Response to anonymous Referee #1

Referee comment	Author's response	Rework performed
General comment: The paper is well written and of good quality, with a considerable number of new interesting topics and techniques, and shall certainly be published. There remain a number of issues that I think would improve the quality of the paper, also in comparison with similar papers of other missions and instruments (for comparison), in line with the comments and suggestions provided below. After these comments and suggestions have been addressed, the paper shall certainly be published. Since this is the second submission I leave it to the scientific editor to decide if and how he/she wants to proceed with the implementation of the comments and suggestions below.	We would like to thank the referee for his/her valuable review.	Added acknowledgement.
Comment 1 a + b: Table 6: I (still) think that the uncertainties presented in table 6 for absolute radiance, absolute irradiance and BSDF are unrealistically low at 1-sigma, given the complications and issues described in the text in this paper: a. There is a relatively large unexplained measurement discrepancy that dominates ABSRAD and BSDF. b. There is a relatively large unexplained FEL lamp discrepancy in ABSIRR and BSDF.	The errors given in the table exclude the errors that are caused by the stimulus. We know these errors exist, but cannot quantify them due to the lack of independent measurements. It is true that this is a large source of error in the uncertainty budget, which we do not understand, but nonetheless is real because we could quantify it by measurement repetition.	Made unknown contributions explicit.
Comment 1 c: The external diffuser calibration accuracy is quite low. It is mentioned in the text that the diffuser(s) have been calibrated twice at two different institutes, but no results or comparisons are given. I would suggest to add a small section with these external diffuser BSDF calibration results, since this is an important contributor to the BSDF accuracy.	We assume that with 'low' you mean good? The obtained accuracy is in our opinion the best that can be achieved for these kind of measurements with the current technology. The error reported is the combined result of both calibrations.	Clarified in the text.
Comment 1 d: The text refers in some cases to the fact that the preferred sun beam simulator method could not be used for more accurate BSDF calibration accuracy. This seems to suggest that with a well working sun beam simulator the results could have been much more accurate, while the results with the FEL lamps are already now quite accurate. This seems strange.	If the sun beam simulator setup had worked as anticipated, the resulting accuracy of the instrument BSDF calibration would have been much better than it is now. This is due to the fact that with a properly working sun beam BSDF calibration setup all common instrument errors should cancel out, and the main	No rework performed.

	source of error remaining is the external diffuser calibration error. As mentioned the latter error is in the range of 0.5% to 1.0%. Also the SNR would have been much higher than the FEL improving the UV and UVIS. Using the FEL lamps however, more error sources and noise enter the calibration and the resulting error is significantly higher, as listed in the table.	
Comment 1 e: On page 25 the impact of stray light in TROPOMI measurement data is mentioned, which is explained further in sections 6.6 (in-band stray light) and 6.7 (out-of-spectral-range stray light). It is not clear from the text if stray light was (had to be) corrected for the calculation of the radiometric CKD for which the accuracies are given in table 6. It is not clear what additional uncertainty this would add to the accuracies in table 6. Please explain this in more detail in a few sentences and, if necessary, add a line with uncertainty due to all types of stray light. See also table 7 and 8, which suggest that stray light can add up to 1-5% uncertainty before correction and 1-3% after correction, or figure 25, which seems to suggest even higher uncertainties due to stray light in some wavelength areas (e.g. 450-500 nm, and in NIR).	We confirm that both the in-band and out-of-band straylight corrections were applied during the calculation of the radiometric CKDs. This is guaranteed by the use of the production L01b data processor in CKD derivation. The radiometric calibration itself includes the total response of the instrument, and is therefore internally consistent as long as straylight is corrected. The fact that the straylight correction is not perfect, does not affect the conclusion because it is considered true signal and only has to be consistent with the ISRF and PRF.	No rework performed.
Comment 1 f: The accuracies listed in table 6 suggest that with the use of the CKD presented in this paper in-orbit comparisons of TROPOMI L1b measurement data to sun irradiance, earth radiance and earth reflectance spectra should agree within some 1-2% at 1-sigma. Do the authors think that this will be the case? To me that seems unlikely, given the above uncertainties.	The reported uncertainties reflect all error sources in the on-ground calibration as far as we could identify. But they exclude the unknown errors due to known sources. We agree that the reported errors are potentially underestimated. We look forward to the in-orbit comparison for further validation.	We have clarified this at various places in the text.
Comment 2: Section 9: The conclusions are still very qualitative. I think it would be useful to add a few quantitative numbers for some of the key parameters / CKD.	Agreed.	We repeat the important numbers now in the conclusion.

Referee comment	**Author's response**	**Rework performed**
Specific comment 1: Page 15, line 4: In figure 3 the smear correction appears after the dark current correction, while here it seems to be the other way around.	The smear correction in the L01b data processor is implemented after the dark-current correction, which is the correct order. Dark-current is however a special	No further rework.

Please explain, since I understood from the text that the idea is to use the operational 0-1b data processor in the process of deriving CKD.	case because in order to derive the dark-current CKD, the measurements have to corrected for smear first as mentioned in the text. This special condition is handled in the calibration framework during the derivation.	
Specific comment 2: Page 15, line 17: For the detector exposure smear correction the reader is referred to ATBD KNMI (2017) (issue 8.0.0). In that ATBD a rather complex matrix inversion method is described. However, I doubt that this smear correction has been implemented / activated in the 0-1b operational data processing software. Please confirm that this is the case, or else describe (in a few sentences) how the smear is corrected with a simplified approach.	In addition to the rather complex matrix method a simplified algorithm is also described in the ATBD. However it is indeed not explicitly clear that the latter has been selected for implementation in the L01b data processor.	The ATBD is currently being revised, and we will make it more explicit in the next release. No rework for this paper.
Specific comment 3: Page 25, line 20. Please clarify in a few words what this “optical feature” is (since it seems to be quite important, since it causes stray light).	Agreed, the optical feature is now thought to be due to scattering at the inside of one of the mounts that holds the last lens to form the image on the detector.	Added to the text.
Specific comment 4: Page 27, line 15ff: I have some concern with the CKD smoothing procedures. This is normally avoided, to avoid that important instrument spectral features are removed from CKD and then show up in L2 fit residues. What are the main reasons for performing this rather unusual CKD smoothing? Is it diffuser (speckle) features, or also other effects? Are the spectral diffuser features mainly from the external diffuser plate, or also (partly) from the internal diffusers? Please explain this is a bit more detail in the text.	We agree that smoothing of CKD should be avoided. The smoothing here is indeed for the external diffuser features (speckle) that should not enter the instrument calibration.	Clarified in the text.
Specific comment 5: Page 31, lines 2-4: This argument about the FEL lamp cross-hair is a bit strange, since the lamp is calibrated by NIST using the same cross-hair target. I therefore don’t quite understand how this specific effect can lead to uncertainties in the distance. Please explain.	The point here is that the lamp is not used at the same distance as it was calibrated at NIST. To port the calibration to another distance the 1/r2 law is used, but this is only valid for true point-sources. The FEL lamp’s double helicoil is not a point source at the distance used in our setup, and thus an error will be introduced.	We clarified the sentence, and removed the confusing remark about the cross-hair.
Specific comment 6: Page 35, figure 18: It is essential to also provide a plot with the BSDF CKD plotted as function of wavelength, showing also the band overlap regions, in order to see if uncertainties exist in the band overlap spectral ranges. Please add such a plot (this should be easy to produce from the data that is already there).	Agreed.	Figure added.

Referee comment	Author’s response	Rework performed
Technical correction 1: Page 18, figure 7: The horizontal units are electrons. This is understood and agreed. The vertical units seem to be unitless [-], with values up to 15000. This is not understood. Is this really unitless, then how is this to be interpreted? Or should it be electrons after all? Please clarify.	Agreed, the unit is [electrons]	Figure updated.

Response to anonymous Referee #2

Referee comment	Author's response	Rework performed
General Comment: The reviewed Revision 5 is significantly improved over Revision 3. Sections are better organized and balanced, and more background information is provided to contextualize the presentation. While no further rewrite of the paper is needed, I have several comments that I would like addressed prior to publication.	We would like to thank the referee for his/her valuable review.	Added acknowledgement.
Comment 1: The paper contains references to rows and columns throughout. While such terminology is appropriate when discussing detector effects, it is less so for issues such as radiometric response. Some figures, especially but not limited to Fig. 18, would be more informative if plotted versus wavelength.	Many calibrations are indeed detector related. For these no unique wavelength can be assigned due to the row dependent spectral smile. Figures that could be updated are 11, 12, 13, 15 and 18 because they only display the optical axis detector row.	We have updated figures 11, 12, 13, 15 and 18. Data are now plotted as a function of wavelength. Fixed a typo in caption of Fig 18.
Comment 2: The authors are still underplaying the importance of the BSDF calibration. The revised discussion makes clear there were pre-launch calibration problems that prevented a high-quality set of BSDF calibrations. But the instrument still has BSDF calibrations, as presented in Section 6.3. It is noteworthy that this section precedes sections dealing with relative radiance and irradiance, suggesting that these somehow do not affect the BSDF. There are three basic spectral frequency regimes of interest for the BSDF, listed with increasing importance to science products: wavelength-independent errors, broad wavelength dependence, and spectral structure. The authors present a relative BSDF uncertainty for each sensor channel, but do not say over what spectral frequency these values apply. The authors state the BSDFs are spectrally smooth, so the stated uncertainties clearly do not describe residual spectral structure. Fig. 18 suggests (it would benefit from improved quality) the BSDFs are not perfectly smooth, so the authors should provide an uncertainty component for the residual structure. This section or the conclusions would benefit from some critical discussion of the results. By simply looking at Fig. 18 and with some knowledge of aluminum reflectance it is possible to see significant problems in the UV channel. Are the other channels better understood?	We see that in the process of restructuring the section, it may have become unclear that the RELRAD is also part of the instrument BSDF. We have clarified this in the text. We have improved the quality of figure 18, and agree that the claim to spectral smoothness is not fully justifiable. The overall curves of ABSRAD and ABSIRR are similar for each spectrometer, but the shape of these curves differ strongly per spectrometer. These are real instrument properties defined by mirror reflectances, dichroic transmission curves, (graded) coatings and detector coatings. When calculating the BSDF these features do not all cancel out. This could be caused by folding mirror properties or residual speckle on the internal or external diffusers, or artifacts of the optical stimulus. These features are not understood, and no further validation is possible due to the lack of proper	We have updated the figure, clarified the use of RELRAD, and extended the discussion of the BSDF result in line with our response.

	independent on-ground measurements. Therefore, this constitutes an unknown error we cannot quantify and include in the error budget.	
Comment 3: Section 6.6, entitled In-band straylight, contains a nice description of terminology but never tells us what in-band refers to. Is it all stray light that is not out-of-range? Or all stray light that is not out-of-band? The authors (in Section 6.7) seem to use these two terms interchangeably. In other words does in-band stray light ignore inter-band stray light, e.g. from VIS wavelengths into UV wavelengths?	Agreed, text is unclear.	We have made the list on page 41 line 6 more specific and added an additional bullet for clarification. Also added specifically the term in-band where needed in the text (page 41 line 29).
Comment 4: Section 6.6 contains the statement, "The SLRF describes the relative straylight response of the system and is derived from on-ground calibration measurements." Which measurements were used and how were they used?	Agreed, this is not clear. The laser measurements were used to derive the SLRF.	We have clarified the text. (page 42, line 8)
Comment 5: The discussion in Section 6.6 centers on the EWLS measurements, which evidently did not involve spectral cutoff filters. The omission of information about spectral stray light is glaring, regardless of how the requirements were defined. The numbers shown in Table 8 and in Fig. 25 are therefore rather misleading to the average reader. The authors state at the end of the section that laser PSFs form the basis for the processing correction, yet no quantitative assessment is provided. Even if the authors cannot estimate the effectiveness of the correction, they should report the total stray light correction at several wavelengths within each band.	The order in which we have presented the different calibrations may be confusing. We have chosen to discuss the less important characterizations first, and only later in the section we come to the main calibration of near-field straylight which is done with the laser stimulus. The EWLS is used for various characterizations and also to create validation scenes; cut-off filters have also been used as described on page 41 line 26. At the end of the section we come to the in-band-near-field straylight which fully covers all *spectral* straylight (and also a spatial component). These are corrected simultaneously by a 2D convolution method. This accuracy of this method is assessed quantitatively as presented in table 8.	We have added a few lines to the text to make this more explicit. Fixed double reference to a single figure on page 44 line 15.
Comment 6: Table 8 needs more explanation, either in the caption or in the text. What is the definition of stray light percent (what is the value in the denominator)? Same comment applies to Fig. 25.	For table 8 the definition is the total integrated (straylight)signal outside the direct region divided by the integrated (EWLS) signal inside the direct region. For the figure 25 the definition is similar to the system requirement for a hole-in-the-cloud scene: the straylight at the center of the hole divided by the EWLS signal outside the hole.	Updated caption table 8 and figure 25.

Referee comment	Author's response	Rework performed
Technical Comment 1: Figure 17 appears to be mislabeled. Different lines should be columns instead of rows.	Agreed, caption is correct, but labels in figure are wrong.	Figure updated.
Technical Comment 2: Section 6.6, paragraph 3. Tol et al. (2018) should be followed by a period or semi-colon	Agreed.	Added semicolon.
Technical Comment 3: Figure 23 caption. The fifth sentence is poorly formed.	Agreed.	Caption updated.