corresponding P values.