Response to referee #2's comments

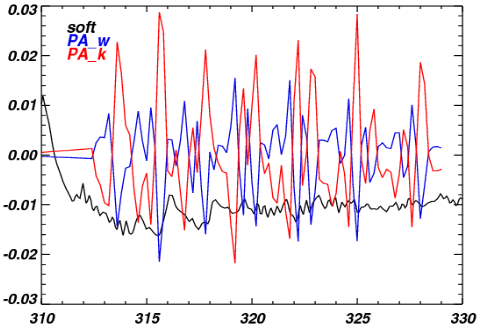
The author would like to thank Anonymous referee #2 for the constructive and helpful suggestions on this manuscript.
We replied to 3 major comments and 11 technical comments.

General Comments

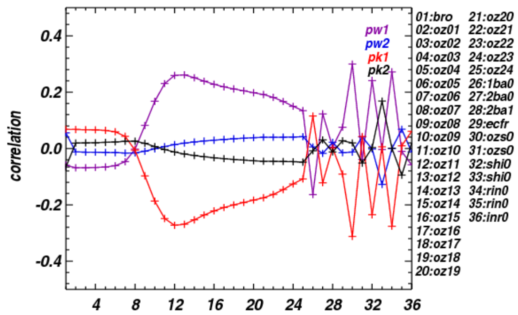
This paper is well organized to describe a methodology for reducing the spectral fit residuals. The subject of the paper is appropriate to AMT. Below are a few comments concerning clarifications / extensions for consideration in the final publication in AMT.

Major comments

C1. The PROFOZ algorithm applies the pre-estimated, pixel dependent "soft calibration" factors to the normalized radiances, while conducting the ozone profile retrievals. The "soft calibration" factors seem, by design, accounting for the imperfectness of OMI L1B earthshine radiances and solar irradiances calibration, parameterization of the pixel & wavelength dependent ISRFs, and forward model parameters (absorption cross-sections, surface albedo) etc. The PROFOZ also fits scaling factors for the pre-estimated mean spectral fit residuals (Liu 2010 a, b) for UV1 and UV2 bands accordingly, to account for the remaining systematic errors that were not fully removed from "soft calibration" process. This work suggests fit additional ISRF PA coefficients is necessary for OMI ozone profile retrievals. It seems there might some degeneracy among these approaches. The authors should elaborate whether employing pixel & temporal dependent 'soft calibration' factors, or fitting the mean spectral residuals could also achieve the goals same to employing the presented PA approach, in terms of reducing the spectral fit residuals. Are the Jacobians of these PA coefficients, orthogonal to the pre-estimated mean fitting residual spectra, or any other Jacobians of parameters in the retrieval vector?
R1. - As this review pointed out, the soft calibration could partly take into account the remaining systematic errors including the spectral structures due to slit function errors, but it should be taken as a last resort after the known physically treatable errors are considered separately. The applied soft spectra were derived from clear-sky tropical measurements in July 2006 and then applied to everywhere and every day. However, the PAs are calculated at each satellite pixel based on the physics associated with slit convolution proposed in Berlie et al. (2017), and iteratively adjusted with the retrieved coefficients. Therefore, the presented PA approach works much better than soft calibration to reduce the fitting residuals and retrieval errors caused by slit function errors.
- Several peaks of soft spectrum are matched with those of PA jacobians, but the soft spectrum is uncorrelated with PAs (with correlation less than 0.1 in UV1 and 0.3 in UV2) because of other dominant factors causing much higher spectral residuals in the soft spectrum (S1). In addition, PA spectra show a weak correlation with other Jacobians within 0.3 for UV1 variables, but for within 0.1 for UV2 variables (S2). In the revised manuscript, this discussion has been added such as "It should be noted that these spectral structures are weakly correlated with the partial derivatives of radiances with respect to other state vectors (ozone, BrO, cloud fraction, surface albedo, radiance/irradiance shift, radiance/ozone cross section shift, Ring/mean fitting residual scaling factor) within $\pm$ 0.3 and $\pm$ 0.1 in the UV 1 and UV 2, respectively."



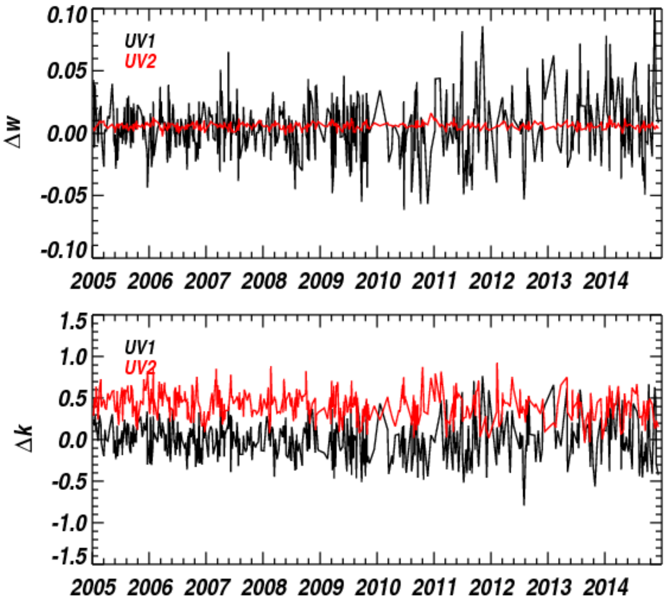
S1. Comparison of soft spectrum with PA jacobians with respect to $\Delta \boldsymbol{w}$ and $\Delta \boldsymbol{k}$, respectively.



S2. Correlation of jacobians from the other fitting parameters with PA jacobians with respect to $\Delta \boldsymbol{w}$ and $\Delta \boldsymbol{k}$ for UV1 and UV2, respectively.

C2. The authors should obtain time series of retrieved ISRF PA coefficients. Do they show trends similar to Figure 1? At least for Nadir pixel, if not all pixels.

R2. Fig. 1 show the time series of slit function parameters derived from solar irradiance measurements. While the PA coefficients show the deviation from those in Figure 1, so they are not expected to show similar trends as shown in S3. In addition, the PA coefficients can vary from spatial to spatial pixel, and vary along the track for the nadir pixel, so it is not as straightforward to obtain the time series. However, this time series also show the larger variation later in OMI mission, especially in the UV1 due to radiometric calibration issues.



S3. Time-series of PA coefficients for UV1 and UV2, respectively, spatially collocated to 4.3°E, 50.8°N.

C3. A) The authors evaluated the impacts of with/without retrieving PA coefficients on the bias/RMS between retrieved ozone and in-situ ozonesonde measurements (Figure 9). However, the evaluation only made for the period of 2005 to 2008, when OMI instrument was within design lifetime. The authors should also evaluate the performances using the satellite-ozonesonde measurements in other time periods including 2010 and 2012-2013, when the ISRF characteristics were significantly different than the earlier years, as shown in Figure 1. B) The authors should also add some discussions on the possible reasons

causing these sharp changes of ISRF characteristics.
R3-a. As well known, there has been concern over the row anomaly effects appearing in 2007 and becoming serious in early 2009, causing trend errors of OMI tropospheric ozone as reported in Huang et al. (2017). Therefore, the period of 2005 to 2008 is focused on the evaluation of including PAs on ozone profile retrievals to avoid any interference with row-anomaly impact. “This evaluation is limited to the period of 2005 through 2008 to avoid interferences with row-anomaly effects appearing in 2007 and becoming serious in early 2009 (Schenkeveld, et al 2017)” has been added in Section 3.2 of the revised manuscript to clarify why the period of 2005 to 2008 is targeted.

R3-b. To explain the sharp changes of ISRF characteristics, “The sharp change and random-noise of these derived slit function parameters might be influenced by the decreasing signal-to-noise ratio (SNR) of solar spectra later in the OMI mission and radiometric errors in solar irradiance due to row anomaly (Sun et al., 2017).” has been added in the revised manuscript.

Technical comments

C1. Have the authors evaluated the impacts of this methodology on the L2 retrieval throughput/yields?
R1. There is no significant impact on throughput. The number of successful retrievals for one orbit measurements is 10880 (standard Gaussian, w/o PA), 10880 (super Gaussian, w/o PA), and 10884 (standard Gaussian, with PA), and 10883 (super Gaussian, with PA)

C2. Line 29, use the statistical numbers on the bias/RMS differences to replace the word “substantial”.
R2. The manuscript has been revised to accept this comment as followings.
- (Abstract) “Comparisons with ozonesondes demonstrate noticeable improvements with the use ofwhen using PAs for both standard and super Gaussians, especially for reducing the systematic biases in the tropics and mid-latitudes (mean biases of tropospheric column ozone reduced from -1.4 ~ 0.7 DU to 0.0 ~0.4 DU) and reducing the standard deviations of tropospheric ozone column differences at high-latitudes (by 1 DU for the super Gaussian).”
- (Line 329) “clearly shows that including PAs to account for ISRF differences significantly reduces mean biases below 10 km” ➔ “clearly shows that including PAs to account for ISRF differences significantly reduces mean biases ofby up to ~ 5 % below 10 km”
- (Line 383) “Using super Gaussians, the TCO comparison shows significant improvement in mean biases in mid-latitudes and in standard deviations in high-latitudes. Using standard Gaussians, the TCO comparison also shows significant improvement in mean biases in the tropics” ➔ “In the TCO comparison between OMI and ozonesonde, the mean biases are reduced by 0.2 (0.6) DU and 0.6 (1.4) DU in the tropics (mid-latitude) when super and standard Gaussians are linearized, respectively.”

C3. Line 47, the authors should consider to revise “by narrow and weak absorption features of the temperature-dependent Huggins bands (320-360 nm)” to “by the 320-330 nm absorption features residing in the temperature-dependent Huggins bands.”, since neither this work nor the referenced studies utilized spectral region > 330 nm in the OMI ozone profile retrievals. “narrow and weak” are general terms and might subjective, e.g., this statement will break down. When comparing within the Chappuis bands, the refereed portion of Huggins bands (>320 nm) is no longer weak.
R3. According to this comment, the indicated sentence has been revised to “by the 310-330 nm absorption features residing in the temperature-dependent Huggins bands”.

C4. Line 50, I will suggest to cite the following studies on OMI ozone profile retrievals, since [1] they made use of the ISRFs from Dirksen et al., [2006] cited a few times in this work, [2] the quality evaluation have been conducted by the comparison with in-situ ozonesonde measurements, [3] same to Liu et al., 2010 cited in this work, these studies were conducted prior to the era of including PA coefficients in the retrieval vector.

R4. We appreciate this suggestion. The suggested references have been cited such as "For space-borne instruments, ISRFs are typically characterized as a function of the detector dimensions using a tunable laser source prior to the launch (Dirksen et al., 2006; Liu et al., 2015; van Hees et al., 2018) and directly used in ozone profile retrievals (e.g., Kroon et al., 2011; Mielonen et al., 2015; Fu et al., 2013; 2018)"

C5. Line 60, might be a typo (radiance repeated twice)?. Do the authors mean "differences in stray light between radiance and irradiance" or "differences in stray light among OMI measurements"?

R5. It is printed-word. It should be "differences in stray light between radiance and irradiance"

C6. Line 61, It seems that "intra-orbit instrumental changes" is duplicating the statement of "the instrument temperature change". Please clarify (or remove one).

R6. It has been clarified such as "Slit function differences between radiance and irradiance could exist due to scene heterogeneity, differences in stray light between radiance and radiance, and intra-orbit instrumental changes (such as instrument temperature change)."

C7. Figures 1, 2, 4, 5, 7, 8, 9 and 10, increase the tick length for improving their visibility.

R7. All figures have been revised for better visibility.

C8. Figures 5, 6, 8, 9 and 10 captions, state the date/time range of the data presented in the figures. It is not where they are all for 1 July 2006, shown in Figure 4 caption.

R8. In the revised manuscript, all captions include the date/time range of the data.

C9. Figure 9, create a table and move the statistical values to the table. Having all these numbers on the plots resulted in the plots being too busy to read.

R9. The corresponding figure has been changed to Table1.

C10. Figure 10 # please spell out the "MB" and "SD" in the x axis title, - space suffice to hold the full name and they were not defined in the caption. # Add two panels to show the differences among data sets, as a function of altitude?

R10. The a-axis titles have been changed to Mean Bias and Standard Deviation, respectively.

C11. Finally, please keep the 'style' of all figures in a similar fashion. e.g., the panel index of Figure 2 (a), (b) and (c) are inside the plots, while the other figures are outside of the plots. I understand that there is no space for the subtitles outside Figure 2b and 2c, due to the x axis labels. The authors should consider to remove those x axis labels, since all panels could share the one of panel c. Similarly, there are unnecessary axis labels in other figures, e.g., Figures 4, 5, 6, 7, 8, 9, and 10, when some subpanels having an identical scale/range across a row and/or a column, authors should consider remove the unnecessary labels in x or y axis, to help readers easily catch key information presented in the figures.

R11. Thanks for this detailed suggestion. All figures have been revised for better visibility.