Overall Comments:

I find that this is a carefully crafted paper, which I appreciate. I can tell the authors put a lot of effort into error assessment and to describing their algorithm improvements. They describe a v2 product that is a significant improvement over an already good performance by the v1 product. However, I believe the authors vastly understate the importance of their soft calibration approach in the success of the OMPROFOZ product. Nor do the authors adequately discuss the information content of their product given their soft calibration approach. The derived ozone profiles are effectively normalized to equatorial MLS profiles once per year. Therefore, the OMPROFOZ product is one that describes extra-tropical ozone profile variability and intra-annual ozone variability relative to MLS ozone profiles. There is nothing wrong with this approach, nor is it even new to this version. A product such as OMPROFOZ that better describes UTLS ozone variability is quite valuable. But the authors should clearly state up front (i.e. in the abstract or in the conclusions) what this product is and what it is not. In particular, it is not a product that independently measures long-term changes in ozone profiles. Their soft calibration approach is central to the definition of this product, yet the authors treat it almost as an afterthought with little mention throughout this paper. One gets the impression they themselves do not fully appreciate the role this normalization plays.

.→ We appreciate the sincerely comments, and tried to improve the manuscript, in accordance with reviewer's suggestion.

Specific Comments:

C1. Section 1, Line 51 Remove "been" C2. Section 1, Line 56 Insert "to" before "evaluate"

R1 & R2 Accepted.

C3. Section 1, Line 66 Since this is the first reference to MLS in this paper, please indicate that you refer to the AURA instrument.

R3. The manuscript has been edited to accept this comment; "*Ten-years of the OMPROFOZ product were assessed in-depth in Huang et al. (2018;2017) through the spatiotemporal validation using global reference dataset collected from balloon-borne ozonesondes and space-borne Microwave Limb Sounder (MLS), which is one of the payloads onboard the Aura satellite, along with the OMI instrument*."

C4. Section 2.1, Line 122 In the interests of full disclosure the authors should inform the readers that the row anomaly affects ALL of the UV1 channel. There are no rows that are reliably free of the effects of the anomaly, though longer wavelengths tend to be less effected than shorter ones, and lower rows less than higher ones.

R4. Thanks to specify the row anomaly issue in detail. The row anomaly is a well-known issue. The anomaly affected measurements are uncorrectable. In this section, we need to address which flagging are applied, rather than the detailed explanation of the anomaly behavior for assuring the conciseness and consistency of the context. Therefore, Schenkeveld et al. (2017) is referenced for the detailed behavior of the row anomaly in the revised manuscript.

C5. Section heading 2.1 is duplicated. Section 2.2 is missing.

R5. The number of the section has been corrected.

C6. Section 2.2, Lines 171-175 The authors state that the measurement noise reported in the OMI L1B product underestimates the true noise. This is a well-known problem across multiple instruments. It occurs because pseudo-random systematic errors, either in the measurements or in the model, are much larger than detector noise. The authors go on to state that their assumed minimum errors are uncorrelated between wavelength, but this seems to be a rather poor assumption. Many modeling and measurement errors are spectrally correlated, e.g. cloud modeling errors. Can the authors comment on the effect such an assumption has on their product?

R6. We agreed with what this reviewer addressed about the assumption on the measurement noises. The detector can be interfered between adjacent pixels. However, there is no data source about the information on the correlation behavior between wavelengths for OMI and other similar instruments. It is common to construct the measurement error covariance as a diagonal matrix, where measurement errors are assumed to be uncorrelated among wavelengths, as suggested by Rodgers (2000) who introduced an inverse theory based on an optimal estimation. There is no literature about the BUV ozone profile retrievals in which the spectral correlation of measurement errors is applied, based on our knowledge.

C7. Section 3.2, Line 298 Begin this sentence with, "In Version 1 these meteorological variables ..."

R7. Accepted.

C8. Section 3.5, Line 358 Unlike other BUV instruments, normalizing by the OMI measured solar irradiance does not reduce optical degradation errors. With OMI there are several optical elements not shared in common between Earth radiance measurements and solar irradiance measurements. In the case of solar irradiance these elements (diffuser and folding mirror) represent the primary sources of degradation. As a consequence, the Coll. 4 degradation corrections for radiance and for irradiance are completely separate. Furthermore, the degradation correction for irradiance measurements was derived by assuming constant solar irradiance over the mission (Figure 5c demonstrates this point). This is a reasonably good calibration approach for wavelengths longer than 300 nm, but not for the UV1 channel. There are clearly benefits to normalizing OMI radiances with a time-dependent solar irradiance, but cancellation of long-term optical degradation is not one of them. Having said all this, it's not clear why the authors are concerned about optical degradation given their soft calibration approach outlined in Section 3.8. Perhaps the authors should make it clear in Section 3.5 that their interest in improved solar irradiance measurements is solely to address seasonal variations in instrument calibration.

R8. Thanks for specifying the calibration characteristics for OMI. In previous version, the data processing uses the climatological solar spectrum derived from three years of daily OMI Level 1B product (2005-2007), which is switched to the monthly solar irradiances in the new version. with . As addressed by this review, applying the monthly solar irradiances does not reduce the degradation errors, but offers several benefits in reducing short-term noises and in addressing the seasonal variations of instrument characteristics that are common in radiance and irradiance measurements, compared to applying either climatological spectrum or daily spectrum. As shown in Fig.7.b, the fitting quality is improved due to switching from daily irradiances to monthly mean irradiances due to improved signal to noise ratio after averaging. According to this comment, we have revised the manuscript to address the seasonal variations in instrument calibration, with the improved solar irradiance spectrum as follows.

"*In order to reduce the short-term noise of individual measurements, the earlier algorithm implemented the use of climatological solar spectra derived from three years of OMI collection 3 irradiances (2005-2007). In the newer algorithm, collection 4 irradiances are tabled as monthly averages to address seasonal variations of instrument characteristics that are common in both radiance and irradiance measurements.*"

C9. Section 3.1, Lines 369 - 371.What does "implementations are identically applied" mean? If you mean to say that Coll. 3 and Coll. 4 data are treated the same for this experiment, please say so.

R9. As described in the original manuscript, Figure 7 shows the retrieval experiment to see the impact on the spectral fitting residuals due to switching OMI L1b data from collection 3 to 4. The radiance is mostly unchanged and so this comparison represents the impact due to switching irradiance. We stated that "In this experiment, the v2 implementations are identically applied without radiometric corrections (soft calibration and common mode correction are commonly turned off)". It means that both ozone profiles are retrieved with the v2 algorithm, but soft calibration and common mode correction are commonly turned off.

C10. Section 3.6 The authors could instill confidence in their pseudo-absorber approach to scene inhomogeneity if they were to demonstrate that their empirical parameters correlate with scene reflectivity changes or the small pixel column results contained in the OMI Level 1B product. The largest slit function errors will occur in the along-track direction at cloud edges, which can be identified via scene reflectivity or the small-pixel data. However, if such a comparison has already been shown in the referenced paper the authors may ignore this comment.

R10. We appreciate this important point. In the reference paper (Bak et al., 2019b), we already showed the correlation between pseudo-absorber coefficients ($\triangle$w, $\triangle$k) and scene reflectivity changes as well as the improved fitting residuals with their pseudo-absorbers being implemented.

C11. Section 3.8 Per the authors' description in lines 454-463, OMI calibration is adjusted so that the retrieved ozone profiles match MLS + LLM profiles in the tropics. Per the description provided, this normalization occurs during the northern summer every year. While such an approach should help deal with systematic biases caused by the row anomaly, a once-per-year correction is inadequate to deal with the variable nature of the row anomaly. The authors should address the question of what intra-annual TCO errors may remain after the once-per-year soft calibration corrections.

R11. We would like to clarify the soft calibration. The soft calibration is intended to correct the slowly varying systematic biases, not to correct the rapidly varying biases. Therefore, soft calibration is composited of the spectral residuals between measurements and simulation, where the forward model simulation could be as accurate as possible, such as the summer tropics at nearly clear sky conditions. The ozone profile input is one of the main sources of errors in the forward model calculation and thereby the summer tropics is targeted with the fact that the daily ozone variability is relatively small. We prepare ozone profile inputs by merging MLS in the stratosphere and LLM in the troposphere, with the profile shape adjustment based on OMI column ozone. The soft calibration is an empirical correction and therefore they should be carefully derived and carefully applied. Therefore, we derived the soft spectra only as a function of instrument dimensions for each year, assuming that the soft spectra are representative of the slowly changed degradation errors, and implemented the soft calibration to correct the daily measured spectrum by interpolating the yearly given soft spectra. Both soft calibration and common mode correction either cannot account for the remaining calibration errors or induce the artificial errors somewhere. These could be the answer for some intra-annual TCO errors. However, Figure 14 still demonstrates the improvement of the seasonal bias with the update to PROFOZ version 2.

C12. Section 3.9 The authors should attempt to provide a physical explanation for the observed intra-orbital variations in residuals. Without a reasonable explanation how can the authors or the readers be confident a static CMC correction is appropriate and adequate? The most likely explanation for the observed residual variation is additive errors (e.g. stray light) and the row anomaly. Will the CMC as the authors have implemented it address errors introduced by stray light and the row anomaly?

R12. As mentioned in the original manuscript, the intra orbital variations in residuals implies the existence of a spectral dependence of the radiometric calibration and detector sensitivity on the signal represented by solar zenith angle, which is not completely treated by the soft calibration dependent only on CCD dimension. As mentioned in this comment, the solar zenith angle dependent features could be related to the straylight. It is assumed that the row anomaly contaminated pixels are screened out in this retrieval experiment. We observed the consistent component of fitting residuals, with the dependence on the solar zenith angle (Figure 14 d/e/f). Therefore, we carefully composited the common mode spectrum and then fitted them as a pseudo absorber. Comparing Figure 14 d-f and Figure 14 g-I demonstrated the improvement of the spectral fitting quality and comparing Figure 8.f and Figure 12.c also confirmed the improvement of the retrieval quality at cross-track positions 1-5.

C13. Figure 13 The MB and Std. Dev. labels appear to be reversed.

R13. Corrected.

C14. Figures 13 & 14 Given that the OMPROFOZ product is tied via soft calibration to MLS+LLM, it will be helpful to show readers similar comparisons of MLS+LLM to the same ozonesonde measurements (or at least measurements from the same stations). This may provide insight into how much of the observed TCO-sonde difference arises from the choice of soft calibration.

R14. As addressed in replying to comment 11, our soft spectra are derived as systematic biases due to slowly varying radiometric calibrations, not due to the forward model calculations in which MLS + LLM is used as an input once per year. Therefore, our ozone profile is not tied much to MLS + LLM. This paper aims for describing the algorithmic updates and its verification in details. In this study, the validation is limited at three ozonesonde stations due to the computation limit in the local machine. The PROFOZ v2.0 processor is being tested in the NASA OMI operational facility for reprocessing the entire mission. When the new product is available, we are planning to perform the global and long-term validation activities for both tropospheric and stratospheric ozone retrievals against ozonesonde and MLS measurements.

C15. Section 4.0, Line 583 The authors should avoid referring to the flags as "TOMS-based" and instead continue to reference the OMUANC product as the source of these flags.

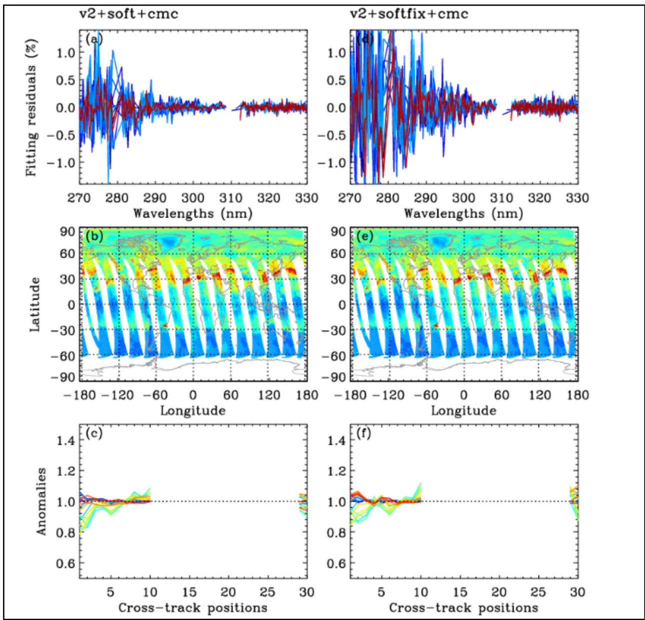
R15. We have revised the manuscript to accept this comment as follows: "Those spikes could be attributed to row anomaly-contaminated retrievals unscreened with the row anomaly flags taken from OMI collection 3 L1b product. The related improvements in OMPROFOZ v2 retrievals are contributed by applying the stricter flags taken from the OMUANC product."

C16. Section 5.0, Lines 628-630 This brief mention of the soft calibration understates the role it plays in the performance of this product. The text suggests that its role is to merely keep the simulations close to the measurements, perhaps to keep them in a more linear regime. Surely, in an iterative retrieval algorithm that typically requires no more than 2-3 iterations (Section 2.2) dependence on the initial guess is not strong. The authors should acknowledge that the primary role played by their soft calibration is to eliminate the long-term drift observed in v1, and to remove some of the static and slowly varying row anomaly errors that have hitherto stymied all other attempts at retrieving ozone profiles from OMI data.

R16. The lines 628-630 provide a summary of the implementation changes from v1 to v2. Therefore, we stated the update to the soft calibration as "*The empirical soft calibration spectra are re-derived annually to be consistent with the updated implementations to remove the systematic differences between measured and simulated radiances.*". In both v1 and v2, the soft calibration plays an important role in eliminating the CCD-dependent systematic errors, in particular for reducing the stripes on the tropospheric ozone retrievals and the spectral fitting residuals. The main update to soft calibration (v1→2) is expanding the dimension of calibration spectra to account for the degradation error. Actually, this work was initially experimented in the frame of collection 3 and the soft calibration played the primary role for eliminating the long-term drift. However, this role is reduced in the frame of collection 4. Therefore, the soft calibration plays a role in accounting for the remaining degradation errors.

C17. Section 5.0, Lines 643-648 The authors imply that the improved long-term drift is somehow related to switching from Coll. 3 to Coll. 4 and to some unidentified implementation details. Given the soft calibration approach in v2 it is unlikely that the improved calibration in Coll. 4 plays any role in this.

R17. The major updates from collection 3 to collection 4 is mostly related with the solar irradiance degradation correction as well as the improved quality flagging. The soft calibration is employed for accounting for the systematic biases of the normalized radiance (radiance/irradiance) as a function of CCD dimensions for each year. As indicated by this reviewer, the dependence of the data quality (PROFOZ v1 → v2) on the soft calibration is reduced with respect to the long-term consistency, due to the improvement of L1b data quality (collection 3→4). However, the temporal dependent soft calibration plays an important role in improving the data quality and consistency. Therefore, we decided to implement the yearly soft calibration in the v2 implementation. Please take a look at the following figure, comparing the retrieval results with yearly soft calibration and the 2005 soft calibration, respectively.



Same as Figure 12 in the manuscript. (a, d) Spectral fitting residuals (%) averaged in the latitude of 60ºS and 60ºN from OMI measurements on 15 June 2018, (b,e) the global distribution of tropospheric column ozone (TCO, DU), and (c,f) anomalies of TCO as a function of 18 latitude bands, but for comparing the retrieval results on 2018m0615 with yearly soft calibration and the 2005 soft calibration, respectively Note, the 2005 soft calibration represents the correction spectrum derived from measurements on 2005m0711-17.