RC2 - answers

General comments

In general, I find this manuscript version improved over the initial submission. The authors have done a better job of organizing results to make clear when corrections are and are not applied. More explicit statements for each figure would still be appreciated. It is clear that the new instrument characterization described in this paper is improving the accuracy of the radiance data (thereby reducing the magnitude of soft calibration corrections), but some of the details regarding how those improvements were achieved is lacking. It is only by providing such details that readers can evaluate the quality of the characterizations and learn something about instrument performance.
We thank the reviewer for the comments and appreciation. We addressed each comment, updating the manuscript where necessary.

Section 2.2, Lines 135-138

This text describes the contents of Figure 2, which presents a compelling case that the Mg line signals have increased over time relative to background signals. Unfortunately, the reader has no idea from this figure what the time scales are, since the paper provides no assignment of dates for S5P orbit numbers. Two common reasons these Fraunhofer line changes occur is solar activity and improper correction for additive signal errors. The most likely additive errors are detector dark current and spectral stray light. All three phenomena are valid explanations for why the Mg line depths shown in Fig. 2 are decreasing. The authors state that the Fig. 2 results are independent of solar activity effects, but offer no details supporting this statement. Was a correction applied, and if so what was it? They do not discuss their background signal corrections, which must also increase as the instrument ages. These are glaring omissions if they expect the reader to accept the most surprising conclusion, that internally-scattered stray light is increasing within TropOMI. There are few plausible mechanisms for postlaunch increases in spectrometer scattering (primarily caused by the grating), so the authors need a convincing argument why the least likely of the 3 explanations is in fact the actual cause of the observed signal changes. One convincing metric would be to compare the Mg II core-to-wing ratios from solar irradiance and Earth radiance spectra at different times in the mission. In the absence of additive errors these two ratios should remain in lock step with each other. If, however, spectral stray light is increasing, the effect on Earth radiances should be much greater than on solar irradiances.
The authors highly appreciate the reviewer's detailed analysis. We did not want to exclude that additive effects are present. At line 141 (previous manuscript), we therefore remarked that the additive effects will be discussed later in the manuscript, in the Sections about the straylight and residual correction updates. We agree that all three

phenomena can cause the Fraunhofer line changes (solar activity, uncorrected additive straylight, uncorrected additive dark signal), however it is difficult to disentangle them.
Our claim that straylight is one (and not "the") actual cause is based on the following: 1) in-flight growth of signal in the straylight rows; 2) the large changes in the first E2 years, when solar activity is still low (before 2020); 3) the smooth behavior of far-field straylight, showing a single peak at 280nm in comparison with solar activity changes showing a double peak; 4) the improved correlation between Mg and Ca lines time series when implementing the dynamic straylight correction (image included in Appendix A1(b)).
The effect of detector dark current signal in irradiance is small, but it is true that it can be relatively more important for radiance. The residual correction, described in Section 2.3, is implemented in both radiance and irradiance measurements and it is meant to correct for remaining additive effects (therefore also detector dark current and RTS). The soft-calibration spectra, pictured in Figure 9c-d, also show improvements regarding across-track biases when using input data with residual correction implemented.
Regarding the suggestion of comparing the Mg line indices of radiance and irradiance: only if all additive errors - especially in radiance - have been successfully removed, then the correlation of the time series of the line indices are perfectly aligned. However, we do not claim that that is the case, otherwise the bottom panel in Figure 9 would be much smoother.
We indeed reprocessed ~200 orbits over the entire mission that includes both radiance and irradiance. Due to restrictions, it was not possible to regenerate a consistent set of the (ir)-radiances without the mentioned three improvements: the radiance measurements of the existing dataset have been processed with CKD that have been updated several times, with large sensitivities at the 280nm FL. The part of the timeseries that is consistent, from April 2024 onward, indeed shows better correlation (not shown).
<u>Manuscript adjustment:</u> Figure 2 has been modified by adding the dates and showing the effects of the dynamic straylight on the peaks at 280nm and 286nm. The text of Section 2.2 has been also updated to add more explanations on the reasoning behind the exclusion of solar activity, with companion plots in Appendix A1 and A2, showing improved correlation between the Mg and Ca line indices for irradiance measurements. The correlation improves, however it is not perfect.

Section 2.2.1, Line 153

The statement that the SL correction algorithm is a 2D correction that only addresses in-band stray light is confusing. I suspect the authors mean the correction address both spatial and spectral stray light, but the latter is limited to only photons from within the same band. If so, a clarification of this point would be helpful.
The kernel is 2D, and its convolution with the 2D (spatial-spectral) detector image results in a straylight image containing both the spatial and spectral straylight. The kernel operates on an entire detector image, e.g. UV, VIS, NIR or SWIR detectors, with each

detector consisting of two electronic bands. Perhaps here the term "band" is causing confusion.
Manuscript adjustment: we deleted the phrase in line 154 as the out-of-spectral range straylight is not relevant for the manuscript and it interrupts the flow regarding the in-band straylight. We updated also the text of this section to enhance clarity.

Section 2.2.1, Lines 179-189

The authors should provide a clearer explanation of how the SL convolution kernels are adjusted. What criteria were used to assess the correct spectral and spatial distributions? Scattering occurring after the entrance slit typically has no preferential spatial or spectral direction. The 2D scattering PSF is usually symmetric. Therefore, the large difference in the two orthogonal cross sections of that PSF (shown in Fig. 3) is rather surprising and the authors are encouraged to provide more detail about why they feel the shapes should behave so differently. I presume that telescope stray light (i.e. pre-slit scattering) has been excluded from the plots in Fig. 3a. Please state so explicitly.
Yes, symmetric scattering has indeed been our first assumption. The problem is that the on-ground calibration measurements (white light source measurements with small spatial range) led to non-unique solutions for the convolution kernel. Giving the on-ground measurements, non-axisymmetric kernels were in principle equally valid as solutions.
The first E2 radiance measurements were clearly over-corrected in band 1, resulting in negative radiance signals after the straylight correction. Given that radiance images have more variation in the spectral than spatial direction, the revised kernel shape, stronger in the spectral dimension, would be less strong in the spectral direction while keeping all its other properties deduced from the on-ground calibration measurements. This is why the elliptical kernel shape was considered an improvement over the original symmetrical shape: it is still a valid (alternative) solution allowed within the entire set of on-ground measurements, and giving better corrections on the in-flight measurements.
Manuscript adjustment: We modified section 2.2.1 to clarify the choice of the elliptical kernel shape. A new image has been added in Appendix to show the comparison among the original, the elliptical kernel shape, and a third case with a long tails kernel shape. All these kernel shapes are valid given the on-ground calibration measurements.

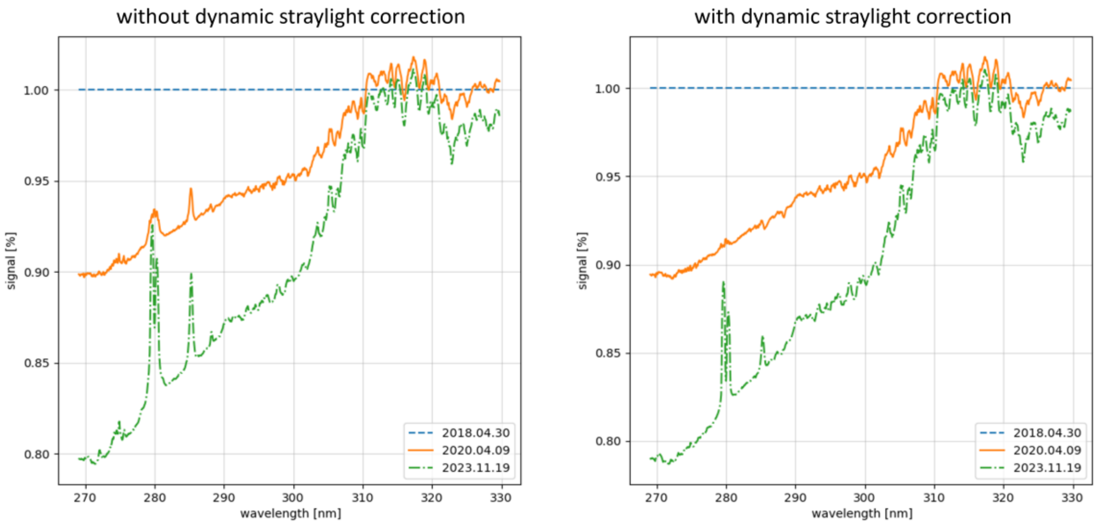
Section 2.4

The authors mention the introduction of time-varying SL correction to account for the increasing Mg II signals and the SL row signals, but they provide no details about how this is implemented. Are the kernel shapes altered in time? If so, how? Are the spatial and spectral components of the kernel locked relative to each other or allowed to vary independently? Whatever the approach, it will be helpful if the authors can demonstrate the effect of these dynamic corrections on the metrics shown in Figures 2 and 4. Figure 10 does indicate the effect on RTM residuals of these instrument corrections, but it does not address the effect on irradiance or stray light rows.

Thanks for the comment and the opportunity to improve the clarity of the text and the analysis. The convolution kernel shape (both the original and elliptical version) is static in time. The dynamic straylight correction is instead based on the instantaneous measurements obtained from the straylight rows, hence the attribute "dynamic". The only difference between the original and elliptical kernel is the asymmetry in the spectral dimension, as the mass (sum of all elements) of the convolution kernel is the same in both cases. Again, we mention that we make use of some freedom in proposing slightly different kernels that are still valid solutions with respect to the on-ground measurements. To show the effect of the introduction of the dynamic correction on the degradation of un-corrected irradiance measurements, we updated Figure 2.

Regarding Figure 10: the soft-calibration is applied on radiance measurements.

Manuscript adjustment: Figure 2 has been updated with the following new figure, showing the effect of the dynamic straylight correction on the degradation of un-corrected irradiance measurements. Section 2.2.1 was also updated to clarify the difference between the updated convolution kernel (static in time) and the newly added dynamic (time dependent) straylight correction.



(a) (b)

Section 4.3

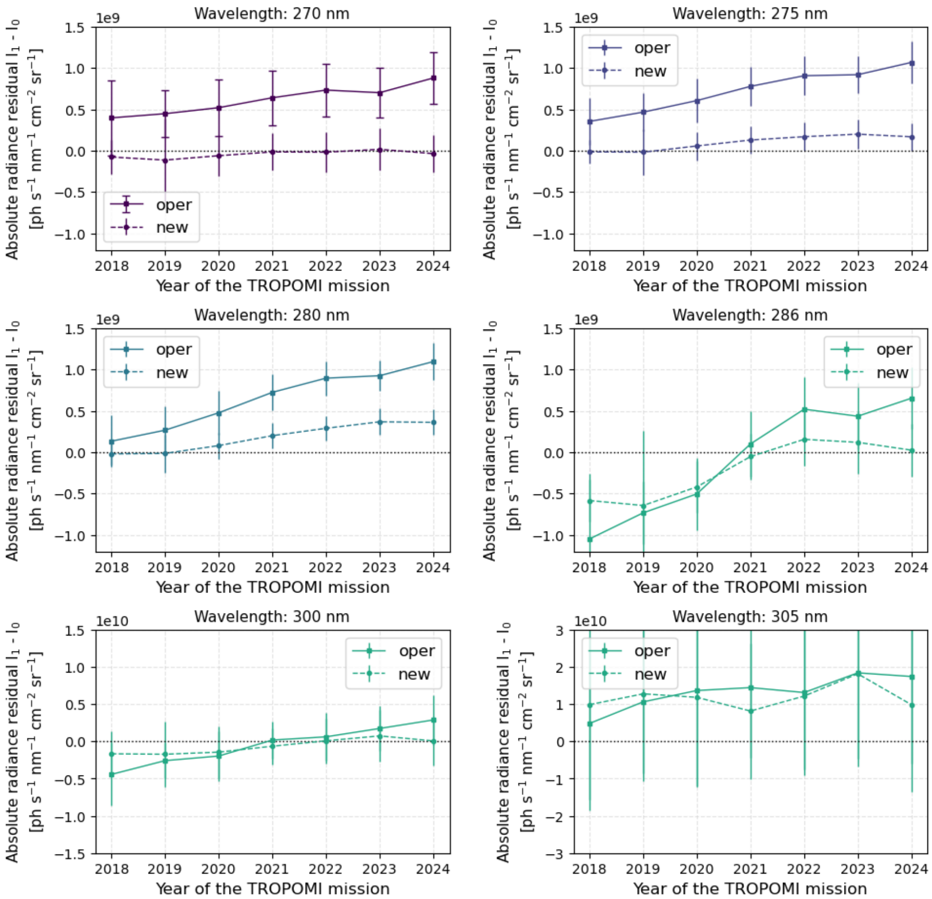
Regarding the observed residuals (as represented by soft calibration corrections) presented in Figures 7 & 8 the authors do not discuss whether or not a solar activity correction has been applied. The increase seen in Figure 7a is consistent with the activity increase over the mission timeline. Figure 8 also appears to be consistent with the increase. An accurate solar activity correction must be applied prior to residual analysis, therefore the authors should provide details about any correction that was applied.

What data was it derived from? What is the magnitude of the changes in time at key band 1 wavelengths?

There is no solar activity correction applied to the radiance and irradiance spectra used for the comparison with forward model calculations.

Section 3.1 describes the correction steps applied to the L1 (bands 1-2) measurements before the soft-calibration correction (see Fig. 5), as pre-processing steps of the ozone profile retrieval: spectral calibration, spatial regridding, spectral binning, signal-to-noise ratio floor, polarization & raman correction.

We add the following image to show the magnitude of the absolute difference between the un-corrected radiance and the modeled radiance at key band 1 wavelengths and at two band 2 wavelengths, for a single central ground pixel number 35. The solid line ("oper") stands for the current processing, while the dashed line ("new") shows the time change of the absolute residuals using the updated bands 1-2 measurements. These figures also clearly show the decreased magnitude of the bias and the decrease time dependence of the bias over the mission.



Manuscript adjustment: we realized it might not have been clear that the bands 1-2 pre-processing described in Section 3.1 were the corrections applied to the input spectra before the soft-calibration correction. Therefore, we moved the description of the measurements pre-processing under the main section on the *TROPOMI UV soft calibration*. Moreover, we changed the last step in Figure 5 from “Radiometric calibration correction” to “Soft-calibration correction” to enhance clarity.