1 *Author comment on* "New method to determine the instrument spectral response function, applied to TROPOMI-SWIR" *by Richard M. van Hees et al.*, manuscript amt-2017-438, Anonymous Referee #1

We would like to thank Referee #1 for the very useful comments to improve our manuscript. In this document, we provide our reply to the comments. The original comments made by the referee are numbered and typeset in red. Page, line and figure numbers refer to the old version of the manuscript. After the reply we provide a revised version of the manuscript, each section is adjusted according to the review as described below. In this process, the original text has been substantially rewritten and/or reorganized. The output of 'latexdiff' is considered confusing and nevertheless provided. The revised version of the manuscript is part of this authors comment.

A brief overview of the major changes of the revised manuscript compared to the original manuscript:

- Title has been changed to "Determination of the TROPOMI-SWIR instrument spectral response function". We do no longer claim to present a new method.
- Clear distinction is made between what has been measured, what is determined from these measurements, and what is delivered. Inconsistent usage led to confusion.
- Thanks to the reviewer question to explain "laser artefacts", we discovered that:
  - a significant amount of irradiance measurements had to be rejected from the analysis due to an unexpectedly large variation of the laser signal yielding detector saturation.
  - we rejected radiance measurements from partly illuminated rows.
- The Pearson VII shape parameter m was fixed to 1.25 in our analysis presented in the original manuscript, because then the shape of the model matched the far wings observed in TROPOMI SWIR stray light very well. However, simulations – requested by the reviewer – showed that this assumption introduces significant errors in case the true ISRF has wings with a different shape. A new approach has been developed and verified with simulations that requires two successive ISRF fits per stage: first only fixing the tail fraction (from the previous stage), then only fixing w (improved guess from the previous ISRF fit). The method requires an initial guess of the tail fraction, which can be obtained from stray light or ISRF data (determined in stage 1).
- As a result of the above 2 points the analysis was improved such that the irradiance ISRF and radiance ISRF are more consistent.
- Analysis of the ISRF as determined from on-ground measurements with the on-board diode-lasers is included in the manuscript.

1.1 General comments

The paper addresses the determination of Instrument Spectral Response Functions (ISRF) of the recently launched Tropomi/Sentinel-5P mission. ISRF uncertainty is a notorious limitation of past and future space-borne atmospheric chemistry missions (e.g. GOME-2, OMI, Sentinel-4, Sentinel-5), as well as missions targeting greenhouses gases (OCO-2, MicroCarb, Sentinel mission). The SWIR band of the Tropomi/Sentinel-5P is used for retrieval of CO and CH4, where ISRF knowledge is most critical (together with the NIR band) because of strong, narrow absorption features. In this sense, the paper addresses a critical aspect of trace and greenhouse gas retrieval and therefore the topic is of high scientific relevance.

GC-i The introduction of a new method for ISRF characterisation by on-ground calibration measurements, as promised by the title, would be of high interest for planned future missions. However, the manuscript fails to deliver key elements for introducing and making a case for a new method. Neither do the authors describe other techniques, nor do they motivate the introduction of a new one, explain the difference, or compare it with previous calibration/validation measurements. If the objective of the paper is the introduction of novel method for ISRF determination, as suggested

by the title, the differences to previous instrument calibrations (e.g. SCIAMACHY, OMI, OCO-2) have to be pointed out. Ideally, a comparison shall be offered pointing to the advantages of the new approach.

Adjusted. We do no longer claim to present a new method, updated the title, abstract, section "Methodology" and conclusions accordingly. Citations have been added where relevant. A brief discussion on ISRF determination of several relevant other missions is added to the Introduction.

GC-ii In fact the reader is left wondering which are the new elements of the proposed approach: On the instrumental aspect: Is it the first time an OPO was used? If so, why is it expected to yield better performance (than e.g. monochromator)? What are the instrumental limitations of the approach? The impact of key parameters of the measurement data (e.g. signal-to-noise ratio) is not discussed. No details on the instrumental setup are provided, and the quoted reference does not contain them either. On the modelling aspect: Is the novelty of the approach the mathematical model for the ISRF in terms of a peak and a tail function? Then, why were these particular functions chosen and not any other? Are there physical reasons why the ISRF tails should follow a Pearson type VII distribution? All these questions are not addressed in the manuscript. In large parts, the text resembles an technical document (or ATBD), which describes a mathematical algorithm without explaining, why certain steps are taken.

Adjusted. The manuscript presents the determination of the TROPOMI-SWIR ISRF. All questions and remarks are addressed in the updated manuscript and in the detailed comments below.

GC-iii The fitting procedure to determine the ISRF seems to suffer from an under-determined equation system, although this is never mentioned. This is mitigated by "fixing" some parameters selectively in the inversion during four iterative steps. Again, the authors do not justify the presented sequence of partial fits, other than it results in "good fits" (defined by low residuals). This is particularly worrying as the approach is verified with two synthetic ISRF shapes, which were presumably computed using the same mathematical model as in the inversion (albeit this is not clear from the text). In theory, if the information content of the measurements would be sufficient to estimate all model parameters, this should result in perfect agreement between fit and forward model. However, the fit residuals clearly exhibit systematic features, indicating a weakness in the fit procedure. Nevertheless, the authors conclude "compliance" as the fit residual's amplitude is below the accuracy threshold (1% requirement).

Adjusted. We obviously have failed to present the method to determine the ISRF clearly, leaving the reader confused. Therefore, we have reworked our manuscript to cover these questions. The comment and questions are repeated as detailed comments below.

GC-iv The iterative fit procedure yields ISRF shapes for all detector pixels, which are not uniform and clearly show high-frequency variation. At this point the authors argue that the ISRF is only determined by the spectrometer optics (PSF), which varies smoothly with wavelength and swath angle. This assumption is not in line with the definition of ISRF used in other publications, and even with the introduction of the present manuscript. Defined as the spectral response of a single pixel as a function of wavelength, the ISRF includes the detector PSF (convolved with the slit boxcar and the optical PSF). The detector PSF is mainly driven by cross-talk, which may indeed introduce a pixel-to-pixel or column-to-column variation. While such systematic features are clearly visible in the estimated parameters across the detector (e.g. Fig. 5), the authors indirectly dismiss them as artefacts (by the above assumption) and eliminate them with an elaborate "smoothing procedure".

Partly adjusted. We only partly agree with the reviewer. By assuming that the ISRF is determined by spectrometer optics (PSF) –which should varies smoothly with wavelength and swath angle by design– we can identify detector PSF features. This is illustrated in Fig. 5b, where clear features of possible SWIR detector PSF are visible in the residuals (Fig. 8b in de revised manuscript). However, these are not the obvious column-to-column effects (the vertical stripes), present in Fig. 5, which are measurement artefacts, because these are absent (or at different locations) in other measurements (radiance and/or diode laser measurements). Variation of the detector PSF across the array is reflected in the light blue areas in the dark blue background on the right side of the detector. The same patterns are seen in the photo-sensitivity of the SWIR detector (Fig. 6 of Hoogeveen et al., 2013). An estimate of the error introduced by the

bivariate polynomial fit, which neglect detector PSF, is preseted in Fig. 15 in the revised manuscript. However, it can be shown that this error is small, see Sect. 4 "Discussion of results".

GC-v The smoothing equations (bi-variate polynomial fit) are reported, without justifying the choice of parameters (merely stating to yield "smoother and better" results). It is impossible to judge the impact (and therefore significance) of the smoothing, since no statistics are provided as of how much the individual parameter have been smoothed. The difference w.r.t the original fit results is not presented. Pixel-dependent effects (e.f. cross-talk) are "smoothed out" by this procedure, but presence of such effects does not mean that the individual ISRFs determined in the previous step were less accurate.

Adjusted. We agree with the reviewer that our description (in Sect. 3.5) is confusing and have updated the manuscript (Sect. 3.3) accordingly: (i) the need to smooth and interpolate the ISRF parameters is explained in the text. Figures on the difference between parameters of the local ISRF and the smoothed ISRF are added; (ii) the order of the bi-variate polynomial fit used, is determined by the variance of the residuals (= ISRF parameters - ISRF smoothed parameters). The lowest order is used for which the variance stabilizes, because then we can identify contribution by the detector PSF; (iii) we are aware that small local variations are neglected/ignored, any remaining residuals are checked by examining the results.

GC-vi The use of smoothing procedures gives the impression that the authors do not have much confidence in their technique. They regard the smoothed ISRF parameters as more accurate, and report the observation that the individual rms deviation from the measurements have increased as "counter-intuitive" (although seems to be expected).

Adjusted. It is certainly not our intention to leave the reader with this impression. This section has been rewritten, see also DC-36.

GC-vii It is understood that for practical purposes (operational Level-2 processing), calibration key data cannot take into account ISRF variations at the pixel-to-pixel level. However, the authors should clearly state where such compromises are made(e.g. between level of detail and computational resources) and justify them by quantifying or estimating the impact on Level-2 accuracy. Currently, the ISRF is defined such as to match the fit and smoothing procedures, effectively establishing a definition which is only valid for the presented approach. This compromises the ability to compare with other instrument calibrations and the applicability for future instruments.

Adjusted. We are confident that the delivered ISRF calibration key data are of high quality, based on the residuals seen and the simulations done with synthetic data. We have found no evidence for residual pixel-to-pixel effects, except for a correlation with photo-sensitivity of the SWIR detector, see discussion in Sect. 4. These residual pixel-to-pixel effects are small and can be ignored. Measurement artefacts have been identified by a comparison of the irradiance, radiance and diode-laser measurements. The improved results wrt the original manuscript, and the many more simulations done further underline the original conclusion.

GC-viii Any local variation is interpreted as resulting from not further specified laser artefacts, which are smoothed out by a bi-variant polynomial fitting. This needs to be further justified, answering the following questions:

- Are the laser artefacts repeatable? Have measurements been repeated for some of these "bad data"?
- Do they occur at given angles and wavelengths (patterns) or are they randomly distributed?
- What is the likely instrumental root cause (e.g. speckle or wavelength instability)?
- Why are such instrumental effects absent in the "good data"? Could these also be affected by "laser artefacts".

The paper lacks a discussion of error sources of the new method including instrument effects (only fit residuals are considered). A true error analysis of the technique would involve a rigorous analysis of instrumental error sources, such as: - SNR of the laser measurements; - laser stability; - speckle amplitude (of integration spheres, diffusers); - straylight correction efficiency; - non-linearity (resp. knowledge thereof); - pixel-to-pixel cross talk variation.

Adjusted in the revised paper. Answers to the questions listed in this comment: (i) Laser artifacts are explained, measurements affected by saturation due to too strong laser signals are rejected from analysis. Detailed data analysis was performed after the measurement campaign was finalized, denying the possibility to improve or redo certain measurements. However, measurement artefacts have been identified by a comparison of the irradiance, radiance and diode-laser measurements; (ii) Mostly random, hoewever, at some wavelength ranges the OPO was better behaved, compared to other ranges. Then measurement artefacts are present in the results of both radiance and irradiance measurements at nearly the same wavelength ranges; (iii) There is no evidence that speckle is present in the radiance and irradiance measurements. In the original manuscript, we use "laser artefacts" for too strong laser signals, wavelength instabilities and intensity instabilities; (iv) Yes, but only minor effects are seen in ISRF fits with a small rms of less than 0.003.

All detector calibrations, such as memory, offset, dark current, background signal, PRNU are corrected using the operational level 1b processor. Examples of "good data" affected by wavelength instabilities are presented in the revised paper.

GC-ix The methodology for evaluating the suitability of the approach is questionable. The only figure-of-merit considered is "good fit quality". The amplitude of fit residuals is interpreted as the accuracy of the method, compared with the requirement of 1% of the ISRF peak. However, fit residuals only provide information about the consistency between the mathematical model and the measurement. If the measurement is affected by systematic instrument error (say, from stray light), the fit absorbs this error into the estimated parameters. This does, however, not mean that the true ISRF was determined more accurately. In fact,the authors identify "different treatment of stray light" as the likely cause of discrepancies between the ISRF for radiance and irradiance ports (although the correction technique and its accuracy is not presented), which in theory should be identical. But instead of interpreting this as a limitation, the two different stray light backgrounds are fitted by the model and both ISRFs are regarded as true ones. It is clear that instrument effects are fundamentally unavoidable (and obviously present here). They should be identified as such, and not "absorbed" into parameters of the model and declared part of the "true" ISRF.

Adjusted. In the revised paper, we present the rms values, not a flag for good fit quality.

Stray light is corrected with the operational level 1b processor. In irradiance measurements, where the light is imaged as a vertical line, the ISRF at one row could be affected by stray light from other rows. Hence, these data are corrected for stray light. In radiance measurements, where the light is imaged as a spot, there is no stray light from other rows and the stray-light correction is not applied. This was intended with the remark "different treatment of stray light". The fact that radiance measurements are not corrected for stray light is mentioned in Sect. 3.2, page 4, line 21.

GC-x Due to a poor (not further explained) laser performance, the irradiance measurements were used to compute key data, while the radiance fits classified as "good" were used for validation. Some validation results are reported, but only median values for selected columns, from which the conclusion is drawn, that radiance and irradiance ISRF are identical. This raises the question, if ISRF characterisation (which is typically a cost driver for imaging spectrometers), can be restricted to irradiance measurements only. This would greatly reduce effort (one wavelength scan versus 100 fpr each spatial sample) and cost. A discussion and quantitative analysis would greatly enhance the impact of the paper.

Adjusted. This comment has been addressed in the revised paper. Laser artefacts are explained and examples provided. Furthermore, Local ISRF data and fits are presented, not medians.

For SWIR, the spectral features introduced by the on-board diffuser in irradiance data have much larger periods than the ISRF range, so they are constant per ISRF. Because they are absent in radiance data, the radiance measurements were intended to be used for the key data, but the irradiance measurements turned out to be better suited. The difference found between the irradiance and radiance ISRF is gone with our improved method; it was an artefact of the fixed tail parameter m.

To characterize the ISRF only with irradiance data indeed saves a lot of measurement time, but denies the possibility for a proper comparison between the two measurements. For example, confirmation on presence of detector effects.

GC-xi The paper includes a short section on in-flight calibration of the ISRF with Tropomi's on-board calibration system, comprising five tuneable laser diodes. This part of the paper has the potential to significantly raise its impact, as this aspect is relevant for several upcoming missions (see list above). However, this opportunity is missed as no comparison between on-board and external diffuser/laser ISRFs is provided. The authors merely the state, that "The ISRF measured with the diode lasers is in close agreement with the ISRF calibration data,...", without presenting any plot, table, or statistics. Since data using the on-board calibration system were acquired and are available, it is strongly suggested to extend this part of the paper by providing quantitative comparison.

Adjusted. The on-ground measurements with the on-board diode lasers are included in the revised paper. However, the method for in-flight monitoring the ISRF stability will be presented in a separate paper after in-flight data has become available and has been thoroughly analysed.

GC-xii Finally, the authors do not give adequate credit to previous work. The reference list is rather short and limited to Tropomi-related publications. Tropomi/S5P not the first grating spectrometer for which extensive ISRF calibration has been performed, and also not the first covering the SWIR spectral range (e.g. SCIAMACHY, OCO-2). A more extensive discussion on previous work(actually needed for the introduction of a new method) should include a literature review covering the following missions (non-exhaustive list): -SCIAMACHY -GOME-2 -OMI (Dobber et al. 2004) -OCO-2 (Crisp et al., lasers used for ISRF calibration in SWIR and NIR) They should also mention upcoming missions for which a new technique may be come relevant, like e.g. -Sentinel-4 -Sentinel-5 -FLEX -MicroCarb -future Copernicus mission for antropogenic CO2 Also, the definition of ISRF and ISSF functions in the introduction, although accurate, needs to make reference to previous discussions in the literature (see detailed comments).

Adjusted. Thank you, for providing literature suggestions. A better overview of the field has been added, together with the suggested citations, where relevant.

GC-xiii Overall, the draft paper in its current status cannot be considered for publication in a peer-reviewed scientific journal. However, as the topic under study is highly relevant for interpreting results from Tropomi/S5P, as well as for a wealth of future space-borne push-broom imaging spectrometers, they are encouraged to thoroughly rework the manuscript along the lines indicated in the above comments, and the more detailed ones below.

We agree with this conclusion and thank the referee for the general and detailed comments. The manuscript has been revised along the suggested lines.

Detailed comments

DC-1 Abstract: from the abstract it is not clear what is the difference be-tween "deriving" an ISRF, as proposed in this paper and applied to the SWIR band of the TROPOMI instrument on one hand, and "measuring an ISRF", which is reported to be done for all bands. Please rephrase or add a sentence, clarifying the topic of the work presented here. It is not clear from the abstract (and in large parts of the paper), if a new generic technique for determining an ISRF is proposed, or if the it is simply reported how it has been done for the Tropomi instrument. If the main topic is the introduction of a new technique for ISRF derivation, which can be used for future instruments, then the abstract shall contain quantitative statements about the advantages of the new technique. Will the proposed new method or be used to derive the in-flight ISRF?

Adjusted. We rephrased the text in the manuscript (not only the abstract) to clearly distinguish what has been measured, what is determined from these measurements, and what is delivered. Inconsistent usage led to confusion.

No, the method is TROPOMI-SWIR specific. The abstract has been updated accordingly.

DC-2 P.1 L.13: incomplete sentence "The latter..."

Adjusted according to the suggestion.

DC-3 P.1 L.14: "entrance slit"; also add that light from ground scene is collected by a common telescope (not mentioned)

Adjusted according to the suggestion.

DC-4 P1 L.19: Why does CO "have to be measured" with this acc. ? If this is a scientific requirement, please clarify and reference it (e.g. from a study)

Adjusted. Replaced text by: "The TROPOMI-SWIR band is used for the retrieval of the atmospheric trace gases carbon-monoxide, methane and water vapor. Simulations have shown that in particular the methane retrieval is very sensitive to errors in the instrument spectral response function (ISRF or instrument line shape). As a result, the requirement on the ISRF is formulated that it should be known with an accuracy of 1% of its maximum where the ISRF is greater than 1% of its maximum (Hu et al., 2016)."

(a) Add a sentence clarifying the term ISRF (in US literature instrument line shape) and how it is defined

Adjusted. The term ISRF is defined in the introduction, here we also refer to the term "instrument line shape".

(b) Better clarify the link between the requirement for XCH4 accuracy (1%) and the ISRF requirement. Is this error the only contributor to the budget?

Not adjusted. This is outside the scope of this manuscript. Please consult Hu et al. (2016).

DC-5 P.2 L. 2-17: The discussion of the difference between ISRF and ISSF on is important. It is basically paraphrasing a similar discussion in: J. Caron, et al.: THE CARBONSAT CANDIDATE MISSION: RADIOMETRIC AND SPECTRAL PERFORMANCES OVER SPATIALLY HETEROGENEOUS SCENES ICSO 2014 Tenerife, CanaryIslands, Spain International Conference on Space Optics, 7–10 October 2014. Please cite this paper in the introduction

Not adjusted. The CarboSat paper focuses on the distortion of the ISRF in spatially heterogeneous scenes, an important subject, but it does not provide a clear discussion on the ISSF and ISRF.

DC-6 P.2 L.5: It shall be pointed out what is the role of the ISSF, which is composed of ISRFs of neighbouring detector pixels. Is only the ISRF useful for ...

Adjusted. The sentence is indeed confusing: we have measured many ISSF samples, to determine high resolution ISRF. Rephrased in the revised manuscript.

DC-7 P.2 L.27-28: Previously it has not been mentioned that the instrument also measures solar irradiance. Please clarify that in the introduction, so that the reader understands the difference in "irradiance ISRF" and "radiance ISRF".

Adjusted. Added the following text to the introduction, page 1, lines 14–16: "The TROPOMI instrument measures sunlight reflected by the surface and atmosphere of the Earth via the radiance port. Direct sunlight is measured via the irradiance port and internal diffuser for calibration purposes."

DC-8 P.2 L.9: "In the spectral dimension, about 4–5 points have significant signal. This is the spread function of the instrument for this wavelength". If "This" refers to 4–5 points of signal, it must be replaced by "these"? It shall be added, that an ISSF cannot be measured continuously, but only sampled, while the ISRF can be measured continuously.

Partly adjusted. Added to the following text, page 2, line 18: "In the spectral dimension, about 5 pixels have significant signal as expected from the spectral oversampling." It is already explained in the text and Figure 1 that "an ISSF cannot be measured continuously, but only sampled, while the ISRF can be measured continuously."

DC-9 P.2 L.9: Recommendation: Cite a publication about the SCIAMACHY instrument, which pioneered space based measurements in the SWIR spectral range. Also cite the NASA mission OCO-2, which also deployed tuneable lasers in ISRF calibration.

Adjusted. Differences and advantages with previous analysis of instrument ISRF have been added to the introduction, page 2 lines 6–13.

DC-10 **GENERAL REMARK:** The introduction(and abstract)fail to motivate a new in the technique for ISRF determination, as is promised by the title. Why is a new approach needed ? Is the 1% accuracy requirement not reachable by previous techniques? What is the basic idea of the new approach? Please add a section motivating the new approach and its basic idea (and why it is only applied to the SWIR).

Adjusted. The usage of measure, determine and derive have been reconsidered and we have aimed to used them consistently throughout the document. For example, we measure an ISSF, from which we determine the wavelength and signal intensity of the laser for this measurement.

We dropped the claim to present a new technique, because we only adopted existing techniques to our measurements and 1% accuracy requirement. Updated the Title, Abstract, Introduction and Conclusions accordingly.

DC-11 P.2 L.31: Calibration measurements: Please provide a figure depicting the calibration setup (incl. OPO, integrating sphere, collimating optics, etc.). This would make it easier to follow the given description

Adjusted according to the suggestion.

DC-12 P.3 L.8: 1.1 deg. of the FoV corresponds indicates that more then one detector pixel is illuminated (216 pixels/(108 deg/1.1 deg) = 2.2 pixels). In the introduction it is stated that the ISRF is determined for individual pixels. How does image distortion (smile, frown) influence the ISRF measurements over >1 pixel?

Not adjusted. The image distortion is subpixel in both spatial and spectral direction and can therefore be neglected locally.

DC-13 P.3 L.11-12: The sequence of measurements is not clear. With 165 sec. of measurements at 10Hz we have 1650 acquisitions, so 20 for each wavelength?

Adjusted rephrased the sentence, see answer to next comment

What is the spectral sampling of the calibration measurement (2nm/80 = 0.025 nm, equally spaced)?

Adjusted as explained in Sect. 2.1, see answer to next comment

Is the wavelength adjusted in steps or continuously(by temperature variation?). Please clarify in the text.

Adjusted. Each automated wavelength scan of about 2 nm, or 20 spectral pixels, took about 165 seconds. The data is collected at 10 Hz, yielding 1650 ISSF samples in total and about 80 samples per pixel. The laser scan speed was not constant, despite the small adjustments of the piezo voltage during the wavelength scan. Due to the large number of samples taken, this has no negative impact on the ISRF determination.

DC-14 P.3 L.14: Please explain in the text why the "irradiance ISRF" is expected to be different from the "radiance ISRF" of the same spectrometer. Theoretically, the ISRF is a spectrometer property determined by the optics after the entrance slit. The only difference is the pre-slit illumination via the Sun diffuser (which btw. is not mentioned in the instrument description given in the introduction). The cited paper (Tol et al.2017) does not provide a sketch neither.

Adjusted. rephrased text on page 2, lines 30–32: "The ISRF is determined from measurements using the radiance port and the irradiance port. Although differences between both data sets are not expected, they could not be ruled out. Light entering via the irradiance port follows a different path, via a diffuser."

DC-15 P.3 L.14: "Calibration measurements via the radiance port and irradiance port have been performed to verify that they are identical." Were the irradiance and radiance ISRF measurements identical? To what extent? If this is reported later, indicate it in the text.

Adjusted. The explanation is moved to the introduction.

DC-16 P.3 L.16: The wording "100 wavelength scans per manual wavelength setting... to cover the swath..." is somewhat confusing. Is the manual setting a coarse adjustment which is then scanned in finer steps? Or is it that 100 swath positions are manually aligned and then scanned over wavelength?

Adjusted. During the radiance measurements each automated wavelength scan (of 2 nm) had to be repeated for each swath position. At each swath position the wavelength was scanned in the opposite direction to save time.

DC-17 P.3 L.18: "...was performed up and down" -> better "...was performed with increasing and decreasing wavelength scan direction."

Adjusted according to the suggestion.

DC-18 P.3 L.25: "block distribution" is not a commonly used expression, recommended to use the term "boxcar function" instead.

Adjusted. We do no longer refer to the block distribution. However, in our opinion, the term 'boxcar function' is associated with time series and is not a distribution. Therefore, we prefer the term 'uniform distribution'.

DC-19 P.3 L.28: "In the end, the Pearson type VII distribution resulted in the best fit". What criteria determines what is the best fit? In general, the criteria for the quality of the determined ISRF are not clearly defined at this stage. Define or reference "Pearson type VII distribution"

Adjusted. A justification why the Pearson VII distribution was used is added to the section "Methodology". Several other possible distributions have been tested, including the super-gauss, an asymmetric version of the exponential power distribution (Beirle et al. 2017). They do not represent the TROPOMI-SWIR ISRF with enough precision. The asymmetry of the super-gauss is provided by combining separate functions for two halves of the peak, which results in a continuous curve, but a discontinuous derivative, leading to larger residuals in the peak. The tails of the super-gauss are far too steep. Pearson type VII is defined as Eq. 4 in the text.

DC-20 Please clarify why the particular representation of the ISRF was chosen. The idea of the Peak function is sketched (convolution of a perfect slit image with a function representing image blur by the optics. However, a Gaussian is not a perfect representation of the spectrometer PSF and the detector cross talk, and other functions maybe used. Please comment on how the shape components were chosen (e.g. why a Pearson distribution). How robust are the results w.r.t. other representations?

Adjusted. this comment is elaborated in Sect. 3.1.

It is also not shown (nor stated) that eq. (2) and (3) represent a convolution of a skew-normal distribution with a boxcar function (which is suggested on P.4 L.1-2). Please clarify.

Adjusted. this comment is elaborated in Sect. 3.1.

DC-21 P.4 L.21: "The measurement data are corrected for... and straylight (irradiance only)". Why are the irradiance data stray-light corrected, and radiance data (apparently) not? Generally, the stray-light correction raises the question of definition of straylight, which is linked to the definition of the ISRF. In some ESA and NASA missions, stray light is defined by the spectral extent of the ISRF. E.g. for OCO-2, spectral stray-light is defined as the light detected at wavelengths beyond 6 times of the FWHM of the ISRF. Inside this range it is part of the ISRF. The authors should highlight the relation between stray light and ISRF definition, and explain the definition used for Tropomi.

Adjusted. Added to the introduction: "For Tropomi, spectral stray-light is defined as the light detected at wavelengths beyond the equivalent of 4.5 pixels. Inside this range it is part of the ISRF." From Sect. 2.3: "In irradiance measurements, where the light is imaged as a vertical line, the ISRF at one row could be affected by stray light from other rows. Hence, these data are corrected for stray light using the operational processor. In radiance measurements, where the light is imaged as a spot, there is no stray light from other rows and the stray-light correction is not applied."

DC-22 P.4 L.21: Have the data been corrected for detector non-linearity as well? This could be important for measurement of the far wings of the ISRF, where signal levels are low.

Adjusted. From Sect. 2.3: A non-linearity correction is not implemented in the operational processor. It is also not needed for the ISRF characterization, as the error is small: detector non-linearity was measured to be about 0.1% to 0.2% (Hoogeveen et al., 2013).

DC-23 P.4 L.22: "Readouts from bad pixels are discarded in the analysis." The manuscript often uses terms from requirement definitions, which may not be familiar to all readers. Please briefly clarify the term "bad pixels" (e.g. reduced spectral detection efficiency).

Adjusted. A description of bad/dead pixels is added tot Sect. 2.3.

DC-24 P.4; L.23-24: "Frames where the light source was off or very weak are discarded." Switched off light source is not expected in nominal calibration procedure. Specify how many data were lost or, if insignificant, remove sentence.

Adjusted. Indeed confusing, we have removed sentence.

DC-25 P.4; L.24-25: "In each remaining frame, the column with the maximum average signal is determined and the columns up to 7 pixels from this peak column are selected, to include the faint signal of the tails." Are 7 pixels sufficient to capture the faint signal of the tails? The ISRF could be deliberately saturated to increase the sensitivity further out to the wings. The criteria should be the definition of the ISRF boundary (1% of the peak). Please comment (and add in text) that 7 pixels cover that range.

Adjusted. Too detailed information, not relevant. This paragraph has been removed.

DC-26 P.5 L.2–18: This short (5 lines) section should be merged with the next one, or even the next two, which can be combined in a Section "ISSF and ISRF fitting".

Adjusted according to the suggestion.

DC-27 P.5 L.4: "The ISSF is assumed to be the mirrored version of an ISRF,..." With growing wavelength distance (and changing optical PFS), the measured ISSF should differ from a mirrored ISRF centered at c0. Deviations from the symmetry assumptions could potentially affect the far wings of the ISRF. Please comment on the assumption and justify be demonstrating negligible error. If possible, present data (plots) supporting the validity of the assumption.

Adjusted. Following text has been added to the introduction: "By design, the ISRF and dispersion of the TROPOMI instrument should vary only smoothly in the spectral and spatial dimension. This assumption will be validated in this study by determining the local ISRF for many pixels of the SWIR detector."

DC-28 P.5 L.4: "which can be modelled with the function AR(c;d,-s,w,.,.,m,c0) using Eq. (1); only its skew parameter s has the opposite sign". The introduction of the function AR in this sentence is confusing. Just define it by R with reversed skew parameter to reflect the mirror shape w.r.t. the ISRF. Maybe even add the defining equation AR(s) = R(-s).

Adjusted. Intended is "aR() where the normalized function "R()" is defined in Eq.1 and a is a scaling parameter. Text in the manuscript is updated accordingly.

DC-29 P.5 L.5: "In each frame, the ISSF of an illuminated row is fitted to the ISRF shape to normalise the signals and to find the wavelength peak position expressed in pixel units (Fig. 1b)". The process of fitting the ISSF to the ISRF shape involves the estimation of the 8 parameters indicated in Eq. 1-4. In a previous sentence, the number of useful pixels of a measured ISSF frame is only about 5: P.2 L.9: "In the spectral dimension, about 4–5 points have significant signal". Therefore the inversion of an ISRF profile seems under determined. Please explain how the ISRF parameters can be still be estimated (assumptions on constant parameters?).

Not adjusted. Rephrased this paragraph, the ISSF is fitted with at most 4 free parameters (stage 1) and in later stages with only two parameters: wavelength and intensity of the signal. At most 6 parameters are fitted in the ISRF fit.

DC-30 P.5 L.18: "The square root of the fit variance is the rms value." This can be assumed to be known by the reader (recommended to be removed).

Adjusted. A definition and explanation of "rms" as used in the revised paper: "An alternative measure of the fit quality turns out to be an rms value calculated as the square root of the sum of the squared difference between the ISRF fit residuals, using points where the fit function is larger than 6% of its maximum, divided by the number of ISRF data points minus the number of free fit parameters. The threshold of 6% is arbitrary, but a lower value would include more

of the tails where the residuals are always very small, which would make this measure less sensitive. The advantage of this measure is that it is less sensitive to small outliers, and sensitive to large outliers which can corrupt the fit procedure. Therefore, we use this rms as a measure of the fit quality."

DC-31 In general, the description of the fit procedure is rather sparse. It reports processing steps, which are not further justified, leaving the reader wondering why a certain step is taken in a particular way: For example: 1) "As the laser-wavelength scan is not regular, the ISRF data points are not on a regular grid. Therefore, the points in the scan range are collected in bins of 1/32 of a spectral pixel and a median is applied to the data points in each bin. Empty bins are discarded". Does the sampling (1/32) automatically follow from the non-regularity of the grid? How irregular are the wavelength steps of the laser? Why can not the fit be performed on a non-regular grid, when the functional shape and the relative wavelength are prescribed anyway?

Adjusted. We removed irrelevant information from the description of the method.

2) The quality of the fit is determined by calculating the fit variance, the sum of the squared fit residuals where the fit function is larger than 6% of the maximum, divided by the number of degrees of freedom (number of points minus the free fit parameters)". Please explain (or reference) the fit quality parameter. In absence of explanation it appears to be an arbitrary choice (e.g.why 6% threshold). Since a new method shall be introduced here (according to the title), please extend the discussion of the fit procedure, explaining and justifying all steps.

Adjusted. The description is extended in item DC-30.

DC-32 P.5 L.21: "However, the ISRF fits are valid locally (at location (r, c)) and not available for all pixels". A previous sentence seems to indicate that the fit is done for all pixels: P.2 L.24: "A description of the method and algorithm used to derive the ISRF for all illuminated pixels is presented in Sect. 3" Please clarify.

Adjusted. It always has been our intention to measure the ISRF for all pixels, however, we could not determine the ISRF for all pixels due to measurements with too strong laser signal, bad pixels and laser artefact. Therefore, it is necessary to interpolate the local ISRFs.

DC-33 P.5; L.21-22: "It is expected that the fit parameters that define the local ISRF vary only smoothly over the surface of the detector as this is determined by the spectrometer optics". This is only true if detector effects can be neglected. This assumes that all pixel-to-pixel effects (like PRNU, DRNU) are perfectly calibrated. Cross-talk effects, which determine the detector PSF (one component of the ISRF and therefore not eliminated by calibration), may vary from pixel-to-pixel. Please justify the assumption that the ISRF is determined by optics only. If possible, provide empiric evidence for constant cross-talk.

Adjusted. All detector calibrations are performed with the operational level 1b processor. We have found no evidence for residual pixel-to-pixel effects, except for a correlation with photo-sensitivity of the SWIR detector, see discussion in Sec. 4 of the revised manuscript.

DC-34 P.5 L.25, Eq. 5: Similar to the previously described fit procedure, this particular smoothing function seems to "fall from the sky". Please provide justification for this particular function and how its parameters are determined (e.g. are the constants 255 and 999 the detector pixels in spatial and spectral dimension?

Adjusted. The function is given in a simpler form (but it still represents exactly the same shape). We have added some explanation and mention that 255 and 999 are the detector pixels. Order 'M' is the maximum of the sum of the exponents of variables 'r' and 'c'.

DC-35 P.6 L.1-3: "To obtain good results for the ISRF parameter fitting, obvious outliers in the individual ISRF-fit results should be rejected before the bi-variate polynomial fit is performed. Given the distribution of outliers (in columns at the same wavelength), it is judged that most of them are caused by laser artefacts". Typo: Replace "artifacts" by "artefacts"

Adjusted. Thanks to the reviewers question to explain "laser artefacts", we discovered that:

- a significant amount of irradiance measurements had to be rejected from the analysis due to an unexpectedly large variation of the laser signal yielding detector saturation.
- we rejected radiance measurements from partly illuminated rows.

Note that "artifact" is not a typo, depends on the language (US, British).

DC-36 Comment: Each ISRF has been determined by a multitude of ISSF measurements at many wavelengths, which have been used to fit a composite shape function with relatively few parameters. Random laser effects (which ones?) should already be smoothed out by this procedure. It is somewhat surprising, that these obtained parameters need to be further smoothing. Please elaborate (in the text) on: -the size and distribution of outliers -why it is judged that the outliers are caused by laser artefacts. The likely cause of "laser artefacts" (e.g. can better lasers improve the method?)

Adjusted. Smoothing is needed to derive ISRF for every pixel. Minor problems with the OPO are still present in the ISRF data points (see Figs. 3–5), which may affect the ISRF fit yielding minor deviations of the fit parameters. Assuming that these deviations are random and that the shape of the ISRF varies only smoothly over the surface of the detector (as it is determined by the spectrometer optics), then the quality of the determined ISRF would benefit when the ISRF parameters are smoothed and interpolated using bivariate polynomial fitting.

Measurement artifacts are identified by obvious outliers: $s > 5$, nearby saturation (measurements of a wavelength scan where saturation occurred), comparison with other measurements e.g. radiance and/or diode laser measurements.

DC-37 P.6; L.8: "unrealistic curve-fit solutions are rejected..." Please report the fraction of rejected fits.

Adjusted. We provide these numbers in Sect. 4

DC-38 P.6; L.9-10: Please report the fraction of rejected fits. And again, please justify the numbers of the rejection filter (why rms <= 0.0065 and not any other number?) Question: What was the impact of "bad" and "dead" pixels in the procedure? This may provide important guidelines for detector cosmetics requirements.

Adjusted. The rms is arbitrary indicator, see DC-30

We do not have "dead" pixels and only 260 "bad" pixels.

DC-39 P.6 L.12: "all automated scans are performed twice: scanning up and down in wavelength". Please report if a systematic difference was observed between the two scan directions (hysteresis effect).

Adjusted. No hysteresis effect was found in the measurements.

DC-40 P. 6 L.16-17: "The irradiance data has a better coverage in both spectral and spatial directions,..." Why is this the case? The difference between Sun and Earth ports is the diffuser before the slit. This may extend the illumination in the spatial direction (full slit illuminated), but not in the spectral. "...so a higher order M=7 could be applied on the parameters d and s, which show much more structure than the other fit parameters.". It would be useful to compare(in plots and by statistics) the variability of the different ISRF parameters.

Adjusted. The radiance measurements were intended to be used for the determination of the calibration key data, but the irradiance measurements turned out to be better suited (Sect. 4). In the updated method, there is no difference between order M used for radiance and irradiance measurements.

DC-41 P.6 L.21-25: "The quality of the parameter fitting is determined by comparing the measured ISRF data points with the ISRF that results from the parameter model". This implies that the measurements are independent from the chosen mathematical model of the ISRF. However, each "measurement point" is already the result of fitting the ISRF shape model to the measured ISSF. "In general, the parameter smoothing will result in better and smoother ISRF calibration key data due to averaging and interpolation". This seems to imply that "smoother" automatically means "better". However, exaggerated smoothing could systematically affect the accuracy of the derived ISRF shapes. Please comment on the chosen balance between smoothing and accuracy. Have measurements been repeated at outlier positions to verify they are due to random instrumental artefacts? "Possibly counter intuitive, the rms value will be slightly larger as the ISRF

data points are now compared with a smoothed ISRF instead of an optimized local ISRF that might be influenced by measurement imperfections". What does "slightly" mean? Please provide numbers. This "counter-intuitive" observation actually may indicate exaggerated smoothing. Since the individual fits are closer to the truth, I would certainly expect the rms to increase.

Adjusted. As stated in the revised manuscript (Sect. 3.3) "Minor problems with the OPO are still present in the ISRF data points (see Figs. 3–5), which may affect the ISRF fit yielding minor deviations of the fit parameters. Assuming that these deviations are random and that the shape of the ISRF varies only smoothly over the surface of the detector (as it is determined by the spectrometer optics), then the quality of the ISRF would benefit when the ISRF parameters are smoothed and interpolated using bivariate polynomial fitting". Smoother only means better, when the parameters of the true ISRF are smooth. Then a fit through noisy data will result in a better estimate of this parameter. The results presented in Sect. 4 show that small local deviations are present in the data.

DC-42 Please report on the overall statistics of the smoothing procedure. What is the scatter of the original ISRF parameters around the smoothed value used in the key data ? Please provide a table for all ISRF parameters.

Adjusted. Figures of the residuals of all ISRF parameters are provided in the revised manuscript

DC-43 P.6 L.31: "block width" -> "boxcar width"

Adjusted. We do no longer refer to the block distribution. However, in our opinion, the term 'boxcar function' is associated with time series and is not a distribution. Therefore, we prefer the term 'uniform distribution'.

DC-44 ISRF parameter iteration: This section provides some explanation for the question raised above: How can 8 parameters be estimated from ISSF of 4-5 significant pixels? 1)"Once the ISRF has been fitted, the skew and tails are known approximately, and can be included as fixed properties...". 2) "Therefore, the refitted block-width as a function of row and column is smoothed and used as a fixed property in the final ISRF fit". However, the approach of the "passes or stages" cannot solve an under-determined measurement problem. Parameters are estimated in stages, limiting the number of unknowns in each step to avoid under-determination. However, the results of each stage impact the ones of the next one (by fixing parameters estimated from an incomplete model). Each stage necessarily yields errors (incomplete description of then shape profile by fixing parameters), and these are propagated into the next stage. In addition, smoothing seems to be applied between the 4 steps (see 2)), which also introduces systematic errors, propagated into the next step. Please include a discussion, justifying the choice of the sequence, in which parameters are estimated.

Adjusted. We never fit the ISSF with many parameters, max 4. We use many ISSF fits to estimate the (local) ISRF, which can be determined with many (6-7) parameters. We do only fix one of the ISRF parameters (η or w). The method converges as is shown with synthetic data, see Sect. 3.2 of the revised manuscript.

DC-45 P.7 L.3-8: How have the "realistic ISRFs" been simulated, using the same parametric model? This paragraph seems to describe a convergence test of the fit procedure for ideal data (no noise, no parameterization errors, same ISRF). What are the start values chosen before stage1? How far from the known true values can they be? Please perform a test to demonstrate the convergence range of the technique.

Adjusted. Thanks to this request, we have extended the simulations and updated the method presented in the revised paper (Sect. 3.2).

DC-46 P.7 L.9-14 The discrepancy of the derived ISRF from the known true one may indeed indicate the numerical problem highlighted above and result from the approach (under-determined problem "solved" by fixing parameters and repeated iterations).

Not adjusted In our opinion, there is no under-determination problem. It is not surprising that the determined ISRF does not always match the true ISRF perfectly, although the true ISRF is generated with the same ISRF model used for the fit. This is because details of the true ISRF are lost in the poorly sampled ISSF fit. According to an extra simulation, the determined ISRF matches the true ISRF nearly perfectly when the ISSF signal is measured with 10 instead of 5 spectral pixels.

DC-47 P.7 L.14: "However, the differences between the true ISRF and the derived ISRF are less than 0.25% and are considered acceptable". Acceptability could be stated if this were the maximum possible error. However, it seems to be assumed that the true ISRF is computed by Eq. 1-4, so it is consistent with the mathematical model and no modelling error is included. Please test the approach with ISRF profiles deviating from the chosen mathematical model to demonstrate robustness of the approach. Please also quantify the impact/sensitivity to measurement noise. This latter would give useful information on the required quality of the calibration system.

Adjusted. Simulations with noisy data have been performed, and using a different mathematical model: the super-gauss (Beirle et al. 2017). The noise simulations were available for the original manuscript, but because the impact of noise is very small, it was not included in the discussion. Simulations with other values for tail parameter m revealed the sensitivity of the method to fixing this parameter and we have updated our method, now fitting both tail parameters. Changing the mathematical model by using a different function for the main peak still shows that convergence is very good, but the residuals are slightly larger wrt our model and wrt to the truth. Which is no surprise, and therefore not included in the revised manuscript.

DC-48 P.7 L.22: "A median has been taken over all rows illuminated." What does this mean (a medium of what parameter)?

Adjusted. A median has been taken of the ISRF data-points of all illuminated rows. And an ISRF fit is performed on these point.

DC-49 P.7; L.22: "From visual inspection of the displayed ISRFs, one can conclude that: (i) the ISRF is sharper and higher at higher column number (longer wavelength)". A large (.20% change) in ISRF width (2.3 - 2.7 pixels) should have been predicted by the optical design analysis. Is the magnification changing in spectral direction? See also comment on Figure 3 and 4. "(ii) the ISRF fit resembles the ISRF data very well, e.g. the residuals are very small,except where small artifacts can be identified in the ISRF data". The log plots show significant discrepancy in the wings and the residual plots show periodic structures, whose peak-to-peak amplitude correspond to almost half of the requirement (1%). I would change "very well" to "satisfactory" to "compliant". Since a new method is proposed (according to the title), the question arises if it provides superior performance over previous calibration campaigns. "(iii) the fit residuals of the irradiance ISRF are nearly a factor 2 smaller compared with the radiance ISRF". Please provide explanation.

Adjusted. (i) This is a visual inspection, but on the next page first bullet is the explanation (original paper): "Block width w of the ISRF is determined by the projection of the slit onto the detector and therefore decreases as a function of wavelength. As expected, no variation is seen over the spatial dimension of the detector (swath angle)".

This article is not about the requirements on the instrument.

(ii) The method is updated to perform a more accurate fit on the tails. Note that the significant discrepancy in the wings, observed by the reviewer, has no impact on the ISRF knowledge as defined by the requirement on the ISRF. It is agreed that we should have used "satisfactory" instead of "very well". With the updated method, the results have improved, therefore, in the revised manuscript we use "agrees well". (We never claimed to have a superior method, only a new method.)

(iii) We only observe this and cannot explain it.

DC-50 P.7 L.27: "The difference is likely due to differences in stray light in these measurements". It was stated before that all measurements (both for radiance and irradiance) are based on stray light corrected data. Now stray light is identified as the cause for a discrepancy between radiance and irradiance ISRFs. This needs to be commented to avoid confusion. How accurate is the stray light correction? Is the observed discrepancy (apparently averaged across the entire detector) explainable by the limitation of stray-light correction. All this would be part of an accuracy analysis of the "new method", which is currently missing.

Adjusted. The radiance measurements are not corrected for stray light, as is mentioned in Sect. 3.2, page 4, line 21. It is agreed that the quoted sentence is confusing and it is rephrased.

DC-51 P.7 L.28: "In all subsequent fitting, shape parameter m is fixed to 1.25 to enhance convergence of the curve-fitting routine...". Why is this value fixed whereas all others are (partially, sequentially) fitted? If it represents the "stray light level", as suggested in the text, why should it be constant? How much do the fitted values of m vary and deviate from the median value? In absence of this discussion this appears an arbitrary reduction of parameters for better convergence. Please comment and justify.

Adjusted. This is argued in the two sentences before the quoted text. And yes, this value of m is rather arbitrary and the only parameter fixed in the ISRF fit.

We should have mentioned in the original manuscript, that for the computations to determine the TROPOMI-SWIR ISRF, the tail parameter m was fixed to 1.25, based on the shape of the tails found in stray-light measurements.

In the revised manuscript, we present the results from an updated iteration scheme where both tail parameters are fitted to the ISRF data.

DC-52 P.7 L.30: "It has to be noted that the contribution of the tail to the ISRF is small ($< 10\%$) and only significant 1-2 pixels away from the peak". This seems in contrast to studies analysing the impact of ISRF shapes on CH4 retrieval, for which ISRF far wings are very important.

Not adjusted. Indeed, far wings are important for methane retrieval. However, the far wings are part of the stray light, not ISRF.

DC-53 P.7 L.31: "It has been verified that fixing the m parameter has negligible effect on the resulting ISRF and the fit residuals expressed in the rms value". Please provide evidence. Level 2 processing (not visual inspection) defines when an effect on the resulting ISRF is negligible. Has this been verified by retrieval simulations?

Not adjusted. No, level 2 is outside the scope of article.

DC-54 P.8 L.10: "Block-width of the ISRF is determined by the projection of the slit on to the detector and therefore decreases as a function of wavelength. "Again, please explain why a 20% reduction of slit width image is expected. In fact, the spectral sampling requirement (>2.5 pixel) seems to be violated across a wide spectral range.

Adjusted. Updated the text in the revised document: "The peak width due to the projection of the slit on the detector is constant when expressed as a wavelength interval, but expressed as a column distance it decreases towards larger columns (longer wavelengths), because the spectral dispersion changes.". The spectral resolution is about 2.5 nm, which is consistent with the range we have determined of 2.3–2.7 nm.

DC-55 P.8; L.19: "However, width parameter has been designed such that no errors are introduced by the ISRF parameter fit". What does it mean to "design" a parameter? Is it the choice of xi in Eq. 2?

Adjusted. The skew-normal width d is defined (not designed) such that no errors are introduced by the ISRF parameter fit. With sigma the interpolation of s would introduce errors, however, with d defined as a function of sigma and s such errors are not introduced.

DC-56 P.8; L.28: "The quality of the ISRF fits as determined with the parameters from the bivariate parameter-fitting models shown in Fig.6b. "Typo, "is" is missing.

Adjusted according to suggestion.

DC-57 P.8 L.30: "There area few small regions which coincide with the fine-scale structures visible in the skew-normal width, see for example around row 50 at columns 525 and 610". It is very difficult (if not impossible) to see the described fine-scale structures in Fig. 5 and 6 (a single row is not visible).

Adjusted. We replaced "fine-scale structures" with "patches". There are a few small regions in Fig. 6b such around row 64 column 600 and which corresponding to similar regions visible in Fig. 5b. See also Fig. 6, Hoogeveen et al. 2013.

You should be able to see single pixels (rows and columns) in the figures of the revised manuscript.

DC-58 P.9 L.3-4: “In general, the laser performed worse during the radiance measurements, yielding radiance ISRF measurements of poorer quality than the irradiance measurements”. Please provide details as to why and how much the laser has performed worse.

Not adjusted. This was a general impression.

DC-59 P.9 L.13-14: “On the left side of the detector, the block width of the radiance ISRF tends to be smaller than that of the irradiance ISRF. This subtle difference is attributed to the non-optimal scanning of the laser at these wavelengths.” Please indicate in which way the laser-scanning was non-optimal. Provide recommendation for optimum laser scanning.

Based on the poorer quality of the radiance fits, irradiance measurements are used for key data generation, while the radiance measurements merely serve for validation. What was the calibration time partition between irradiance and radiance measurements? Noting that - significantly more time was spent on radiance (100 scans versus 1) and - the differences are stated to be negligible one conclusion could be that ISRF calibration can be reduced to measuring irradiance only. Please comment.

Adjusted. We think that it was a very good decision to perform radiance as well as irradiance measurements, see also GC-x. Only afterwards, we could conclude that the stray-light correction performs very well on the irradiance measurements and the different light path has no significant impact on the determined ISRF. The way the measurements were implemented was the most efficient way to collect the data. However, we had to reject a significant number of irradiance measurements due to potential saturation, and reject a significant number of radiance measurements with partly illuminated rows. Therefore, we recommend to reserve more time for testing of the laser before the actual measurements. And for the radiance measurements: a FOV to 2.5–3 degrees (about 4–5 swath pixels) would be optimal, and perform successive swath scans with about one swath pixel overlap.

DC-60 It is appreciated that a section on in-flight calibration is included in this paper. However, more detailed information shall be given here. As a minimum, the text should provide - type of laser diodes (Distributed Feedback (DFB) - their distribution over the SWIR range This is particularly important since the ISRF changes significantly over the spectral range - scan range in nm (not roughly pixels) - mention of a negligible laser bandwidth Also, reference to publications shall be made here, which describe the instrument design of Tropomi.

Adjusted according to the suggestions of the reviewer.

DC-61 The comparison between ISRFs from the ISRF calibration campaign with external laser sources and measurements using the on-board lasers is of high interest in the context of future missions (e.g. Sentinel-5). Therefore, the authors should elaborate on the quantitative comparison between the ISRFs derived form the two sources. The current discussion is too qualitative and does not allow an evaluation of on-board ISRF monitoring. Adding a quantitative discussion (with plots) would significantly enhance the impact of this paper.

Adjusted. The measurements with the on-board diode-lasers are included in the analysis presented in the revised manuscript.

DC-62 P.9 L.26-27: “The scanning range is about 6 spectral pixels so that the ISRF can be monitored for one or two wavelength pixels per laser.” This is confusing, since the term “wavelength pixels” is not defined, or at least not discriminated to “spectral pixels” in the same sentence. I assume the authors mean the spectral range corresponding to two FWHM of the ISRF. Please rephrase.

Adjusted rephrased the sentence as suggested by the reviewer.

DC-63 P.9 L.25-26: “The laser wavelengths are scanned by tuning the temperature of the laser using a built-in thermo-electric cooler”. Builtin what: the (DFB?) laser or the calibration unit? Please describe more precisely.

Adjusted rephrased the sentence as suggested by the reviewer. The TE cooler is inside the laser housing.

DC-64 P.9 L.28: “As the diffuser is not moved during the measurements, there will be speckle”. In adequate wording for a science paper: replace “there will be speckle” by more precise formulation, e.g. “...the measurements will be affected

by speckle patterns due to the coherent laser light". Also, the connection to "moving diffusers" might not be clear to every user. Please add a sentence like: "Such patterns can be reduced by moving (e.g. rotating) the detector during the acquisition. However, this is not foreseen for in-flight calibration due to mechanical constraints (e.g. micro-vibrations)."

Adjusted. A monochromanic laser has speckle which can be suppressed by using a moving diffuser however, moving the diffuser is life limited, therefore, we take a median over the illuminated rows. The suggestion of reviewer is accepted with small modification.

DC-65 P.9 L.29: "Most speckle is removed by taking the median of the data of all illuminated rows". This is not understood and needs more explanation: I assume that the illumination of the on-board diffuser illuminates the Is the median taken over all rows (swath direction) to yield only one ISRF for the entire focal plane? Why the median and not the mean (affected by outliers)? The latter makes more sense for reasons of energy conservation. Should not the on-board calibration enable the determination of the ISRF across the entire swath width at 5spectral positions?

Adjusted. We have added some rationale here. Median is taken per laser. The data are already normalized, therefore a median can be used to reduce outliers (i.e. speckle).

DC-66 P.9 L.29: "During the commissioning phase, in-flight measurements with the on-board lasers will be performed with a moving and a fixed diffuser" Before it was mentioned that the on-board diffuser cannot be moved in flight. This may be different during commissioning phase, but deserves a sentence of explanation. Please provide details to improve clarity.

Adjusted according to the suggestion.

DC-67 P.9; L.31: "The ISRF obtained from these measurement scan be compared with the ISRF measured on-ground using the external laser and the on-board diode lasers to detect any possible changes". Remove "any", as this suggests infinite accuracy.

Adjusted. In this context, "any" means "it does not matter which type", not "every". The sentence is rephrased.

DC-68 P.9 L.31: "The monitoring ISRF is of sufficient quality to check for any degradation of the instrument but cannot be applied in trace-gas retrieval". Remove "any", as this suggests infinite accuracy. Please explain why the on-board ISRF can not be applied in tracegas retrieval. It may actually be useful to correct for launch effects and thermo-mechanical effects (de-focus). If such correction is made, it will be indirectly used in Level-2 processing. In general, be more quantitative in the comparison and evaluation of the on-board ISRF measurements. What is the expected and obtained accuracy, and over which spectral range?

Adjusted. The in-flight measurements with the diode-lasers using a fixed diffuser are intended for monitoring of ISRF stability. Details on the monitoring ISRF will be provided in a separate paper. An improved description of the monitoring ISRF is added to the revised manuscript, see 'Introduction' and 'In-flight monitoring of ISRF' (Sec. 5).

DC-69 P.10 L.3: "With an oscillating diffuser...". Please explain what is meant by "oscillating". I suppose that a calibration disk is moved back and forth by a few degrees (which is not quite oscillation), but the reader has to guess. Provide more details (see comment above).

Adjusted according to the suggestion.

DC-70 P.10 L.4-5: "...except that ISRF parameter smoothing (Sect. 3.5) is calculated from the ISRF fits of the few columns scanned per diode laser". Why is smoothing necessary here? I would assume that the ISRF is determined for the five ISRFs corresponding to the center wavelengths of the diode laser scan ranges. The ISRF fit procedure probably takes the on-ground parameters as start values, so large outliers should not be expected.

Adjusted. We can confirm that smoothing of the ISRF parameters is not needed for diode laser measurements with a moving diffuser. Therefore, no smoothing is used on these measurements presented in the revised manuscripts.

DC-71 P.10 L.4-5: "The column dependence of the shape parameters is neglected and the row dependence is smoothed by a second order polynomial". Does a square law (second order polynomial) describe the variation of optical effects in swath direction?

Adjusted. Yes, it is correct that the ISRF shape changes across the spatial dimension. However, these measurements are intended for monitoring of ISRF stability (see also DC-68).

DC-72 P.10 L.6-7: "Then the median ISRF is calculated from all ISRF data of the central one/two fully-scanned columns, neglecting any row dependence". Replace "one/two" by "one to two". Why is row dependence neglected (it has even been "smoothed" by a polynomial anyway)? Why taking a median ISRF, not a mean? Why are the 5 in-flight ISRFs not determined for all spatial samples across the swath? The approach to in-flight ISRF characterisation appears somewhat immature.

Partly adjusted. Please understand that we are discussing monitoring ISRF, not in-flight ISRF characterisation. Replaced by "one or two".

DC-73 P.10 L.9-10: Insert "the" between "moving" and "on-board".

Adjusted according to the suggestion.

DC-74 P.10 L.10-11: "The ISRF measured with the diode lasers is in close agreement with the ISRF calibration data, thus proving the usability of the method and validating the calibration data". Be more quantitative here. No plot nor table is provided for this important comparison. Please plot the 5 ISRFs measured with the on-board diffuser together with the on-ground ISRF, corresponding to the same detector pixels. Perform the comparison for both, moved and stationary on-board diffuser to quantify the impact of speckle patterns.

Adjusted. We have added the suggested figures. More information will be provided in the upcoming in-flight paper.

DC-75 P.10 L.12: "The monitoring ISRF deviates from the ISRF calibration data as could be expected". This is in contradiction to the sentence before.

Not adjusted. The definition of "monitoring ISRF" has to be taken into account. It is given at the end of page 9, end of line 30: the currently planned in-flight measurements are only suitable for monitoring the ISRF, because for deriving the ISRF a moving diffuser and longer spectral scans are required.

DC-76 P.10 L.12-13: "However, it is believed the method is sensitive enough to be used on board for long-term monitoring, being able to distinguish between changes in the real instrument ISRF and changes in the speckle pattern". Again, please provide a quantitative comparison, to substantiate this "believe".

Not adjusted. Details on the monitoring ISRF will be provided in a separate paper.

DC-77 Editorial: Straylight is written inconsistently in two ways: "stray-light" and "stray light". Any of them is fine (as well as one word), but be consistent.

Not adjusted. As a compound noun (the term on its own), we use "stray light". When used as a compound adjective, a hyphen needs to be added, e.g. "stray-light measurement".

DC-78 P.10 L.15-16: "A new and accurate method using a scanning OPO has been developed and applied to characterize the TROPOMI-SWIR ISRF". Remove "accurate", as this is a qualitative statement, that's need to be substantiated (How accurate? More accurate than other methods ?). In fact, its accuracy should be a results reported quantitatively here.

Adjusted. This sentence is removed, we no longer present a new method.

DC-79 P.10 L.18-20: "An iterative scheme to derive the SWIR ISRF has been developed, where the ISRF determined in a previous iteration is used to improve the ISSF model in the current iteration. The required accuracy of the ISRF is obtained within 4 iterations". -> It shall added here that the "iterative scheme developed" is not estimating free parameters in every iteration (as would be expected for an over-determined problem selectively fixing part of the parameters in every iteration is characteristic to the proposed new approach, so it has to be repeated here (and the consequences as well).

Adjusted. Rephrased the conclusions according to the reviewers comment.

DC-80 P.10 L.18-20: "The ISRF measured through the irradiance port using the solar diffuser has been compared with the equivalent ISRF measured via the radiance port. The differences between the ISRFs derived from both data sets are very small,...". Please be quantitative here "...and largely due to differences in stray-light treatment and laser scan imperfections". The statement that the discrepancy in ISRFs is "largely due to differences in stray-light treatment and laser scan imperfections", is an assertion, not a finding. It is suspected, but not demonstrated in this paper (it is even unclear, in which way stray light was treated differently).

Adjusted. The difference found between the irradiance and radiance ISRF is nearly gone with our improved method. It was an artefact of the fixed tail parameter m. Updated the conclusions accordingly.

DC-81 P.10 L.23-25: "The derived ISRF meets the requirement on ISRF knowledge and should thus be sufficient for methane retrievals". The claim that the derived ISRF meets the requirement on ISRF knowledge is based only on the claim that the fit residuals are smaller than 1% of the ISF peak. However, this only means that the parameters of the chosen mathematical representation can tuned to match the observed shape. It does not mean that the observed shape is accurate. An example is the stray light, which apparently affects the measurements differently in the radiance and irradiance ports. Does it mean that the true ISRF of a system depends on the quality of stray-light correction? By fitting the stray light into the line shape (parameter m), it becomes a feature of the true ISRF. It is proposed to include a critical appraisal of the approach and results in the conclusions. This should outlined also the limitations of the approach. Accuracy (in terms of deviation from a true ISRF) shall clearly be distinguished from consistency between a fit and a measured curve.

Partly adjusted. We are confident that this claim is valid, based on the residuals seen and the simulations done with synthetic data. The improved results wrt the original manuscript, and the many more simulations done only underline the original conclusion. True validation will be performed in-flight by level-2 processing.

As already explained and mentioned in the original manuscript, the radiance measurements are not corrected for stray light, while the irradiance measurements have to be corrected for stray light (as there is hardly any relevant stray light). Therefore, stray light is treated differently, and result in nearly equal residual stray light in both data sets. It is shown in the revised manuscript that the ISRF derived from both data sets are nearly equal.

1.2 Figures and Tables

1. Table 1: Why is the parameters m kept at 1.25

   Adjusted. Obsolete now that there is an updated iteration scheme, see also DC-51.

2. Table 3: Reference (Beers et al.). Please explain why a a publication on "Measures of Location and Scale for Velocities in Clusters of Galaxies" is relevant, resp,. applicable to describing the ISRF variation.

   Adjusted. This table is obsolete. A reference has been chosen that is easily available and that not only defines these general measures of location and scale, but also describes them well. The application for which the article describes them is indeed not relevant.

3. Fig. 3 and 4: Reporting the parameter values in the caption is difficult to associate to the 15 plots and does not provide useful information. Better include in them into the plot legends.The fit residuals (bottom row) clearly exhibit systematic (periodic) structure. This indicates a shortcoming of the ISRF shape model, which does not allow for periodic components. Are there physical reasons why periodic components in the ISRF shape are ruled out? Please comment on this and possibly propose an improvements. - Please include plots showing the difference between radiance and irradiance ISRF and discuss the reasons for differences.

   Adjusted. We agree with the reviewer and different plots are provided. Residuals of a non-perfect fit will always result in periodic residuals (and noise). The residuals in these figures are small, thus the model agrees satisfactorily with the true ISRF.

4. Fig. 5: Unclear Fig. caption: "In the white area, the ISRF fit failed (vertical stripes), the light is blocked by the entrance slit of the spectrometer (top and bottom) or a shield at the detector (left and right)" It does not seem logical, that an

entrance slit blocks light. Improve clarity by adding “white area at the edge” or changing color. The term “white areas” is confusing with most of the middle panel being white (not only the edge). Proposed to change color scale.

Adjusted according to suggestion.

5. Fig. 7: The number of “good” fits is drastically lower for radiance than for irradiance. Please provide explanation why this is the case.

   Adjusted. An explanation has been added to the caption.

6. Fig. A1 and A2: - Axis labels are missing (“Pixel No.”) - Figure captions should be understandable without reading the text. Please extend the Figure captions, briefly explaining the difference between “ISRF fit” and “ISRF parameter fit”. Comment: These plots indicate that there are systematic (not random) features being smoothed by the polynomial fit procedure, especially in the spectral dimension. Without evidence it is not obvious that they result from “laser artefacts”. Speckle effects (not mentioned in the text) should affect the spatial component stronger due to smoothing by spectral dispersion.

   Adjusted. All captions have been updated and additional explanation has been added.

7. The lower panel of Fig. A2 suggests that the “block width”, representing the image width of the entrance slit, varies from 2.3 to 2.7 (pixels ?). This 20% change over the spectral range(for the entire swath) should be readily verifiable by optical analysis (diffraction and spot size PSF). Please check(and report) the plausibility of the result with the optical performance analysis.

   Adjusted. Yes, this is consistent with the design as suggested by the reviewer.

8. Fig. A3 and A4: The plots show large variability of the resulting tail fraction and width from the ISRF fits. However, the ISRF parameter fit (“model”) seems to assume a single value across the detector. Has this also been fixed to the median value (as parameter m)?

   What is the justification, given the large, systematic variability? Convergence? It should be clarified (already in Section X) which parameter have been fixed to avoid the impression that the ISRF shape model has 8 free parameters.

   Adjusted. We thought that it was clear that a fixed value for the tail parameter m was used for all results presented in the original paper. This discussion is no longer relevant for the revised manuscript.

1.3 References

Buscaglione: ESA-SRDs should have no author name on Atmos. Meas.Tech. Discuss., doi:10.5194/amt-2017-438, 2017.

Adjusted. we do no longer refer to the SRD.

1 *Author comment on* "New method to determine the instrument spectral response function, applied to TROPOMI-SWIR" *by Richard M. van Hees et al.*, manuscript amt-2017-438, Anonymous Referee #2

We would like to thank Referee #2 for the very useful comments to improve our manuscript. In this document, we provide our reply to the comments. The original comments made by the referee are numbered and typeset in red. Page, line and figure numbers refer to the old version of the manuscript. After the reply we provide a revised version of the manuscript, each section is adjusted according to the review as described below. In this process, the original text has been substantially rewritten and/or reorganized. The output of 'latexdiff' is considered confusing and nevertheless provided. The revised version of the manuscript is part of this authors comment.

A brief overview of the major changes of the revised manuscript compared to the original manuscript:

- Title has been changed to "Determination of the TROPOMI-SWIR instrument spectral response function". We do no longer claim to present a new method.
- Clear distinction is made between what has been measured, what is determined from these measurements, and what is delivered. Inconsistent usage led to confusion.
- Thanks to the reviewer question to explain "laser artefacts", we discovered that:
  - a significant amount of irradiance measurements had to be rejected from the analysis due to an unexpectedly large variation of the laser signal yielding detector saturation.
  - we rejected radiance measurements from partly illuminated rows.
- The Pearson VII shape parameter m was fixed to 1.25 in our analysis presented in the original manuscript, because then the shape of the model matched the far wings observed in TROPOMI SWIR stray light very well. However, simulations – requested by the reviewer – showed that this assumption introduces significant errors in case the true ISRF has wings with a different shape. A new approach has been developed and verified with simulations that requires two successive ISRF fits per stage: first only fixing the tail fraction (from the previous stage), then only fixing w (improved guess from the previous ISRF fit). The method requires an initial guess of the tail fraction, which can be obtained from stray light or ISRF data (determined in stage 1).
- As a result of the above 2 points the analysis was improved such that the irradiance ISRF and radiance ISRF are more consistent.
- Analysis of the ISRF as determined from on-ground measurements with the on-board diode-lasers is included in the manuscript.

1.1 General comments

The paper "New method to determine the instrument spectral response function, applied to TROPOMI-SWIR" by R. M. van Hees et al. addresses the determination of the Sentinel-5p/TROPOMI instrument spectral response function (ISRF) in the SWIR spectral region. The authors claim that the accuracy of the derived ISRF is well within the requirements for accurate trace-gas retrievals, which is stated to be known with an accuracy of 1% of its maximum where the ISRF is greater than 1% of its minimum.

The paper addresses an important topic, as accurate knowledge of the ISRF shape and FWHM is essential to avoid systematic errors in trace-gas retrievals, especially for missions with stringent requirements on small systematic errors, e.g. greenhouse gas missions such as OCO-2/3 MicroCarb, GeoCarb, and the upcoming future European CO2 monitoring missions. The paper describes an iterative approach to accurately retrieve the ISRF shape from a series of measurements performed with an optical parametric oscillator (OPO) during the TROPOMI on-ground spectral calibration measurements at CSL. The topic is in general of high interest to be published in AMT. However, my impression is, that this paper resembles in large parts a technical document or an ATBD describing the applied mathematical algorithm without explanation or deeper analysis of the applied

steps. I agree in this regard with anonymous Referee #1 that this Manuscript should only be published in AMT after substantial revision. As a very comprehensive review is already given by Referee #1 addressing most of the issues I found in this paper, I will only briefly address some additional issues and my major points of concern.

To avoid further confusion, I will use in the following review the terms ISRF and ISSF as defined in this manuscript.

GC-i The authors claims to introduce a new method for ISRF characterization without giving any evidence for the case. To underline the issue, the authors should perform a more comprehensive literature review on the topic and should cite for instance previous literature like K. Sun et al. 2017, R.A.M. Lee et al. 2017, J.O. Day et al. 2011, Beirle et al. 2017, Liu et al. 2015, Dirksen et al. 2006 and others.

Adjusted. We do no longer claim to present a new method, updated the title, abstract, section "Methodology" and conclusions accordingly. Thank you, for providing literature suggestions, citations have been added where relevant.

GC-ii The authors are representing an iterative approach to derive the high resolution ISRF from a series of ISSF measurements, claiming that the high resolution ISRF could not be measured directly.
However this is only true, if the spectral accuracy, line-width and "intensity" of the used optical stimulus (in this case the OPO) is insufficiently known. I have the impression, that this the case for the used OPO setup as the authors stated in two cases: "During the on-ground calibration measurements, the absolute wavelength of the source is not measured accurately enough" and "The ISRF parameters cannot be retrieved directly from the measurements, because the wavelength and intensity of the signal are unknown and have to be determined via the ISSF". If this is the case, the authors should clearly state in the introduction section of the manuscript, that the iterative approach presented in this paper is required due the insufficient accuracy of the used spectral calibration equipment and is therefore the novelty of the described procedure. However a carefully design of the calibration stimulus should be able to overcome this problem and should allow a direct high resolution measurement of the ISRF for each detector pixel. Nevertheless, construction of the ISSF from such measurements could be tricky, as in addition, detector issues, as for instance differences in pixel to pixel cross-talk, PRNU etc. needed to be considered.

Adjusted. A description and rationale has been added to the introduction and details are presented in the section "Calibration measurements". The calibration measurements had to be performed within a given time slot, which did not allow to perform the measurements with a very accurate wavelength meter and wait for a stable laser at a given wavelength. Instead, the laser slowly scanned during data taking and each frame is treated as the measurement of an ISSF per row, where the column index of the fitted center is used as a wavelength label for the row data. The line-width of the OPO laser is negligible in comparison to the instrument line-width and only smeared by approximately 1/80 px due to the scanning speed.

We are aware that small local variations are neglected/ignored. Any remaining residuals are checked by examining the results, presented in the discussion section, see page 8, lines 12–15 in the original manuscript.

GC-iii The authors fail to justify, why for instance the Pearson VII distribution is used or why the given iterative approach is chosen. There is no comparison with other possible distributions, see for instance Beirle et al. 2017. Also the use of filter parameters is not sufficiently justified. For instance RMS filtering of ISRF fits with an RMS larger than 0.0065 is applied. Why not 0.005 or 0.008?

Adjusted. The iterative approach is now supported by a longer discussion of synthetic data (Section 3.2). A justification why the Pearson VII distribution was used is added to the section "Methodology". Several other possible distributions have been tested, including the super-gauss, an asymmetric version of the exponential power distribution (Beirle et al. 2017). They do not represent the TROPOMI-SWIR ISRF with enough precision. The asymmetry of the super-gauss is provided by combining separate functions for two halves of the peak, which results in a continuous curve, but a discontinuous derivative, leading to larger residuals in the peak. The tails of the super-gauss are far too steep. A justification of the rms and its threshold is added to Sect. 3.1 and Sect. 4.

GC-iv The authors claim in the abstract and the conclusion that "The accuracy of the derived ISRF is well within the requirement for accurate trace-gas retrievals". However, the method described in this paper presents only a fit procedure able

to fit the measurements with a given accuracy. The authors lack to provide a comprehensive error budget, including effects on PRNU, detector non-linearity and other mostly detector related effects to underline that claim. The paper also lacks to provide an independent verification for that claim, see for instance Frankenberg et al. 2015 (doi:10.5194/amt-8-301-2015) for comparison.

Partly adjusted. We are confident that this claim is valid, based on the residuals seen and the simulations done with synthetic data. The improved results wrt the original manuscript, and the many more simulations done only underline the original conclusion. True validation will be performed in-flight by level-2 processing.

Measurement artefacts have been identified by a comparison of the irradiance, radiance and diode-laser measurements. Detector non-linearity has been explicitly addressed in Sect. 2.3. We have found no evidence for residual pixel-to-pixel effects, except for a correlation with photo-sensitivity of the SWIR detector, see discussion in Sect. 4.

GC-v The use of the terms ISRF and ISSF as defined by the authors is confusing and not consisted with the paper of Hasekamp et al. 2016. cited for justification of the ISRF knowledge requirements. Typically the (derived) ISRF function is used to convolute a theoretical high resolution RTM spectrum to lower spectral resolution in the retrieval. The (optical) ISRF is typically defined as the spectrometer response to a uniform monochromatic stimulus and approximated as the convolution of the slit image (typically represented by a boxcar function) with the PSF (or more accurately, the LSF) if the detector properties are neglected, see for comparison also Caron et al. 2017. This definition of the ISRF function is the mirror function of the ISRF function as defined in this paper, when assuming the changes of the (optical) ISRF are smooth over the image plane.

Not adjusted. The mapping between object and image can be performed in two directions. The terms 'spread function' and 'response function' are often used interchangably, but within the TROPOMI project they were coupled to the two different mapping directions. In this paper, the terms are applied to the spectral behaviour of the instrument. The difference is clearly stated, using also a diagram.

The ISSF is measured and the ISRF is based on measured data, so a term like 'instrument spectral measured function' is not distinctive. Both are line-shaped, so the term 'instrument line shape' is also not distinctive. By defining the two abstract terms 'spread' and 'response', the two functions are given distinctive names, but still clearly related. For the purposes of this paper, these terms help explain the algorithm.

The ISRF as determined in this algorithm is the function used by Hasekamp et al., including the orientation (checked with the authors).

GC-vi Section 5 needs a deeper analysis to justify statements given in this section. For instance statements as: "However, it is believed the method is sensitive enough to be used on board for long-term monitoring, being able to distinguish between changes in the real instrument ISRF and changes in the speckle pattern" needs to be justified by analysis or removed.

Not adjusted. We agree that deeper analysis is needed. Therefore, the analysis presented in this paper are on the on-ground measurements. In-flight monitoring of the ISRF will be addressed in a separate paper.

1.2 Specific comments

SC-1 Section 2, P.3 L.1-6: (calibration measurements): Is a Wavemeter and a (spectral response and linearity) calibrated monitoring detector used in the setup? Is there any other type of direct laser wavelength monitoring integrated in the setup. If yes, what is the accuracy? What is the laser linewidth?

Adjusted. A wavelength meter was available, which was only used by the operators to set the start wavelength of each automated wavelength scan. The relative intensity was checked by examining raw images. There was no other useful monitoring detector.

The line-width of the OPO laser is negligible in comparison to the instrument line-width and only smeared by approximately 1/80 px due to the scanning speed.

SC-2 P.2 L.23-24: "using the wavelength assignment derived from an independent wavelength calibration measurement." State how accurate the independent wavelength measurements are, as this has impact on the accuracy of the derived ISRF shape and how is it done? Is a different setup used?

Adjusted. We have used a CO spectrum for the wavelength calibration, using a white light source, a CO gas cell and a molecular line list. This provides a smooth wavelength dependence over the whole detector. It is only used in the last step, when the ISRF has already been determined, to convert wavelengths in pixel units to wavelengths in nm. The absolute wavelength is not needed for this, only relative wavelength distances to the line center.

SC-3 P.3 L.25: convolution of block distribution should be exchanged by convolution with a boxcar function in the entire manuscript.

Adjusted. We do no longer refer to the block distribution. However, in our opinion, the term 'boxcar function' is associated with time series and is not a distribution. Therefore, we prefer the term 'uniform distribution'.

SC-4 P.3 L.26: The optics is "blurring" the image by the spectrometer PSF (LSF) which could be asymmetric and often has also oscillations in the wings. Using a normal distribution for approximation of the PSF/LSF is only a first order approximation.

Adjusted according to suggestion, see Sect. 3.

SC-5 P.4 L.21: The measurement data is corrected for background, PRNU and stray light. Why is the data not corrected for detector non-linearity as the used MCT detector can have non-linearity in the % range over the dynamic range?

Adjusted. We refer now to the paper of Hoogeveen et al. 2013. Non-linearity of the SWIR detector is measured to be very small and, therefore, neglected (also in the operational processor).

SC-6 P.5 L.2-3: Wavelength and intensity of the signal are unknown: are they really completely unknown? What is the wavelength accuracy/knowledge and stability of the used OPO?

Adjusted. See answers SC-1 and GC-ii.

SC-7 P.5 L.11-12: The laser wavelength scan is not regular->why?

Adjusted. Added description, revised manuscript, Sect. 2. The wavelength variation involves changing the temperature of a laser crystal, which is not a linear process. It can only be made regular by using negligibly sized steps and waiting longer, but there was no time for that. It is also not needed.

SC-8 P.5 L.17: larger than 6% -> justify 6%, why not 5% or 7%?

Adjusted. This value is arbitrary, and explained more in the revised manuscript, Sect. 3.1.

SC-9 P.5 L.21-23: "It is expected that the fit parameters that define the local ISRF vary only smoothly over the surface of the detector as this is determined by the spectrometer optics. Therefore, a bivariate polynomial fitting is used to smooth and to interpolate the ISRF fit parameters". This can only be expected for the optical system in case additional detector effects are neglected. However in a real life scenario, the effective ISRF is additionally compromised by insufficiently corrected detector effects. I guess, if all detector effects could be corrected to the required level, the resulting derived ISRF would be a smooth function. Therefore this statement contradicts the claim by the authors that the accuracy of the derived ISRF is well within the requirement for accurate trace-gas retrievals. Also the statement that most of the outliers are at same wavelengths and are caused by laser artefacts / scan imperfections contradicts that claim, as laser artefacts and scan imperfections need to be considered in the total error budget of the ISRF. See also discussion on P9. L.10-13. For a better understanding it would be helpful to show ISRF cross sections measurements from compromised rows, where the fit procedure fails.

Adjusted. The pixel-to-pixel variations of the gain (PRNU) are only about 1 % before correction and orders of magnitude smaller after correction. Outliers (due to saturated measurements and bad laser behavior) are rejected and have no impact on the fit residuals, but not all effects from saturated measurements could be removed. The remaining high-frequency

features are ignored by the ISRF parameter smoothing, but they can be seen in the residuals, which are small. They are presented in the revised manuscript and analyzed looking at the differences between the irradiance, radiance and diode-laser measurements (Fig. 7), and at pixel-to-pixel effects in the residuals before and after smoothing (Fig. 8, 14, 15).

SC-10 P.9 L.3-7: A sketch of the radiance and irradiance calibration setup would be very helpful. Is the on-board diffuser illuminated via the integrating sphere during irradiance measurements or directly by the OPO? I would expect from the text and as a ND filter is used in front of the OPO (P3,L10), that the OPO is directly illuminating the on-board diffuser. This would result in a better SNR as more light is entering the instrument but also in more spectral structures introduced by the laser + on board diffuser combination. The authors state that the integration sphere in combination with the spinning mirror is used to avoid speckles. So please justify, why the irradiance measurement which obviously should have more spectral structures is used for the key-data. This furthermore triggers the discussion what is physically more meaningful, a smaller spread in the data as observed in figure 8 and also stated on P.9 L9-10 for the radiance measurements or a better RMS of the fit as observed for the irradiance measurements, which can for instance be caused by a larger number of fit parameters defining the degree of freedom of the fit.

Adjusted. A sketch of the setup has been added (Fig. 2 in the revised manuscript), the description of the setup for irradiance and radiance measurements have been separated. In both cases there is an integrating sphere between the OPO and the instrument.

For SWIR, the spectral features introduced by the on-board diffuser in irradiance data have much larger periods than the ISRF range, so they are constant per ISRF. Because they are absent in radiance data, the radiance measurements were intended to be used for the key data, but the irradiance measurements turned out to be better suited. The difference found between the irradiance and radiance ISRF is gone with our improved method; it was an artefact of the fixed tail parameter m.

SC-11 P.9 L.14: How is it judged that the observed difference is attributed to the non-optimal scanning of the laser? What is an optimal scan? In fact, the wavelength accuracy and intensity of the laser are imperfectly known as previously stated.

Adjusted. The sentence was found confusing and is rephrased.

SC-12 P.9 L.15: Is a difference of the block width between radiance and irradiance measurements of $\approx 5\%$ as shown in Fig 8c for the left side of the detector really negligible in comparison to the requirements?

Adjusted. See answer SC-11. The difference found between the irradiance and radiance ISRF is gone with our improved method. It was an artefact of the fixed tail parameter m.

SC-13 P.9 L.29: Would taking the median over an entire row not imply the assumption, that there could not be a relative ISRF change along the row?

Adjusted. The sentence was found confusing and is rephrased. We take the median over all data along a swath (all rows). Yes, it is correct that the ISRF shape changes across the spatial dimension. However, these measurements are intended for monitoring of ISRF stability. Details on the monitoring ISRF will be provided in a separate paper.

SC-14 P.10 L.1: The statement is contradicting to the previous statement. In fact, if the laser can be used to recalibrate the ISRF for a significant part of the detector, they can be used for trace gas retrievals.

Adjusted. Improved the description of the monitoring ISRF, see the revised manuscript 'Introduction' and 'In-flight monitoring of ISRF' (Sect. 5).

SC-15 P.13 Figure1(a): it should be added to the axis caption that "column" is the detector spectral direction and wavelength is the wavelength derived for each ISSF measurement by the fitting procedure.

Not adjusted. The wavelength is that of the incoming light from the source. That it is determined in the algorithm is not relevant in the description of the difference between ISSF and ISRF. The column index is indeed in the spectral direction, but in this context it is more important that it corresponds to detector pixels. Since a pixel distance could be both along a column or row, the word 'column' seemed slightly more informative.

SC-16 P.17 Figure 5: The figure caption should be clearer. For instance “white areas on top and bottom of the detector blocked by the slit for stray light correction and DC monitoring” or something similar.

Adjusted. The figure caption(s) are adjusted. The blocked edges of the SWIR detector are not useful for any correction or monitoring.

SC-17 P.22 Table 3: How meaningful is the fit of the skew parameter s with > 100% error?

Adjusted. This was indeed confusing. The range of each parameter across the detector was listed, not the error. This table has been removed from the revised manuscript.

~~New method to determine~~ Determination of the ~~instrument spectral response function, applied to~~ TROPOMI-SWIR instrument spectral response function

Richard M. van Hees[1], Paul J. J. Tol[1], Sidney Cadot[1,2], Matthijs Krijger[1,3], Stefan T. Persijn[5], Tim A. van Kempen[1], Ralph Snel[1,4], Ilse Aben[1], and Ruud W. M. Hoogeveen[1]

[1]SRON Netherlands Institute for Space Research, Utrecht, the Netherlands
[2]Jigsaw B.V., Delft, the Netherlands
[3]Earth Space Solutions, Utrecht, the Netherlands
[4]Science and Technology B.V., Delft, the Netherlands
[5]VSL Dutch Metrology Institute, Delft, the Netherlands

Correspondence: Dr Richard van Hees (r.m.van.hees@sron.nl)

Abstract. The Tropospheric Monitoring Instrument (TROPOMI) is the single instrument on board of the ESA Copernicus Sentinel-5 Precursor satellite. TROPOMI is a nadir-viewing imaging spectrometer with bands in the ultraviolet and visible, the near infrared and the short-wave infrared (SWIR). An accurate instrument spectral response function (ISRF) is required in the SWIR band where absorption lines of CO, methane and water vapor overlap. ~~Therefore a novel method for ISRF determination for an imaging spectrometer was developed and applied to~~ In this paper, we report on the determination of the TROPOMI-SWIR ~~band. The ISRF of all detector pixels of the SWIR spectrometer has been measured during an~~ ISRF during an extensive on-ground calibration campaign. Measurements are taken with a monochromatic light source scanning the whole detector, using the spectrometer itself to determine the light intensity and wavelength. The accuracy of the ~~derived ISRF~~ resulting ISRF calibration key data is well within the requirement for ~~accurate~~ trace-gas retrievals. Long-term in-flight monitoring of ~~the ISRF is guaranteed by the presence of five SWIR~~ SWIR ISRF is achieved using five on-board diode lasers.

1 Introduction

The Tropospheric Monitoring Instrument (TROPOMI) is the single payload of the Copernicus Sentinel-5 Precursor (S5P) satellite mission~~(Veefkind et al., 2012)~~. The instrument maps the Earth's atmosphere ~~in two dimensions~~ using two spectrometer modules behind a common telescope, one covering the ultraviolet/visible (270–495 nm) and near-infrared (675–775 nm), and the other covering the short-wave infrared (SWIR) spectral range 2305–2385 nm. The ~~latter with a~~ spectral resolution of the SWIR spectrometer is about 0.25 nm ~~and~~ with a spectral sampling ~~distance of~~ interval of typically 0.1 nm. The TROPOMI instrument measures sunlight reflected by the surface and atmosphere of the Earth via the radiance port. Direct sunlight is measured via the irradiance port and internal diffuser for calibration purposes (Veefkind et al., 2012). The SWIR spectrometer (developed by SSTL, United Kingdom) consists of a slit, collimator mirror optics, an immersed grating ~~(developed by SRON, van Amerongen et al. (2012))~~(developed by SRON, van Amerongen et al., 2012), anamorphic prism and camera op-

tics consisting of multiple lenses, and a HgCdTe detector (developed by Sofradir, France). The detector has 1000 columns in the spectral dimension and 256 rows in the spatial dimension of which about ~~216 rows and 984 columns are nominally~~ 975 columns and 217 rows are illuminated.

The TROPOMI-SWIR band is used for the retrieval of ~~atmospheric CO and methane columns. The methane concentrations have to be measured with an accuracy of better than 1%. Therefore, the requirement on systematic errors is very strict~~the atmospheric trace gases carbonmonoxide, methane and water vapor. Simulations have shown that in particular the methane retrieval is very sensitive to errors in the instrument spectral response function (ISRF ~~), which is used to include the measured absorption line shapein the modeled Earth spectra (Hu et al., 2016). Therefore~~or instrument line shape). As a result, the requirement on the ISRF is formulated that it should be known with an accuracy of 1% of its maximum where the ISRF is greater than 1% of its maximum ~~(Buscaglione, 2011; Hu et al., 2016)~~(Hu et al., 2016). To reach the required accuracy, ~~the ISRF has~~ data have been measured with high spectral resolution ~~during on-ground calibration measurements~~ using a scanning monochromatic light source covering the SWIR band.

~~In the literature, the instrument spectral~~ For a comprehensive overview of various approaches applied to determine the ISRF for relevant past and future space-borne missions we refer to Sun et al. (2017). In summary, for the pioneering mission GOME on ESA's ERS-1 satellite and for the SCIAMACHY instrument on board ESA ENVISAT, no high-resolution ISRF was measured at all on ground. Information was derived from an on-board spectral light source with discrete line emissions (Schrijver et al., 2009). For the later OMI instrument on board NASA's EOS Aura satellite a white light source followed by a high-resolution monochromator was used to create a comb of narrow spectral lines (Dobber et al., 2006). A similar approach was followed for the other three bands of TROPOMI (Kleipool et al., 2018). The NASA OCO instruments had the ISRF measured with a monochromatic line source (Day et al., 2011; Sun et al., 2017).

Often no distinction is made between a *spread* function ~~(ISSF) and instrument spectral~~ and a *response* function~~(ISRF) are often confused~~. In this paper, ~~we define a "spread function" and a "response function"~~ the functions are defined as follows: a ~~"spread function "~~ spread function maps an object to image space, which involves many detector pixels; a ~~"response function "~~ response function maps an image to object space, which is a property of a given detector pixel. The ~~ISSF~~ instrument spectral *spread* function (ISSF) is measured simply by illuminating the spectrometer slit homogeneously with a monochromatic source and taking a detector image (frame). In the spectral dimension, about ~~4–5 points~~ 5 pixels have significant signal, as expected from the spectral oversampling. This is the spread function of the instrument for this wavelength. In Fig. 1b it is shown as a red cross section. When the wavelength is scanned in small steps over a set of frames, the signal in those frames for a given pixel ~~(that is illuminated in at least some of the frames)~~ forms an ISRF, with an arbitrarily fine sampling. This is shown as a green cross section in Fig. 1c. There is an infinite number of ISSFs (one at each wavelength) and a finite number of ISRFs (one for each pixel).

The ISSF consists of one sample from each ISRF of a few neighboring pixels on a row. If the ISRF varies negligibly between these pixels, the ISSF is a sparsely sampled version of this ISRF. ~~However, Fig.~~Figure 1a shows that the samples taken with increasing column index are ISRF points from the right side of the peak to the left side: the ISSF samples a mirrored ISRF, indicated by the light-green line ~~at bottom left~~in Fig. 1b.

~~During the~~ By design, the ISRF and dispersion of the TROPOMI instrument should vary only smoothly in the spectral and spatial dimension. This assumption will be validated in this study by determining the local ISRF for many pixels of the SWIR detector. The assumption is also used in the data analysis to interpolate the ISRF to pixels for which there are no reliable measurements and to reduce the effect of outliers.

The ISRF is determined from measurements using the radiance port and the irradiance port. Although differences between both data sets are not expected, they could not be ruled out. Light entering via the irradiance port follows a different path, via a diffuser.

The ISRF measurements are part of the on-ground ~~calibration measurements , the absolute wavelength of the source is not measured accurately enough~~calibration campaign performed at the Centre Spatial de Liège (CSL) in Belgium (Kleipool et al., 2018). The limited time slot did not allow to perform the measurements with a very accurate wavelength meter and wait for a stable laser at a given wavelength. Instead, the laser scanned slowly during data taking and each frame is ~~seen~~ treated as the measurement of an ISSF per row, ~~and~~ where the column index of the fitted center is used as a wavelength label for the row data. The ISRF of a pixel, based on data from a set of frames, is then the normalized signal as a function of wavelength in pixel units (non-integer values). It should not be confused with the ISSF, which has basically the same horizontal scale but then in integer pixel ~~values~~units. Only at the end of the algorithm is the ISRF of a given pixel converted to a function of wavelength in nm, using the wavelength assignment derived from an independent wavelength calibration measurement. This results in the ISRF calibration key data (CKD) which are used in trace-gas retrievals.

The measurements used for the ISRF characterization are presented in Sect. 2. A description of the method and algorithm used to derive the ISRF for all illuminated pixels is presented in Sect. 3. Details of the algorithm are discussed in Sect. 4, ~~where we also present the ISRF analysis~~ as well as the ISRF results based on the on-ground calibration measurements. ~~The validation of the irradiance ISRF using the radiance measurements is discussed at the end of~~ Comparison of the ISRF as determined from irradiance, radiance and on-board diode-laser measurements are discussed and the choice for the CKD is motivated. The validation of the SWIR ISRF is also part of the discussion (Sect. 4). The in-flight monitoring of the ISRF is briefly described in Sect. 5. This is followed by the conclusions in Sect. 6.

2 Calibration measurements

~~In this section, we describe the setup used during the ISRF calibration measurements, performed at the Centre Spatial de Liège (CSL) in Belgium, during the on-ground calibration campaign (Kleipool et al., 2018). In this setup, the~~

2.1 Measurements with the external laser

The light source employed in the ISRF characterization measurements is a 2 W continuous-wave optical parametric oscillator (OPO), custom-built by VSL (Delft, the Netherlands). The OPO is pumped by a single-frequency distributed feedback (DFB) fiber laser operating at 1064 nm which is amplified to 10 W by an ytterbium fiber amplifier. The OPO wavelength is set coarsely between 2290 nm and 2390 nm by manually setting the temperature of the periodically poled lithium niobate crystal

and rotating the etalon mounted on a galvo. The wavelength is scanned continuously over a range of about 2 nm by applying a changing piezo voltage to the fiber laser and simultaneously changing the crystal temperature with a predetermined dependence on the piezo voltage. ~~To avoid~~ The setup for the radiance and irradiance measurements is shown in Fig. 2.

During the radiance measurements, the power entering the instrument has been reduced with a neutral density filter, just after the OPO, to avoid saturation of the SWIR detector. To suppress speckle patterns on the detector, the light is sent to an integrating sphere via a high-speed spinning mirror with a small angle between the rotation axis and the normal. The light exits the integrating sphere and is collimated with a field stop and an off-axis parabolic mirror. The ~~instrument is mounted on a cradle in order to~~ beam corresponds to a swath-angle coverage of 1.1°, illuminating approximately 2 pixels in the swath direction. To scan all swath angles~~in a~~ , the instrument is mounted on a cradle. The automated wavelength scans were repeated 109 times to cover the range of 108° around nadir. ~~The beam corresponds to a swath-angle coverage of 1.1°, when the radiance port of TROPOMI is used. During irradiance ISRF measurements,~~ At each swath position, the wavelength was scanned in the opposite direction. The detector covers the wavelength range 2300–2389 nm, but due to time constraints, measurements have been collected of wavelengths between 2304 nm and 2386 nm, which covers the performance range 2305–2385 nm.

The field stop and parabolic mirror were replaced by a set of collimating lenses during the irradiance measurements, in which the whole swath is illuminated at once via the on-board solar diffuser. ~~With a neutral density filter just after the OPO, the power entering the instrument has been reduced to avoid saturation. During a wavelength scan , the~~ As there was no need to repeat measurements at different swath angles, each automated wavelength scan was repeated with increasing and decreasing wavelength scan direction. The irradiance data covers the full wavelength range.

Each automated wavelength scan of about 2 nm, or 20 spectral pixels, took about 165 seconds. The data is collected at 10 ~~Hzduring approximately 165 seconds, taking approximately~~ Hz, yielding 1650 ISSF samples in total and about 80 samples ~~. Background measurements are taken by closing a shutter between the filter and the spinning mirror. For more details on the setup, see Tol et al. (2018, Sect. 2).~~

~~Calibration measurementsvia the radiance port and irradiance port have been performed to verify that they are identical. The measurements of the radiance ISRF lasted significantly longer as not only the laser wavelength had to be scanned, but also the swath angle. Therefore,~~ per pixel. The laser scan speed was not constant, despite the small adjustments of the piezo voltage during the wavelength scan. Due to the large number of samples taken, this has no negative impact on the ISRF determination. During the measurements, a dedicated quick-look facility was available to monitor the conditions of the instrument and to show the measured signals in real-time. A wavelength meter was used by the operators to set the start wavelength of each automated wavelength scan. The operators kept a log during the measurement campaign to reported on issues during the measurements. Overall, more problems were reported during the radiance ~~ISRF measurementsabout 100 wavelength scans per manual wavelength setting are performed to cover all swath angles, and due to time constraints one-way wavelength scans were taken. During the irradiance ISRF measurements each wavelength scan was performed up and down~~measurements, mostly due to instability of the source due to drift and mode hopping. Detailed data analysis was performed after the measurement campaign was finalized, denying the possibility to improve or redo certain measurements.

3 Methodology

In this section, we describe our method to characterize the TROPOMI-SWIR ISRF in general. Specific measurement issues will be discussed in Sect. 4.

2.1 ISRF shape

To derive the ISRF from the measured ISSF, an assumption has to be made on the shape of the ISRF. In an iterative way, the ISRF shape can be made more complex to fit its detailed features. At the start of the procedure it is assumed that the ISRF is just a convolution of a block distribution with a normal distribution. This corresponds to an image of the slit on the detector (the block), with the optics blurring the image (the normal distribution). In later iterations, the normal distribution is allowed to be skewed. As the measurements show longer spectral tails, a dedicated tail function is added to the ISRF to account for this. For the tails, several functions were tried. In the end, the Pearson type VII distribution resulted in the best fit.

Mathematically, the ISRF is modeled by the weighted sum of functions for the peak and the tails. Peak function $\mathcal{S}$ is a skew-normal distribution convolved with a block distribution and tail function $\mathcal{P}_7$ is a Pearson type VII distribution.

$$\mathcal{R}(c;d,s,w,\eta,\gamma,m,c_0) = \begin{array}{l}(1-\eta)\,\mathcal{S}(c;d,s,w,c_0)\\ +\eta\,\mathcal{P}_7(c;\gamma,m,c_0)\end{array}$$

The two constituent distributions have been normalized to area 1, so that weight η is the fraction of the total area in the tail function. The weight has to be in the range $0 \le \eta \le 1$. The ISRF model is a function of column index c with its mean at c_0, where both are non-integer numbers.The wavelength assignment is performed when the ISRF calibration key data (CKD) are generated: the SWIR wavelength calibration key data are used to assign a wavelength to the center of a pixel and the ISRF is defined relative to that wavelength.

Peak function $\mathcal{S}$ has three shape parameters, the skew-normal width d, skewness parameter s and block width w:

$$\mathcal{S}(c;d,s,w,c_0) = \begin{array}{l}\frac{\mathrm{erf}\left(\xi_+/\sqrt{2}\right)-\mathrm{erf}\left(\xi_-/\sqrt{2}\right)}{2w}\\ -2\frac{T(\xi_+,s)-T(\xi_-,s)}{w}\end{array}$$

with

$$\xi_\pm = \frac{\sqrt{1-\delta^2}}{d}(c-c_0\pm\tfrac{w}{2})+\delta \quad \text{where} \quad \delta = \frac{\sqrt{2}\,s}{\sqrt{\pi(1+s^2)}}$$

and using Owen's T function (Patefield and Tandy, 2000)

$$T(z,s) = \frac{1}{2\pi}\int_0^s \frac{\exp\left(-\frac{1}{2}z^2(1+t^2)\right)}{1+t^2}\,\mathrm{d}t.$$

~~Tail function $\mathcal{P}_7$ has two shape parameters, γ and m:~~

$$\mathcal{P}_7(c;\gamma,m,c_0) = \frac{\Gamma(m)}{\gamma\sqrt{\pi}\Gamma(m-\frac{1}{2})}\left(1+\frac{(c-c_0)^2}{\gamma^2}\right)^{-m}$$

~~with $m > 1/2$ and $\gamma > 0$. The specific case $\mathcal{P}_7(c;\gamma,1,c_0)$ is the Lorentz distribution with half width at half-maximum γ.~~

2.1 Measurements with the on-board diode-lasers

TROPOMI's on-board calibration system includes five distributed feedback lasers (Nanoplus, Germany) to monitor stray light and ISRF (Veefkind et al., 2012). The wavelength is scanned by tuning the temperature of each laser using the thermo-electric cooler integrated in the laser housing. These diode lasers are tuneable over a range of 7 nm (about 70 pixels), but due to operational constraints monitoring is restricted to 0.6 nm (about 6 pixels). Analysis revealed that the lasers are very stable and can perform very precise wavelength scans are scanned in 430 seconds, taking over 700 samples per pixel (at 10 Hz). The distribution of the five lasers is listed by their operational wavelength (at the center of the scan) and corresponding pixel (between brackets): 2311.8 nm (154), 2328.2 nm (341), 2340.0 nm (471), 2357.5 nm (659) and 2372.2 nm (813). The lasers illuminate the SWIR spectrometer via a dedicated diffuser. The speckle in the laser signal can be suppressed by oscillation of the diffuser around the nominal angle. However, as the diffuser mechanism is a life-limited item, only during the on-ground calibration campaign and during the in-flight commissioning phase, measurements will be performed with a moving diffuser.

2.2 Data preparation

~~The measurement data are corrected for background, pixel response~~ The three measurement data sets (irradiance, radiance and diode-lasers) are corrected for detector features such as memory and pixel response non-uniformity ~~and stray light(irradiance only). Absolute (ir)radiance calibration is not required. Readouts from bad pixels are discarded in~~ (PRNU), using the operational level-1b processor developed by KNMI (Kleipool et al., 2018). Changes of the background signal are removed by dedicated background measurements, which include offset, detector dark current and thermal background signal. These measurements are performed regularly during the measurements with the external laser, by blocking its signal with a shutter (Fig.2). For the on-board lasers the background signal is measured before and after each wavelength scan. A non-linearity correction is not implemented in the operational processor. It is also not needed for the ISRF characterization, as the error is small: detector non-linearity was measured to be about 0.1% to 0.2% (Hoogeveen et al., 2013). In irradiance measurements, where the light is imaged as a vertical line, the ISRF at one row could be affected by stray light from other rows. Hence, these data are corrected for stray light using the operational processor. In radiance measurements, where the light is imaged as a spot, there is no stray light from other rows and the stray-light correction is not applied. The diode laser measurements are also not corrected for stray light, as they are intended to monitor stability.

Calibration key data for pixel quality have been derived from on-ground measurements. The pixel quality is based on several tests to identify pixels with too high dark current or noise. Most of the pixels with low quality exhibit high noise. About 260 pixels with a very low pixel quality ("bad pixels") are rejected from the analysis.

~~The measurement data of each wavelength scan are processed separately. Frames where the light source was off or very weak are discarded. In each remaining frame, the column with the maximum average signal is determined and the columns up to 7 pixels from this peak column are selected, to include the faint signal of the tails~~During data analysis it was found that a significant amount of irradiance measurements had to be rejected from the analysis due to detector saturation, as a result of an unexpectedly large variation of the laser intensity. From the radiance measurements partly illuminated rows were rejected.

3 Methodology

In this section, the method is described to determine the TROPOMI-SWIR ISRF from the on-ground measurements presented in the previous section. A general description is given first, followed by the details of the method.

3.1 ~~ISSF fit~~

~~The ISRF parameters cannot be retrieved directly from the measurements, because~~ The first step in the analysis is to obtain the wavelength and intensity of the signal ~~are unknown and have to be determined via the ISSF. The ISSF is assumed to be the mirrored version of an ISRF, which can be modeled with the function $A\mathcal{R}(c; d, -s, w, \eta, \gamma, m, c_0)$ using Eq. (13); only its skew parameter s has the opposite sign. In each frame, the ISSF of an illuminated row is fitted to the ISRF shape to normalize the signals and to find the wavelength peak position expressed in pixel units (Fig. 1b). The normalized signal and the distance to the fitted peak position yields one point~~ for each frame measured in each wavelength scan. As a first approximation, the ISSF shape corresponds to an image of the slit on the detector with the optics blurring the image. The mathematical model for the ISSF at this stage is a convolution of a normal distribution and a uniform distribution. Non-linear least squares minimalization is used to fit this function to the data. As a result, a wavelength can be assigned to each measurement and the signals can be normalized by the fitted signal intensity. The wavelength is expressed as a non-integer column distance. The wavelength assignment in nm is performed in the final step of the ISRF ~~shape, illustrated by the intersection of the red and green line in Fig. 1a. The normalized signal~~ determination.

The shape of an ISRF at location (r, c), where r is along the spatial dimension and c along the spectral dimension, is given by the normalized signal measured by detector pixel (r, c) as a function of ~~the fitted peak position forms a set of points covering the ISRF shape (Fig. 1c).~~

3.1 ~~ISRF fit~~

~~For a specific pixel in a row, each frame yields one ISSF fit and thus one point of the ISRF shape. While scanning the laser and recording frames, the set of points covering the ISRF shape of a given pixel is generated from the ISSF fits. As the laser-wavelength scan is not regular, the ISRF data pointsare not on a regular grid. Therefore, the~~ wavelength. Typical examples of ISRFs obtained from resp. irradiance, radiance and diode-laser measurements are presented in Figs. 3–5, taken across the SWIR detector from top left to bottom right. The locations vary slightly between the measurement types to avoid measurements affected by strong saturation effects or laser instabilities. In the upper panels, the ISRF data points are shown as black dots.

The irregular distances between the dots clearly show that the scan speed of the external laser is not constant, as explained in Sect. 2.1. This has no negative effect on the curve fitted through the data points~~in the scan range are collected in bins of 1/32 of a spectral pixel and a median is applied to the data points in each bin. Empty bins are discarded. Reducing the number of points like this also speeds up the fitting and reduces noise while no significant details arelost. The ISRF is fitted with the function $AR(c; d, s, w, \eta, \gamma, m, c_0)$ using Eq. (13).~~

~~The quality of the fit is determined by calculating the fit variance, the sum of the squared fit residuals where the fit function is larger than 6% of the maximum, divided by the number of degrees of freedom (number of points minus the free fit parameters). The square root of the fit variance is the rms value~~. Furthermore, the fit is also quite robust against single outliers and missing data points. The peak width due to the projection of the slit on the detector is constant when expressed as a wavelength interval, but expressed as a column distance it decreases towards larger columns (longer wavelengths), because the spectral dispersion changes. The peak height increases to keep an integrated area of 1. The ISRF in the panels (a_1) and (a_2) are clearly skewed, while the other three are, by eye, symmetrical. The signal-to-noise in all three measurements is sufficient to determine the signal accurately up to 4.5 pixels from the center. In the calibration key data the ISRF will be defined in this range only. The ISRF outside this range will be set to zero. Any remaining signal is considered as stray light, in line with the stray-light correction algorithm (Tol et al., 2018).

3.1 ~~ISRF parameter smoothing~~

~~The ISRF calibration key data has to be valid for the whole SWIR spectral range and for all swath angles~~The one-step approach described so far would actually result in ISRF fit residuals much larger than shown in Figs. 3–5. The residuals would show systematic oscillations with a period of one pixel. These occur when a simplified model is applied to a poorly sampled ISSF, leading to errors depending on whether the peak is at a measured point or between two measured points (see Fig. 1b). In addition, the residuals of an asymmetric ISRF would show significant left-right differences, mostly negative residuals to the left and positive residuals to the right, when a symmetric model is used in the ISSF fit.

A TROPOMI-SWIR ISSF measurement has typically 5 spectral pixels with sufficient signal-to-noise, not enough to fit all parameters of the complete ISRF model including a description for skewness and tails. However, the ISRF ~~fits are valid locally (at location (r, c)) and not available for all pixels. It is expected that the fit parameters that define the local ISRF vary only smoothly over the surface of the detector as this is determined by the spectrometer optics. Therefore, a bivariate polynomial fitting is used to smooth and to interpolate the ISRF fit parameters. The model with variables r and c, where the first row and column are mapped to -1 and the last row and column mapped to $+1$, is given by~~

$$E_{\text{fit}}[r, c; \boldsymbol{a}] = \sum_{k=0}^{M(M+3)/2} a_k T_{m-n}\left(2\frac{r}{255} - 1\right) T_n\left(2\frac{c}{999} - 1\right),$$

~~using Chebyshev polynomials of the first kind $T_n(x)$ and indices~~

$$m = \left\lfloor \tfrac{1}{2}\left(\sqrt{1+8k}-1\right)\right\rfloor,$$

$$n = k - m(m+1)/2.$$

derived with the approach so far is closer to the true ISRF than the simplified model used for the first ISSF fit. Using the mirrored shape of the ISRF at a given detector location as the shape of the ISSF at that location, the ISSF is fitted again to yield an updated wavelength and intensity as input to the determination of the ISRF (Sect. 3.2). The residuals of the resulting ISRF fit are much smaller. This procedure should be repeated until the residuals are no longer improving.

~~To obtain good results~~

So far, a general approach is presented to determine high-resolution ISRFs for a spectrometer that measures the ISSF with only a few spectral pixels, using a scanning monochromatic light source. Essential is to use the instrument itself to determine the wavelength and intensity of the light. This method enables the necessary measurements to be taken within reasonable time (days) even for detectors with more than 100,000 pixels. Next, a model is defined for the TROPOMI-SWIR ISRF and a practical implementation is shown of the iterative approach.

3.1 ISRF model

The mathematical model for the ISRF ~~parameter fitting, obvious outliers in the individual ISRF-fit results should be rejected before the bivariate polynomial fit is performed. Given the distribution of outliers (in columns at the same wavelength), it is judged that most of them are caused by laser artifacts. The following data selection was applied before the parameter fitting: no ISRF fit is performed on the first and last two pixels of an automated scan, because the whole peak of the ISRF should be present in the data; curve fit solutions were rejected in case the curve-fitting routine signaled an error. This happened for a small fraction of the pixels ($< 1\%$) , due to bad ISRF data; unrealistic curve-fit solutions are rejected:~~ should be flexible enough to represent the range of ISRF shapes adequately with the smallest number of parameters. The TROPOMI-SWIR ISRF is modeled by the weighted sum of functions for the peak and the tails. The peak function is a skew-normal ~~width $d < 0.1$ or skew $|s| > 5$; poor fits are rejected based on their rms value. Irradiance measurements: reject data of the whole column when its median rms is larger than 0.0065. Radiance measurements: reject ISRF fits with an rms larger than 0.0065. The ISRF of a given pixel can sometimes be derived multiple times, because (*i*) the automated scans overlap in wavelength; (*ii*)for irradiance measurements, all automated scans are performed twice: scanning up and down in wavelength.~~ distribution convolved with a uniform distribution. This corresponds to a (~~*iii*) for radiance measurements, successive scans overlap in~~ possibly asymmetric) image of the slit on the detector, with the optics blurring the image. Beirle et al. (2017) use an asymmetric version of the ~~swath direction by half a pixel~~exponential power distribution ('Super Gaussian'). This function is computationally less demanding and suitable for general ISRF simulations, but in our case the fit residuals are larger than with the convolution above. For the ~~parameter fits, only the ISRF fit with the smallest rms is used.~~ tails a function is needed with an adjustable tail weight. A suitable function is found to be the Pearson type VII distribution, which is a generalization of the Gauss and Lorentz distributions. It can represent the wings of SWIR spectral stray light satisfactorily (see Tol et al., 2018, Fig. 9).

~~Orders M of the bivariate Chebyshev expansions applied to the irradiance and radiance data are equal, except in cases~~ The normal distribution with mean 0 and standard deviation σ is given by

$$\mathcal{G}(c;\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{c^2}{2\sigma^2}\right). \tag{1}$$

The skew-normal distribution is a generalization including an extra skewness parameter s:

$$\mathcal{N}_1(c;\sigma,s) = \left[1 + \operatorname{erf}\left(\frac{s\,c}{\sigma\sqrt{2}}\right)\right] \mathcal{G}(c;\sigma). \tag{2}$$

This function has mean $\sigma\,\delta$ and standard deviation

$$d = \sigma\sqrt{1-\delta^2}, \tag{3}$$

with

$$\delta = \frac{\sqrt{2}\,s}{\sqrt{\pi(1+s^2)}}. \tag{4}$$

To interpret fitting results more easily, the skew-normal distribution is written in terms of d instead of σ and a parameter c_0 is included for the mean:

$$\mathcal{N}_2(c;d,s,c_0) = \mathcal{N}_1\left(c - c_0 + \frac{d\,\delta}{\sqrt{1-\delta^2}}; \frac{d}{\sqrt{1-\delta^2}}, s\right) \tag{5}$$

$$= \frac{\sqrt{\frac{1}{2}\pi + \frac{1}{1+s^2} - 1}}{\pi d}\left[1 + \operatorname{erf}\left(\frac{s\,\xi_0}{\sqrt{2}}\right)\right] \exp\left(-\tfrac{1}{2}\xi_0^2\right), \tag{6}$$

with

$$\xi_0 = \frac{\sqrt{1-\delta^2}}{d}(c - c_0) + \delta. \tag{7}$$

Peak function $\mathcal{S}$ is the convolution of skew-normal distribution $\mathcal{N}_2(c;d,s,c_0)$ and a uniform distribution with mean 0 and full width w (the 'block width'):

$$\mathcal{S}(c;d,s,w,c_0) = \frac{1}{w}\int_{c-w/2}^{c+w/2} \mathcal{N}_2(u;d,s,c_0)\,\mathrm{d}u. \tag{8}$$

This can be written as

$$\mathcal{S}(c;d,s,w,c_0) = \frac{\operatorname{erf}\left(\xi_+/\sqrt{2}\right) - \operatorname{erf}\left(\xi_-/\sqrt{2}\right)}{2w} - 2\frac{T(\xi_+,s) - T(\xi_-,s)}{w} \tag{9}$$

with

$$\xi_{\pm} = \frac{\sqrt{1-\delta^2}}{d}(c - c_0 \pm \tfrac{w}{2}) + \delta \tag{10}$$

and using Owen's T function (Patefield and Tandy, 2000)

$$T(z,s) = \frac{1}{2\pi}\int_0^s \frac{\exp\left(-\frac{1}{2}z^2(1+t^2)\right)}{1+t^2}\,\mathrm{d}t. \tag{11}$$

Tail function $\mathcal{P}_7$ is the Pearson type VII distribution

$$\mathcal{P}_7(c;\gamma,m,c_0) = \frac{\Gamma(m)}{\gamma\sqrt{\pi}\Gamma(m-\frac{1}{2})}\left(1 + \frac{(c-c_0)^2}{\gamma^2}\right)^{-m} \tag{12}$$

with $m > 1/2$ and $\gamma > 0$. This distribution is a generalization of the Lorentz distribution where the tail shape can be changed; the specific case $\mathcal{P}_7(c;\gamma,1,c_0)$ is the Lorentz distribution with half width at half-maximum γ. ISRF function $\mathcal{R}(c;d,s,w,\eta,\gamma,m,c_0)$ consists of a peak function with three shape parameters d, s ~~. The irradiance data has a better coverage in both spectral and spatial directions, so a higher order $M = 7$ could be applied on the parameters d and s, which show much more structure than the other fit parameters. An order $M = 6$ was used on the radiance data. Order M used for~~ and w, and a tail function with two shape parameters γ and ~~η are 5, 2 and 2, respectively. The result of this step yields the~~ m:

$$\begin{aligned}\mathcal{R}(c;d,s,w,\eta,\gamma,m,c_0) =& (1-\eta)\mathcal{S}(c;d,s,w,c_0)\\ &+\eta\mathcal{P}_7(c;\gamma,m,c_0).\end{aligned} \tag{13}$$

The mean is at c_0, the integral is 1 and the integral over the tail part only is η.

The requirement on the TROPOMI-SWIR ISRF states that the ISRF ~~parameter models that are employed to calculate the ISRF calibration key data for each and every pixel. These key data are used in the SWIR retrieval algorithms that derive the gas columns.~~

~~The quality of~~ should be known with an accuracy of 1% of its maximum. A stringent implementation of the requirement would be that the ~~parameter fitting is determined by comparing the measured~~ absolute value of all residuals should be less than 1% of the maximum of the ISRF (about 0.004), but that leaves no room for outliers. An alternative measure of the fit quality turns out to be an rms value calculated as the square root of the sum of the squared difference between the ISRF fit residuals, using points where the fit function is larger than 6% of its maximum, divided by the number of ISRF data points ~~with the ISRF that results from the parameter model. An $\mathrm{rms}_{\mathrm{model}}$ value is determined in a similar way as the rms defined in Sect. **??**. In general, the parameter smoothing will result in better and smoother ISRF calibration key data due to averaging and interpolation. Possibly counter intuitive, the rms value will be slightly larger as the ISRF data points are now compared with a smoothed ISRF instead of an optimized local ISRF that might be influenced by measurement imperfections~~minus the number of free fit parameters. The threshold of 6% is arbitrary, but a lower value would include more of the tails where the residuals

are always very small, which would make this measure less sensitive. The advantage of this measure is that it is less sensitive to small outliers, and sensitive to large outliers which can corrupt the fit procedure. Therefore, we use this rms as a measure of the fit quality.

3.2 ISRF parameter iteration

An ISSF has typically only 4-5 spectral pixels with sufficient signal-to-noise ratio. Therefore, only fitting with a few parameters is accurate, yielding significant errors in the resulting ISRF. An iterative approach has been developed, starting with a symmetric model without tails. Once the ISRF has been fitted, the skew and tails are known approximately, and can be included as fixed properties while the ISSF is fitted again. The ISRFfit also benefits from a reduction of the number of parameters. Therefore, the refitted block width as a function of row and column is smoothed and used as a fixed property in the final ISRFfit. In the end, the ISSF and ISRF are determined in four passes or 'stages'. The differences between the parameter fitting in ISSF In the iterative approach as introduced above we start with a simplified model for the ISSF to determine the central wavelength and intensity of the laser. Many ISSF fits are used to estimate the (local) ISRF, so the ISRF can be determined with many parameters. For the next iterations we use the fact that the ISSF is equal to the mirrored local ISRF, and ISRF fits in the different stages are summarized in Table **??**the only free parameters in the successive iterations in the ISSF fit are the wavelength and the intensity.

The procedure and number of stages This procedure and the number of iterations have been verified using synthetic ISRF data. To illustrate the ISRF parameter iteration, we present the results of simulations using two realistic ISRFs, one nearly-symmetric and the other skewed. The calculations are performed without noise, and the ISRF is kept constant in each data set. A simulated measurement data set was constructed given the shape of About 2000 synthetic ISRFs have been constructed using the ISRF model (Eq. 13) and combinations of the 7 parameters covering most of the ISRF and assuming a laser scanning over about 20 spectral pixels in small steps (100 per spectral pixel) similar to the actual calibration measurements. Table 1 lists the intermediate results of the ISRF fit at the end of each stage, while Fig. **??** shows the parameter space. This data set has been generated with and without realistic noise. The simulations confirmed our guess that not all ISRF parameters can be fitted at the same time, because the tail parameters γ and m are not independent and the fit of these parameters is further complicated by the small signal of the tails. In early computations to determine the TROPOMI-SWIR ISRF, the tail parameter m was fixed to 1.2, based on the shape of the tails found in stray-light measurements. The simulations with synthetic data shows it is better not to fix m, because its true value is poorly known and it improves the convergence of the ISRF model towards the true ISRF.

The ISRF parameters of the skewed model are derived with high accuracy (better than 0.5%) after 3 stages. In contrast, other parameters. The best convergence towards the true ISRF, both in speed and in accuracy, is achieved by fixing peak width parameter w. However, in this case w needs to be known better than 1%, which is unrealistic. Second best is to fix tail fraction η. The tail fraction is nearly constant across the detector and from detailed analysis of the ISRF determined after the first iteration its value is usually be between 0.1 and 0.12. Therefore, assuming a tail fraction of 0.11 would work for most ISRF fitting. To overcome problems in case the parameter iteration has problems deriving the ISRF tail parametersfor the nearly symmetric ISRF, where the tail width is overestimated by more than 20%and the tail fraction is underestimated by 25%. The

curve-fitting routine is not the problem, because Fig. **??**c shows that the fit residuals are almost zero after stage 3. Apparently, in the case of a symmetric function the algorithm converges to a sub-optimal solution. In this case the algorithm has difficulty separating the nearly symmetric peak from the also symmetric tail function. However, true tail fraction is outside this range, the ISRF fit is performed twice. First only fixing the tail fraction (from the previous iteration), then only fixing w (improved guess from the previous ISRF fit). We refer to each repeated step of the method as "stages", because each step contains several computations: one ISSF fit and two ISRF fits.

This approach generates consistent good results after 4 stages. Synthetic data show that the differences between the true determined ISRF and the derived ISRF are less than 0.25% and are considered acceptable. The algorithm converges in two stages, see Fig. **??**b and **??**d. Any further iterations will only marginally improve the results. This is true for noiseless synthetic data, and it should be noted that performing stage 3 does improve the derived ISRF parameters of real measurements. Stage 4 is added because the block width is fixed to its model (derived in the previous stage) in the ISRF fit. This turns out to reduce the variance in the tail parameters significantlytrue ISRF are within 0.0005 (about 0.125% of the ISRF maximum). Adding realistic noise has a very minor effect, due to the large number of measured ISSF samples: the residuals are doubled and the determined ISRF is more slightly symmetrical. Because the true ISRF is generated with the same ISRF model used for the fit, one would expect that the determined ISRF matches the true ISRF perfectly. However, this is not the case due to the fact that details of the true ISRF are lost in the fit of the poorly sampled ISSF. According to an extra simulation, the determined ISRF matches the true ISRF nearly perfectly when the ISSF signal is measured with 10 instead of 5 spectral pixels.

4 Discussion of results

Typical examples of various shapes of the TROPOMI-SWIR ISRF are The convergence of the method is illustrated in Fig. **??** (irradiance measurements) and Fig. **??** (radiance measurements). Shown are the stage 4 results of the ISRF for five columns: 128, 300, 5406, where the ISRF shapes presented in Figs. 3–5 are simulated. For all five simulated ISRFs a quick convergence is achieved. The large residuals between the asymmetric ISRF after stage 1 and true ISRF are due to a shift in the peak position, while the shape of the final ISRF agrees with the true ISRF, 726 and 935, respectively. A median has been taken over all rows illuminated. From visual inspection of the displayed ISRFs, one can conclude that: (*i*) the ISRF is sharper and higher at higher column number (longer wavelength); (*ii*) the ISRF fit resembles the ISRF data very well, e. g. the residuals are very small, except where small artifacts can be identified see Table 1.

3.1 ISRF parameter smoothing

The goal of the on-ground calibration is to determine the SWIR ISRF for all pixels, because trace-gas retrieval needs the ISRF for the whole SWIR spectral range and for all swath angles. However, the ISRF could not be measured for all pixels due to measurements with too strong laser signal (rejected before analysis), bad pixels and laser instability. Therefore, it is necessary to interpolate the local ISRFs. Minor problems with the OPO are still present in the ISRF data ; (*iii*)the fit residuals of the irradiance ISRF are nearly a factor 2 smaller compared with the radiance ISRF.

The Pearson VII distribution has two parametersto control its shape: m and γ. When both are free parameters, the median value of m in the successful ISRF fits is 1.65 for radiance measurements and 1.25 for irradiance measurements. The difference is likely due to differences in stray light in these measurements. In all subsequent fitting, shape parameter m is fixed to 1.25 to enhance convergence of the curve-fitting routine and inter-comparison between ISRFs derived from irradiance and radiance measurements. It has to be noted that the contribution of points (see Figs. 3–5), which may affect the ISRF fit yielding minor deviations of the fit parameters. Assuming that these deviations are random and that the shape of the tail to ISRF varies only smoothly over the surface of the detector (as it is determined by the spectrometer optics), then the quality of the ISRF would benefit when the ISRF parameters are smoothed and interpolated using bivariate polynomial fitting. Any high-frequency detector features still present in the data due to imperfect calibration or local changes in the detector PSF should then be visible in the ISRF is small ($\leq 10\,\%$) and only significant 1-2 pixels away from the peak. It has been verified that fixing the m parameter has negligible effect on the resulting ISRF and the fit residuals expressed in the rms valuedifference between the local and smoothed parameter values.

The selected bivariate polynomial fit uses Chebyshev polynomials of the first kind T_p to minimizes the problem of Runge's phenomenon. The model at a location specified by row r and column c, where the first row and column are mapped to -1 and the last row and column are mapped to $+1$ is given by

$$E_{\text{fit}}[r, c; \boldsymbol{a}] = \sum_{m=0}^{M} \sum_{n=0}^{m} a_{mn}\, T_{m-n}\left(2\frac{r}{n_{\text{row}}-1} - 1\right) T_n\left(2\frac{c}{n_{\text{col}}-1} - 1\right), \tag{14}$$

where $n_{\text{row}} = 256$ and $n_{\text{col}} = 1000$ for the SWIR detector. The order M is the maximum of the sum of the exponents.

Table ?? summarizes the irradiance ISRF characterization. The number of good fits increases with every stage, except for the final stage where, in the ISRF fitThe bivariate polynomial fitting is sensitive to obvious outliers. Therefore, ISRF fits have been rejected when the fit quality is low: rms > 0.003, skew parameter $|s| > 5$, tail parameter γ is outside the range $[0, 3]$ or tail parameter m is outside the range $[0.5, 3]$. The order M of the bivariate polynomial fit is optimized for each ISRF parameter to minimize the variance of the difference between the raw values and the fitted value. An order $M = 6$ is used for skew parameter s, the block width is fixed to the value obtained in the previous stage. The goal of this final stage is , instead, to increase the quality of the tail fitting. This indeed happens as the standard deviation in the tail parameters is reduced by 30–50 %$M = 4$ for peak parameters d and w, and $M = 2$ for tail parameters γ and m.

The final results of the irradiance ISRF are presented in the appendix of this article, shown are smoothed ISRF parameters are employed to calculate the ISRF calibration key data for each and every pixel. Residuals are examined to check whether the excluded local variations are small enough to ignore. Actually, the smoothed ISRF parameters are calculated at the end of each stage, because the ISSF fit of the next stage, in general, benefits when erroneous ISRF fit results are not propagated to the next stage.

4 Discussion of results

Of the 2000 simulations with synthetic data performed, the ~~value of each ISRF parameters from the ISRF fit and after the bivariate polynomial fitting, see Figs. 9~~results of those that closely resemble the measured ISRF examples as shown in Figs. 3–~~??~~. ~~As an example, the skew-normal width is shown in : Fig. 8a its fitted value, Fig. 8c the bivariate polynomial fit result, and in Fig. 8b the difference between the two. The stripe pattern in Fig. 8a is due to scanning imperfections of the laser. They are removed and interpolated in Fig. 8c, but reappear in the difference plot of~~ 5 will be used for illustration. The convergence towards a true ISRF from synthetic measurements is presented in Fig. ~~8b. Conclusions for all fitting parameters are :Block width w of the ISRF is determined by the projection of the slit onto the detector and therefore decreases as a function of wavelength~~6. These simulations are performed with an ideal laser: constant wavelength scan speed (using the average 0.00125 nm/s from the measurements) and constant laser signal. As expected, ~~no variation is seen over the spatial dimension of the detector (swath angle). Skew-normal width d shows more fine-scale structure than the other parameters. The pattern is probably due to variations in the sensitive area of the detector as the pattern is also seen in sensitivity plots of the detector alone (Hoogeveen et al., 2013). The fine-scale structures are smoothed in the parameter fit. This only has a minor effect on the quality of the ISRF calibration key data, as will be discussed later in this section. Skew parameter s shows that~~ the fit residual and the deviation from the true ISRF is still significant after the first stage, but the procedure quickly converges in subsequent stages. The final residuals and differences with respect to the true ISRF are much smaller than the requirement on the ISRF ~~is positively skewed ($s > 2$) at shorter wavelengths, and negatively skewed at longer wavelengths, but with a gradient along the swath. It appears the algorithm excludes s values between -0.6 and $+0.6$ as the difference between these curves is only very small ($< 10^{-3}$) . The parameter fit of s will interpolate this gap, introducing an error in the parameter s. However, width parameter d has been designed such that no errors are introduced by the ISRF parameter fit . Tail parameters η~~ knowledge. Table 1 shows that all ISRF parameters converge towards their true values during the ISRF parameter iteration.

The local ISRFs determined from irradiance measurements (Fig. 3c) and ~~γ show little variation over the swath, but their values can vary greatly from one column to the next. The tails have a low signal (< 0.01) outside the peak (see Figs. ?? and ??) . The low signal hampers the least-square minimization algorithm, leading to significant variation in the tail parameters. The bivariate fitting reduces the variation over swath and wavelengths.~~ radiance measurements (Fig. 4c) have typical fit residuals smaller then 0.004, except for a few outliers. These outliers correspond with the presence of (small irregular) wavelength jumps during the measurements. Some of the fit residuals show systematic features, but they are not consistent between both datasets. Therefore, the ISRF model (Eq. 13) is considered a good representation of the ISRF shape. The fit residuals determined from the on-board diode-laser measurements (Fig. 5c) are less noisy, because the speed of the wavelength scan was much lower and the diode lasers were better behaved.

~~Figure ??a shows the color-coded rms~~

Several fit thresholds are introduced in section 3.1, to optimize the bivariate polynomial fitting of the ISRF ~~fits. The pixels classified as good, i. e. meeting the requirement for ISRF knowledge, have an rms residual smaller than 0.004,equal to 1~~parameters. Low quality fits are identified by a high rms value. The median rms value is about 0.0017 for irradiance measurements, hence a reasonable threshold value is set at 0.003. Saturated measurements are rejected from analysis, but some

ISRF fits are still affected by nearby saturation. Therefore, ISRF fits with an exceptionally large skew-normal width value are rejected.

Of 211,575 pixels of the SWIR detector in the operational wavelength and swath ranges, we have ISRF data measured through the irradiance port of 195,892 pixels. For 7 % of the ~~maximum value of the ISRF. The bands with bad or moderate fits are attributed to laser artifacts,except for the columns at both ends, which are not scanned by the laser. The little spots of just a few pixels are caused by bad pixels. The quality of the ISRF fits as determined with the parameters from the bivariate parameter-fitting models~~ pixels ISRF data was not available or rejected due to possible saturation. The ISRF parameter smoothing at the end of stage 1 is based on about 116,000 local ISRF, because 73,616 ISRF fits were rejected due to the rms condition and about 6000 due to too large skew and skew-normal width. The number of used ISRF fits increased per stage to about 164,000 at stage 4, with only 28,900 rejected due to the rms condition and about 2,400 due to too large skew and skew-normal width. The rms values of the irradiance ISRF are shown in Fig. ~~??b. Here the regions with moderate or bad quality fits are extensions of moderate and bad regions in Fig. ??a. There are a few small regions which coincide with the fine-scale structures visible in the~~ 7a.

The number for the radiance measurements are significantly lower: ISRF data is available for 150,000 pixels. For 30 % of the pixels ISRF data was not available or rejected due to partial illumination of the pixels by the laser spot. The ISRF parameter smoothing at the end of stage 1 is based on about 63,000 local ISRFs, because 84,300 ISRF fits were rejected due to the rms condition and about 3000 due to too large skew and skew-normal width. The number of used ISRF fits increased per stage to about 92,500 at stage 4, with only 51~~see for example around row 50 at columns 525 and 610. However, regions with good quality ISRF parameter fits are distributed over the entire array,and the~~ 300 rejected due to the rms condition and about 2,000 due to too large skew and skew-normal width. The rms values of the ~~ISRF parameters in these regions are consistent with the overall trend.This gives us confidence that our SWIR ISRF characterization is successful, and within the requirement on the knowledge of the ISRF .~~ radiance ISRF are shown in Fig. 7b.

~~In general,~~ The difference between the ISRF data points and the ~~laser performed worse during the radiance measurements, yielding radiance ISRF measurements of poorer quality than the irradiance measurements. This can be seen in Fig. ??, which shows in general higher rms values for the ISRF fits for the radiance measurements~~ISRF model are very small, as the median of the rms is 0.0015 for both irradiance and radiance measurements, but the coverage of the first is much better. Therefore, the ~~radiance ISRF~~ smoothed ISRF derived from irradiance measurements are used ~~for validation of the SWIR~~ to generate the TROPOMI-SWIR ISRF calibration key data~~obtained using the irradiance ISRF measurements.~~. The ISRF model parameters are presented in Figs. 8–13. The skew-normal width parameter is sensitive to the photo-sensitivity of the SWIR detector (Fig. 6 of Hoogeveen et al., 2013). These patterns are easily recognizable in the residual plot (Fig. 8b), where the difference is shown between the ISRF parameter d and its bivariate poynomial fit.

~~For the comparison of the irradiance ISRF with the radiance ISRF, swath angles with the best spectral coverage and quality are selected.From the radiance and irradiance data the rows at ranges 40–72 and 110–140 are selected. Figure ?? shows the median of the~~ Figure 14 shows the differences between the smoothed ISRF and local ISRFs determined from irradiance, radiance and diode-laser measurements at the five locations presented in Figs. 3–5. The fit parameters are compared in Table 2.

The differences between the ISRF shape parameters as a function of column (before bivariate parameter fitting). In general, the spread in the irradiance data is larger. This is due to the fact that all irradiancedata of a column are taken in a single laser scan, while the radiance data of a column is taken from about 100 laser scans, thus averaging out some of the laser irregularities. The skew-normal width and block width show very good agreement at the right side of the detector, including the wiggles which are also visible in Fig. 5a. On the left side of the detector, the block width of the radiance ISRF tends to be smaller than that of the irradiance ISRF . This subtle difference is attributed to the non-optimal scanning of the laser at these wavelengths. The skew parameter tends to be slightly lower for the radiance measurements . However, the impact data points and the smoothed ISRF are for most measurements within the requirement on the ISRFof this difference is almost negligible. The tail parameters show reasonable resemblance in the two data sets, given the noise in both parameters.

In conclusion, the differences between the measured radiance and irradiance ISRFare too small to conclude that there is a significant difference between the two. More likely, differences are caused by imperfections of the laser-scan behavior or due to subtle differences in residual stray light. This justifies the use of the irradiance measurements . In general, all local ISRF fits with rms values less than 0.002 in Fig. 7 have also small residuals against the smoothed ISRF. Therefore, we conclude that smoothed ISRF, used to derive the ISRF calibration key data. Moreover, the radiance measurements effectively provide an independent validation of the derived ISRFs. CKD, agrees well with the local ISRF data.

5 In-flight Monitoring monitoring of ISRF

As knowledge of the ISRF is critical for the science results of the SWIR band, it has been decided to include means to monitor identify possible changes in the ISRF between the on-ground calibration campaign and the first measurements in space, and to monitor the ISRF stability during the 7 years of operational lifetime. For this, five tunable diode lasers in DFB diode lasers, spread evenly over the SWIR wavelength range, are included in the on-board calibration unit. Roughly once per month, the ISRF will be monitored using each laser. The laser wavelengths are scanned by tuning the temperature of the laser using a built-in thermo-electric cooler. The scanning range is about 6 spectral pixels so that the ISRF can be monitored for one or two wavelength pixels per laser. The Technical details of the on-board laser illuminates the SWIR spectrometer via a dedicated diffuser. As the diffuser is not moved during the measurements, there will be speckle.Most speckle is removed by taking the median of the data of all illuminated rows.

During the lasers are given in Sect. 2.1. During the in-flight commissioning phase, in-flight the on-ground measurements with the on-board lasers will be performed repeated with a moving and a fixed diffuser. The ISRF obtained from these measurements diffuser. Hence, these ISRFs can be comparedwith the ISRF measured on ground using the external laser and the on-board diode lasers to detect any possible changes. The early in-flight measurements also act as a reference "reference ISRF" for the ISRF monitoring. The monitoring ISRF is of sufficient quality to check for any degradation of the instrument but cannot be applied in

During the operational phase, dedicated measurements are planned to monitor ISRF stability. These measurements will be performed with each laser, roughly once per month with a fixed diffuser, because the diffuser mechanism is a life-limited item.

The ISRF determined from these measurements ("monitoring ISRF") is less accurate as it is affected by laser speckle patterns. It is used for monitoring only, not for trace-gas ~~retrieval. Should it be necessary, the~~ retrieval. If recalibration of the ISRF is necessary, then the wavelength scan of the on-board ~~diode lasers can be used to recalibrate the ISRF for a significant part of the SWIR band.~~

~~With an oscillating diffuser, the ISSF and pixel ISRF are determined in four stages, as described in Sect. 3, except that ISRF parameter smoothing (Sect. ??) is calculated from the ISRF fits of the few columns scanned per diode laser. The column dependence of the shape parameters is neglected and the row dependence is smoothed by a second order polynomial~~diode lasers has to be maximized and a moving diffuser should be used.

The monitoring ISRF ~~for each of the five lasers~~ is determined with the algorithm presented in Sect. 3 without iterations. It starts with the ISSF fit of stage 4 ~~where~~ using the parameters of a reference ISRF~~is used based on smoothed parameters of the given row. Thenthe median ISRF is~~, determined earlier in the mission with the same diode laser. Then, for each illuminated row, ISRF data of several spectral pixels are combined to generate one ISRF up to 5 pixels from its center. Speckle patterns are reduced by a median calculated from all ISRF data ~~of the central one/two fully-scanned columns, neglecting any row dependence.~~

~~The ISRF determination using the diode laser with and without moving on-board diffuser has been tried during the ground test and calibration campaigns. The ISRF measured with the diode lasers is in close agreement with the ISRF calibration data, thus proving the usability of the method and validating the calibration data. The monitoring ISRF deviates from the ISRF calibration data as could be expected. However, it is believed the~~ points. This monitoring ISRF will be compared with a monitoring ISRF obtained early in the commissioning phase and used in trend analyses. It is expected that this method is sensitive enough to be used ~~on board~~ for long-term monitoring, and being able to distinguish between changes in the real instrument ISRF and changes in the speckle pattern.

6 Conclusions

~~A new and accurate method~~ An approach is presented to determine high-resolution ISRFs for a spectrometer that measures the ISSF with only a few spectral pixels, using a scanning ~~OPO has been developed and applied to characterize the TROPOMI-SWIR ISRF.~~ monochromatic light source. The instrument itself is used to determine the wavelength and intensity of the light, which makes it possible to perform the necessary measurements within a reasonable time (days) even for detectors with more than 100,000 pixels.

The wavelength and intensity of the signal are determined with a fit to the measurements, but the model used for this ISSF fit is decisive for the accuracy of the ISRF determination, as is shown using simulations with synthetic ISRF data. Based on the simulations, an iterative approach is developed to improve the ISSF model from a simple model in the first iteration to the mirrored ISRF model in later iterations. The simulations show satisfactory convergence to the true ISRF in 4 iterations.

The ~~ISRF characterization has been performed on the basis of stray-light corrected irradiance measurements. The~~ TROPOMI-SWIR ISRF is modeled by the weighted sum of functions for the peak and the tails. The peak function is a skew-normal distribution

convolved with a ~~block distribution, and the tail function is a~~ uniform distribution. This corresponds to a (possibly asymmetric) image of the slit on the detector, with the optics blurring the image. For the tails a function is needed with an adjustable tail shape. A suitable function is the Pearson type VII distribution~~. An iterative scheme to derive the SWIR ISRF has been developed, where the ISRF determined in a previous iteration is used to improve the ISSF model in the current iteration. The required accuracy of the ISRF is obtained within 4 iterations~~, which is a generalization of the Gauss and Lorentz distributions. It can represent the wings of SWIR spectral stray light satisfactorily. Each of the five ISRF shape parameters ~~has been~~ are smoothed by fitting a bivariate ~~Chebyshev expansion~~ polynomial to derive the ISRF calibration key data for all SWIR wavelengths and swath angles.

The ISRF measured through the irradiance port using the solar diffuser has been compared with the equivalent ISRF measured via the radiance port. The differences ~~between the ISRFs derived from both data sets~~ are very small, and ~~largely due to differences in stray-light treatment and laser scan imperfections. The derived~~ are mostly due to measurement details, not instrument details. Calibration key data for the ISRF have been derived from the larger irradiance data set. The determined ISRF meets the requirement on ISRF knowledge and should thus be sufficient for ~~methane~~ SWIR trace-gas retrievals.

The on-board calibration unit contains five diode lasers in the SWIR wavelength range. ~~These diode lasers will be employed to verify~~ Accurate measurements with these diode lasers before and after launch will reveal whether the ISRF remained stable and the ISRF calibration key data ~~after launch~~can be applied for data retrieval. During operations, the lasers will be used to monitor the long-term stability of the optical properties of the SWIR module.

7 Data availability

The underlying data of the figures presented in this publication can be found at ftp://ftp.sron.nl/open-access-data/richardh.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors would like to thank the teams of Airbus Defence and Space Netherlands and KNMI for organizing the calibration campaign and in particular the operators for the tireless data acquisition. The authors would also like to thank the anonymous referees for their thorough reviews of the original manuscript.

TROPOMI is a collaboration between Airbus Defence and Space Netherlands, KNMI, SRON and TNO, on behalf of NSO and ESA. Airbus Defence and Space Netherlands is the main contractor for the design, building and testing of the instrument. KNMI and SRON are the principal investigator institutes for the instrument. TROPOMI is funded by the following ministries of the Dutch government: the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, and the Ministry of Infrastructure and the Environment.

References

Beers, T. C., Flynn, K., and Gebhardt, K.: Measures of Location and Scale for Velocities in Clusters of Galaxies—A Robust Approach, The Astronomical Journal, 100, 32–46, https://doi.org/10.1086/115487, 1990.

Beirle, S., Lampel, J., Lerot, C., Sihler, H., and Wagner, T.: Parameterizing the instrument spectral response function and its changes by a super-Gaussian and its derivatives, Atmos. Meas. Tech., 10, 581–598, https://doi.org/10.5194/amt-10-581-2017, 2017.

Buscaglione, F.: GMES Sentinel-5 Precursor — S5p System Requirement Document (SRD), S5p-RS-ESA-SY-0002, ESA, issue 4.1, 2011.

Day, J. O., O'Dell, C. W., Pollock, R., Bruegge, C. J., Rider, D., Crisp, D., and Miller, C. E.: Preflight Spectral Calibration of the Orbiting Carbon Observatory, IEEE Transactions on Geoscience and Remote Sensing, 49, 2793–2801, https://doi.org/10.1109/TGRS.2011.2107745, 2011.

Dobber, M. R., Dirksen, R. J., Levelt, P. F., van den Oord, G. H. J., Voors, R. H. M., Kleipool, Q., Jaross, G., Kowalewski, M., Hilsenrath, E., Leppelmeier, G. W., de Vries, J., Dierssen, W., and Rozemeijer, N. C.: Ozone Monitoring Instrument Calibration, IEEE Transactions on Geoscience and Remote Sensing, 44, 1209–1238, https://doi.org/10.1109/TGRS.2006.869987, 2006.

Hoogeveen, R. W. M., Voors, R., Robbins, M. S., Tol, P. J. J., and Ivanov, T. I.: Characterization results of the TROPOMI Short Wave InfraRed detector, Proceedings of SPIE, 8889, 888–913, https://doi.org/10.1117/12.2028759, 2013.

Hu, H., Hasekamp, O., Butz, A., Galli, A., Landgraf, J., aan de Brugh, J., Borsdorff, T., Scheepmaker, R., and Aben, I.: The operational methane retrieval algorithm for TROPOMI, AMT, 9, 5423–5440, 2016.

Kleipool, Q., Babić, L., Bartstra, R., Braak, R., Dierssen, W., Dewitte, P.-J., Kenter, P., Leloux, J., Loots, E., Ludewig, A., Meijering, P., van der Plas, E., Rozemeijer, N., Schepers, D., Schiavini, D., Smeets, J., Vacanti, G., and Vonk, F.: Pre-launch calibration status of the TROPOMI payload on-board the Sentinel 5 precursor satellite, submitted to Atmos. Meas. Tech., 2018.

Patefield, M. and Tandy, D.: Fast and Accurate Calculation of Owen's T Function, Journal of Statistical Software, 5, 1–25, http://people.sc.fsu.edu/~jburkardt/f_src/owens/owens.html, 2000.

Schrijver, H., Gloudemans, A. M. S., Frankenberg, C., and Aben, .: Water vapour total columns from SCIAMACHY spectra in the 2.36 μm window, Atmos. Meas. Tech., 2, 561–571, 2009.

Sun, K., Liu, X., Nowlan, C. R., Cai, Z., Chance, K., Frankenberg, C., Lee, R. A. M., Pollock, R., Rosenberg, R., and Crisp, D.: Characterization of the OCO-2 instrument line shape functions using on-orbit solar measurements, Atmos. Meas. Tech., 10, 939–953, https://doi.org/10.5194/amt-10-939-2017, 2017.

Tol, P. J. J., van Kempen, T. A., van Hees, R. M., Krijger, M., Cadot, S., Aben, I., Persijn, S. T., and Hoogeveen, R. W. M.: Characterization and correction of stray light in TROPOMI-SWIR, submitted to Atmos. Meas. Tech., 2018.

van Amerongen, A., Krol, H., Grèzes-Besset, C., Coppens, T., Bhatti, I., Lobb, D., Hardenbol, B., and Hoogeveen, R. W. M.: State of the art in silicon immersed gratings for space, Proceedings of the International Conference on Space Optics, 2012.

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H. J., De Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R. W. M., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70–83, https://doi.org/10.1016/j.rse.2011.09.027, 2012.

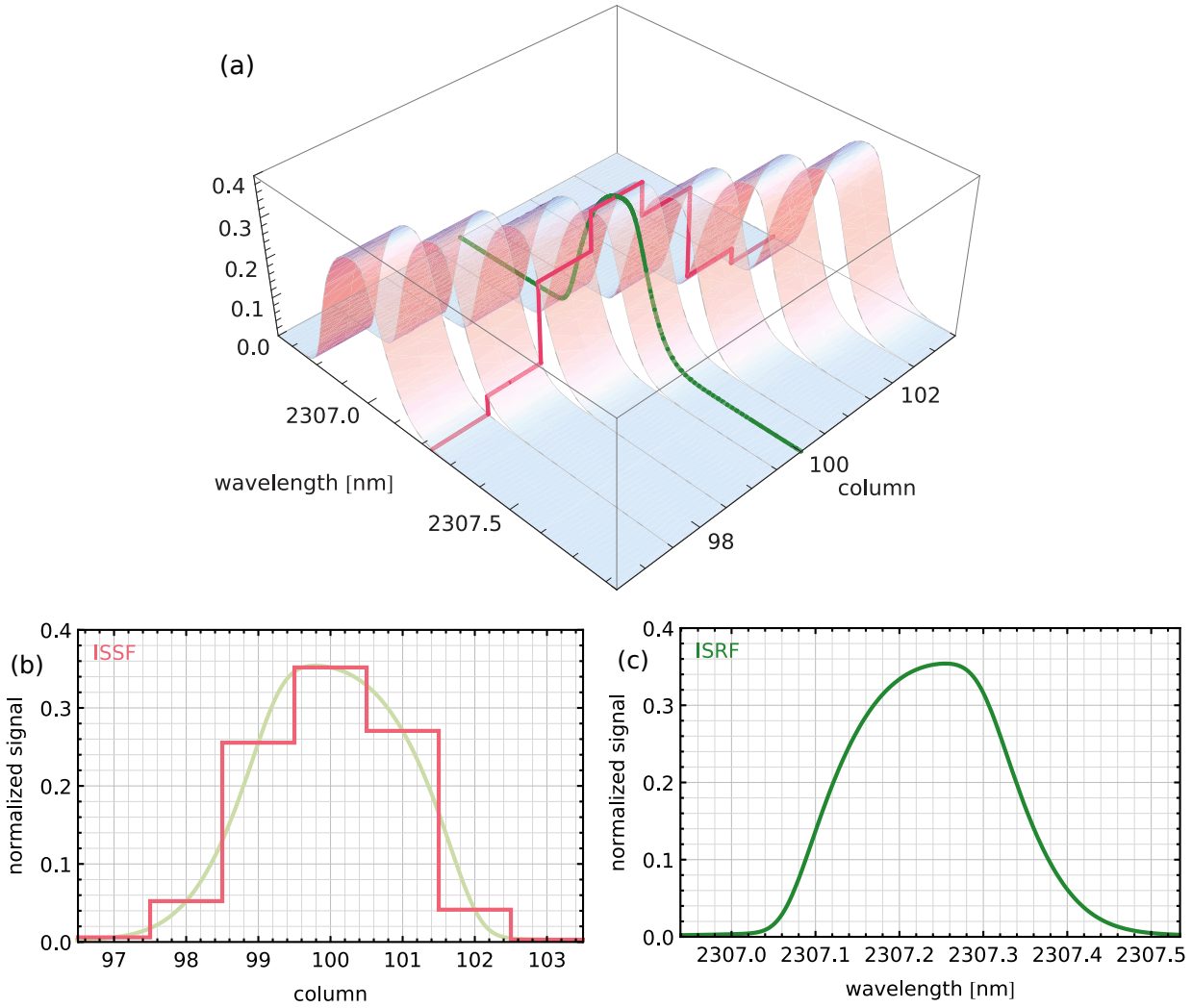

Figure 1. Normalized signal as a function of source wavelength and pixel on an arbitrary row, with two cross sections: the ISSF at 2307.24 nm (red) and the ISRF of the pixel in column 100 (green). In the plot of the ISSF, a mirrored version of the ISRF is shown in light green. The skew of the ISRF has been exaggerated to show the mirroring.

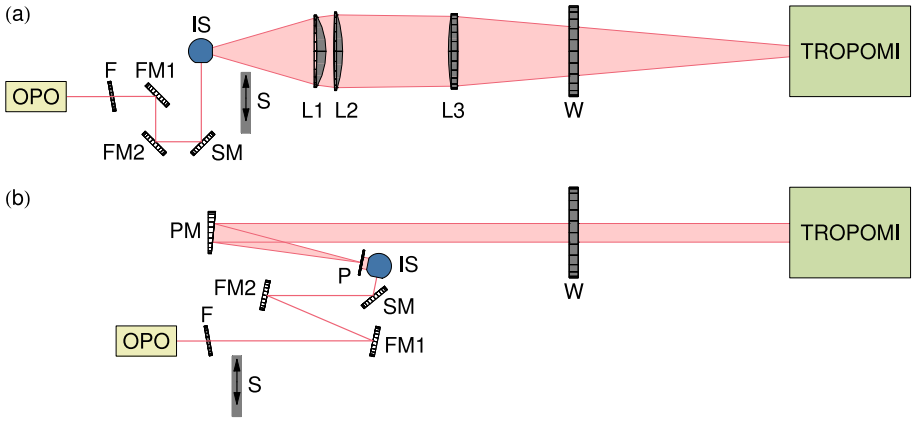

Figure 2. ~~Convergence of the ISRF model towards a synthetic ISRF using the ISRF parameter iteration.~~ The ~~results~~ setup for ~~the 'skewed' synthetic ISRF are shown in panels~~ (a) irradiance measurements and (b) ~~, where panel (a) shows the relative difference between~~ radiance measurements. The elements after the ~~ISRF data~~ OPO are: neutral density filter F, folding mirrors FM1 and ~~the ISRF fit~~FM2, spinning mirror SM, integrating sphere IS, shutter S, lenses L1, L2 and ~~panel (b) shows the relative difference between the ISRF fit~~ L3, field stop P, parabolic mirror PM and window W of the ~~'true' ISRF~~vacuum chamber containing TROPOMI. ~~Panel~~ The light enters (~~e~~a) ~~and~~ the Sun port or (~~d~~b) ~~show~~ the ~~results for~~ Earth port of the ~~'symmetric' synthetic ISRF~~instrument.

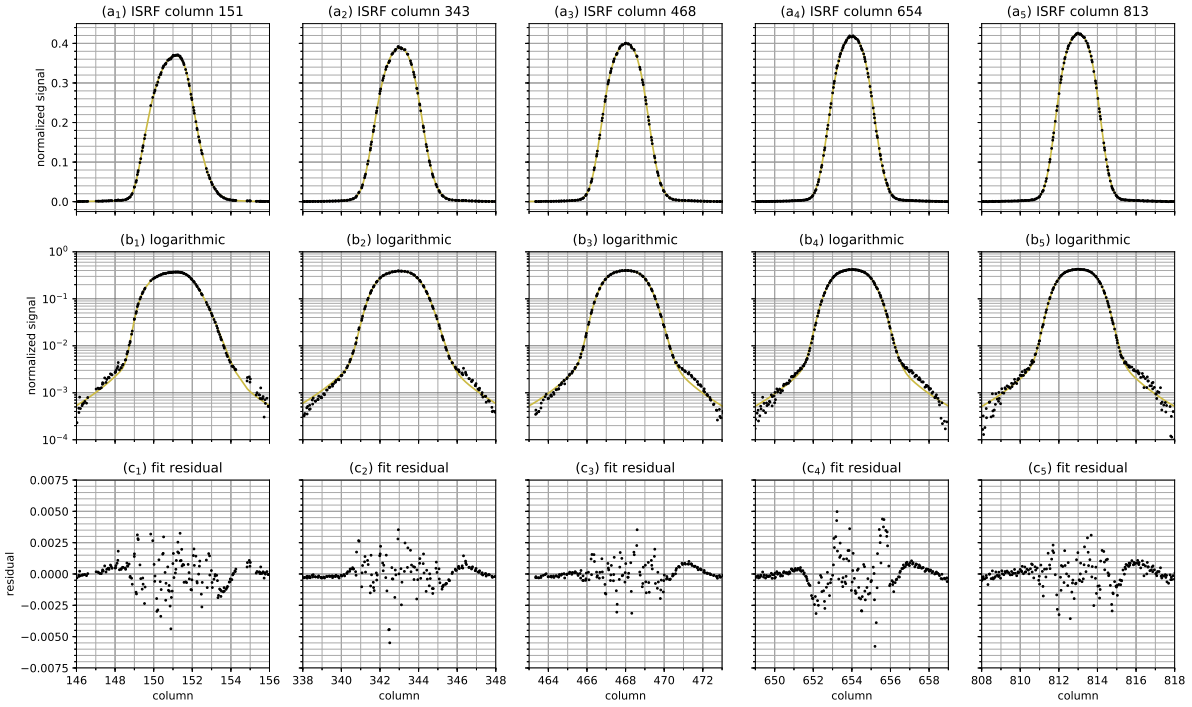

Figure 3. ~~Examples of~~ Five typical SWIR ISRFs determined from on-ground measurements with an external laser through the ~~median~~ TROPOMI irradiance ~~ISRF data with fits~~port. The ~~ISRF fit parameter values~~ upper panels (a) show the shapes of the ~~columns 128, 300, 540, 726 and 935 are respectively~~ISRF at location: ~~(skew-normal width) 0.575, 0.509, 0.444, 0.432 and 0.425; (skew) 2.880~~columns 151, ~~2.062~~343, ~~0.825~~468, ~~0.334~~ 654 and ~~-0.689; (block width) 2.669, 2.594~~813, ~~2.461~~resp., ~~2.388~~ and ~~2.301; (tail fraction) 0.081~~rows 24, ~~0.108~~76, ~~0.101~~118, ~~0.103~~ 155 and ~~0.097; (tail shape) 1.25; (tail width) 1.158, 0.988, 0.972~~191, ~~0.952 and 1.010~~resp. The ~~plots~~ middle panels (b) show the same data ~~and fits~~ on a ~~linear (top) and~~ logarithmic ~~(middle)~~ scale. The ~~residuals are shown in~~ lower panels (c) show the ~~bottom plots~~difference between the ISRF data points and the ISRF fit (end of stage 4).

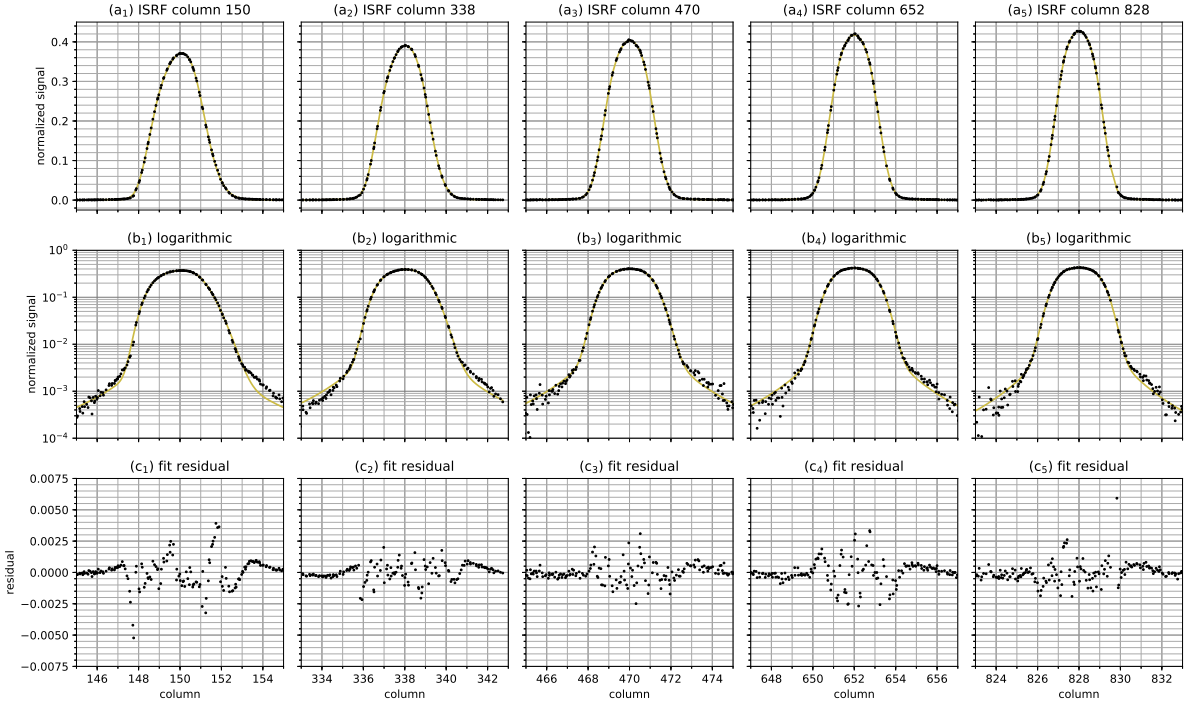

Figure 4. ~~As in Fig. **??**, for radiance~~ Five typical SWIR ISRFs determined from on-ground measurements with an external laser through the TROPOMI radiance port. The ~~ISRF fit parameter values~~ upper panels (a) show the shapes of the ISRF at location: columns ~~128~~150, ~~300~~338, ~~540~~471, ~~726~~ 652 and ~~935 are respectively: (skew-normal width) 0.620, 0.527~~828, ~~0.444~~resp., ~~0.445~~ and ~~0.451; (skew) 2.569~~rows 24, ~~1.650~~76, ~~0.786~~118, ~~-0.701~~ 155 and ~~-0.879; (block width) 2.641, 2.607, 2.452~~191, ~~2.363 and 2.307;~~ resp. The middle panels (~~tail fraction~~b) ~~0.079, 0.108, 0.110, 0.100 and 0.106;~~ show the same data on a logarithmic scale. The lower panels (~~tail shape~~c) ~~1.65;~~ show the difference between the ISRF data points and the ISRF fit (~~tail width~~end of stage 4)~~1.625, 1.238, 1.142, 1.235 and 1.196~~.

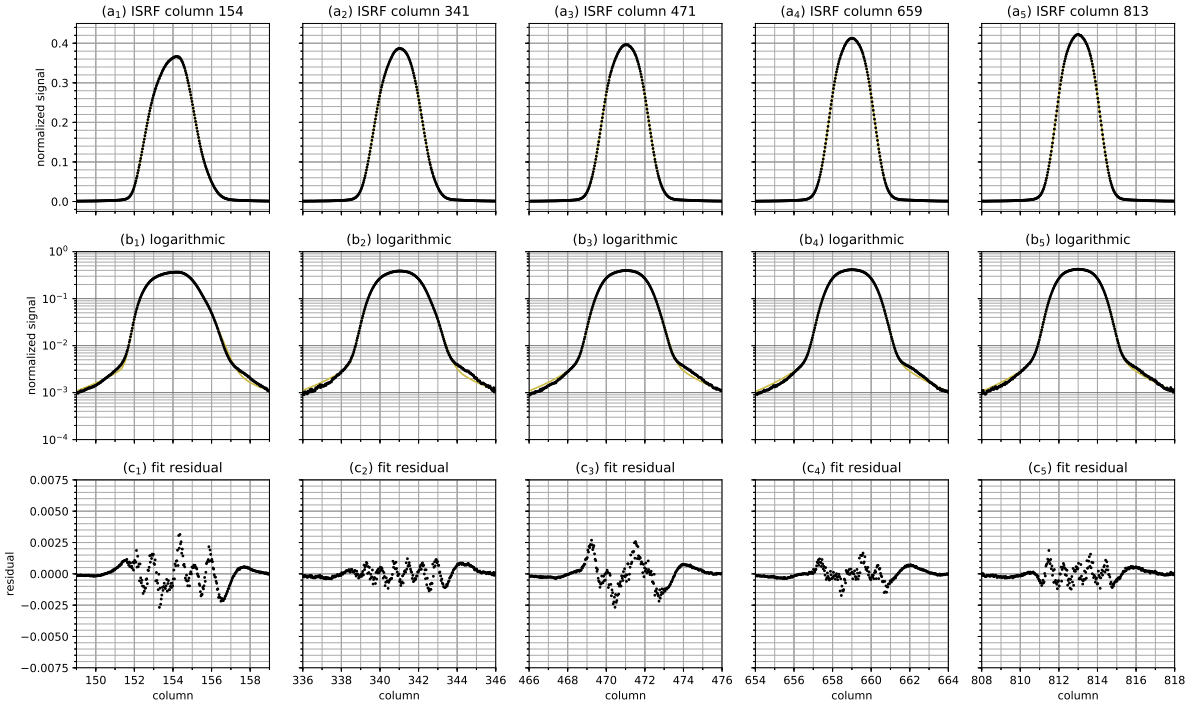

Figure 5. SWIR ISRF determined from on-ground measurements with the five on-board diode lasers. The upper panels (a) show the shapes of the ISRF at location: columns 154, 341, 471, 659 and 813, resp., and rows 24, 76, 118, 155 and 191, resp. The middle panels (b) show the same data on a logarithmic scale. The lower panels (c) show the difference between the ISRF data points and the ISRF fit (end of stage 4).

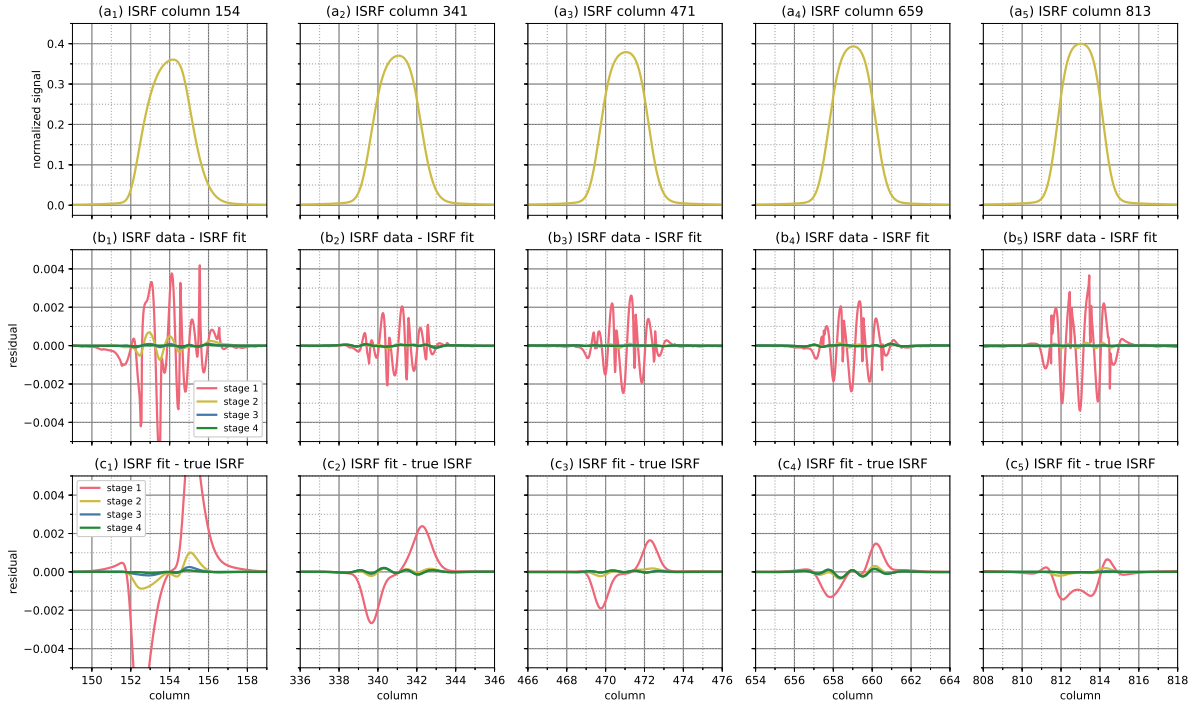

Figure 6. Convergence of five synthetic ISRF determinations in four stages. The upper panels (a) show the shapes of synthetic ISRF closely resemble the ISRF examples as shown in Figs. 3–5. The middle panels (b) show the difference between the ISRF data points and the ISRF fit. The lower panels (c) show the difference between the ISRF data points and the true ISRF.

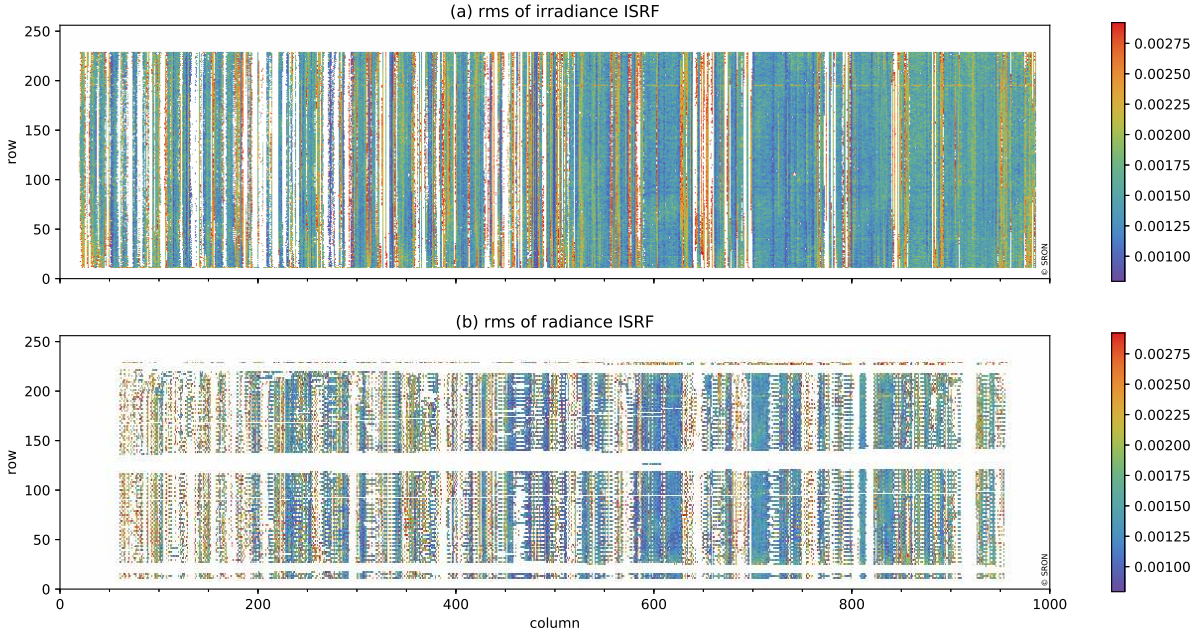

Figure 7. Fit quality of the local ISRF using the (a) irradiance port and (b) radiance port. No measurements are performed in the white edges. Panel (a) white vertical stripes are due to saturation in the measurements, and red vertical stripes are due to nearby saturation or laser instabilities. Panel (b) white areas are due to partly illuminated rows.

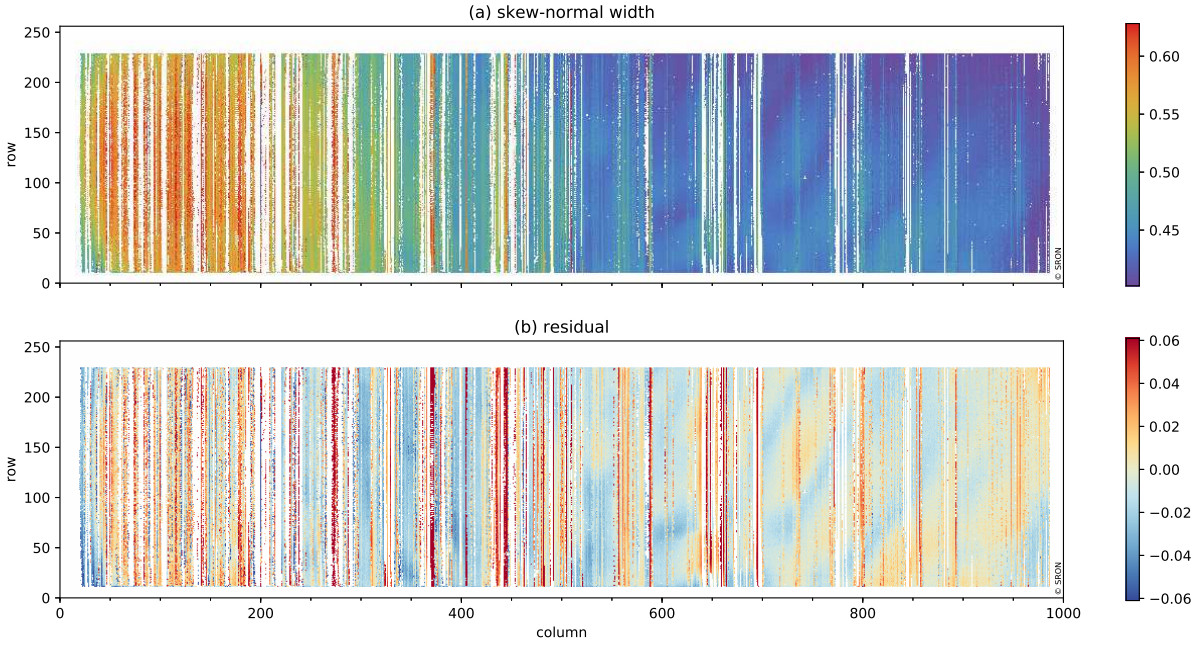

Figure 8. ~~Irradiance~~ ISRF ~~fit results for the peak function (final):~~ parameter skew-normal width d ~~. Presented are (a) ISRF fit, (~~irradiance port)~~ISRF parameter fit, and ~~. Panel (b) shows the difference between ~~the two~~d and its bivariate poynomial fit. ~~In~~ No measurements are performed in the white ~~area, the ISRF fit failed (~~edges. White vertical stripes ~~), the light is blocked by the entrance slit of the spectrometer (top and bottom) or a shield at~~ are due to saturation in the ~~detector (left and right)~~measurements. ~~The white spots~~ Red vertical stripes are ~~bad pixels~~due to nearby saturation or laser instabilities.

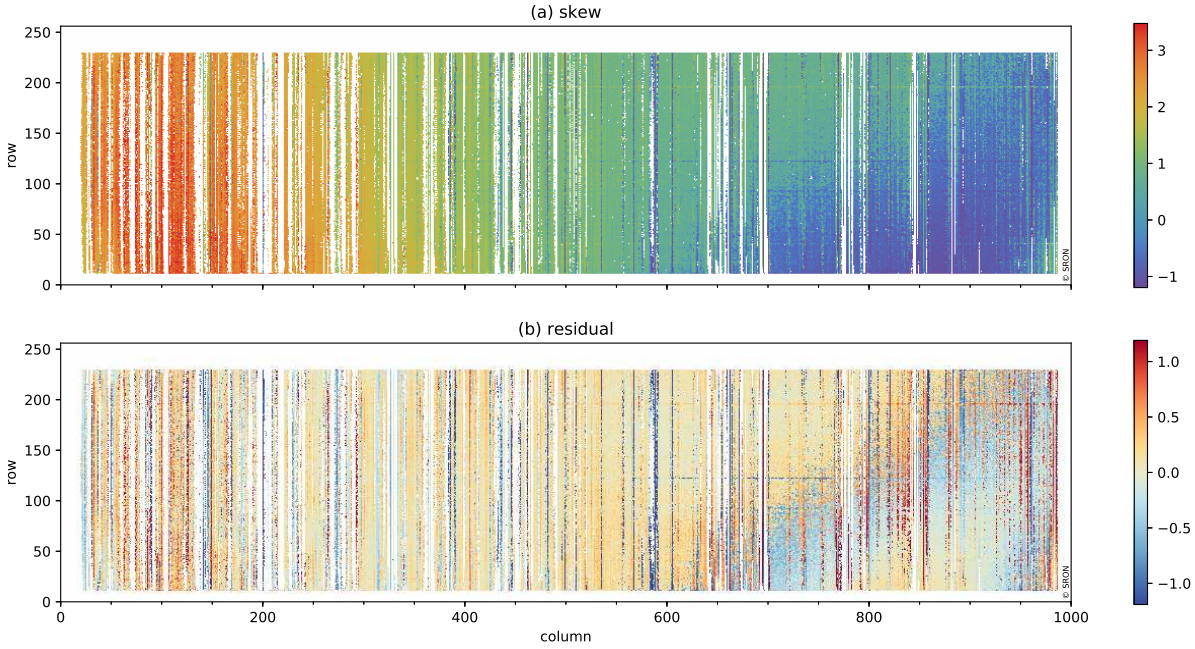

Figure 9. ~~Irradiance~~ ISRF ~~fit results:~~ parameter skew *s* (airradiance port)~~quality of ISRF fit and~~ . Panel (b) ~~quality of ISRF parameter fit, based on~~ shows the ~~rms value. The threshold between good and moderate is 0.004 and~~ difference between ~~moderate~~ *s* and ~~bad 0.0065~~its bivariate poynomial fit. ~~In the gray area, the light is blocked by~~ No measurements are performed in the ~~entrance slit of~~ white edges. White vertical stripes are due to saturation in the ~~spectrometer (top and bottom)~~ measurements. Red vertical stripes are due to nearby saturation or ~~by the shield at the detector (left and right)~~laser instabilities.

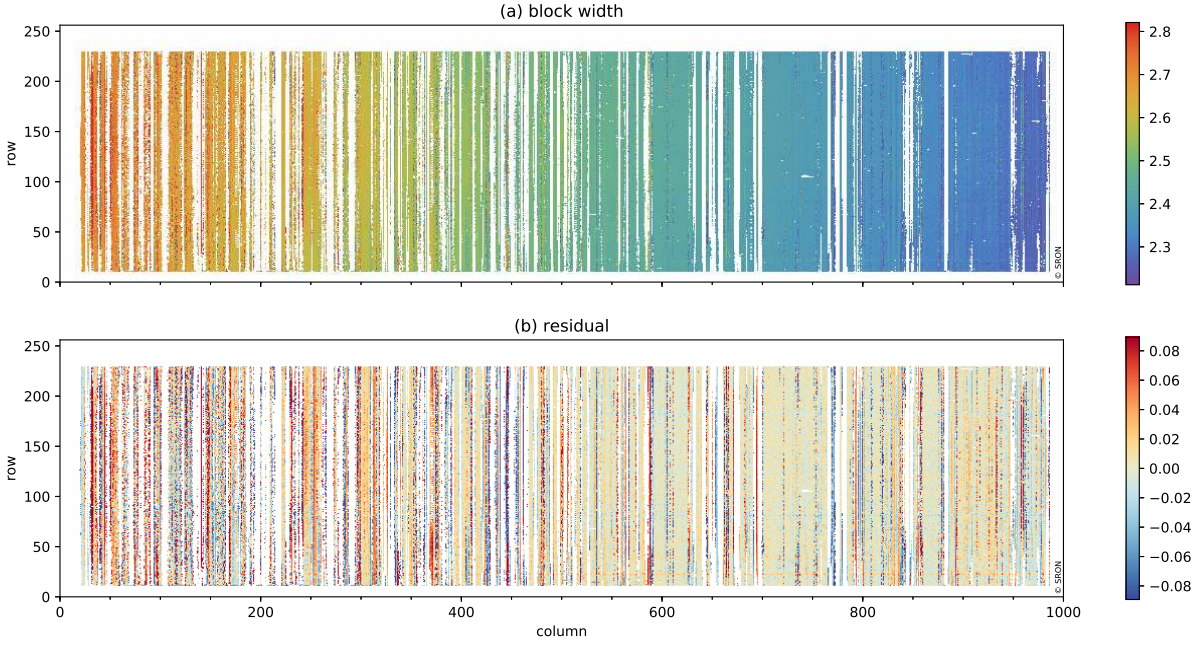

Figure 10. ~~Same as Fig~~ISRF parameter block width w (irradiance port). ~~**??, except for**~~ Panel (b) shows the ~~radiance ISRF~~difference between w and its bivariate poynomial fit. No measurements are performed in the white edges. White vertical stripes are due to saturation in the measurements. Red vertical stripes are due to nearby saturation or laser instabilities.

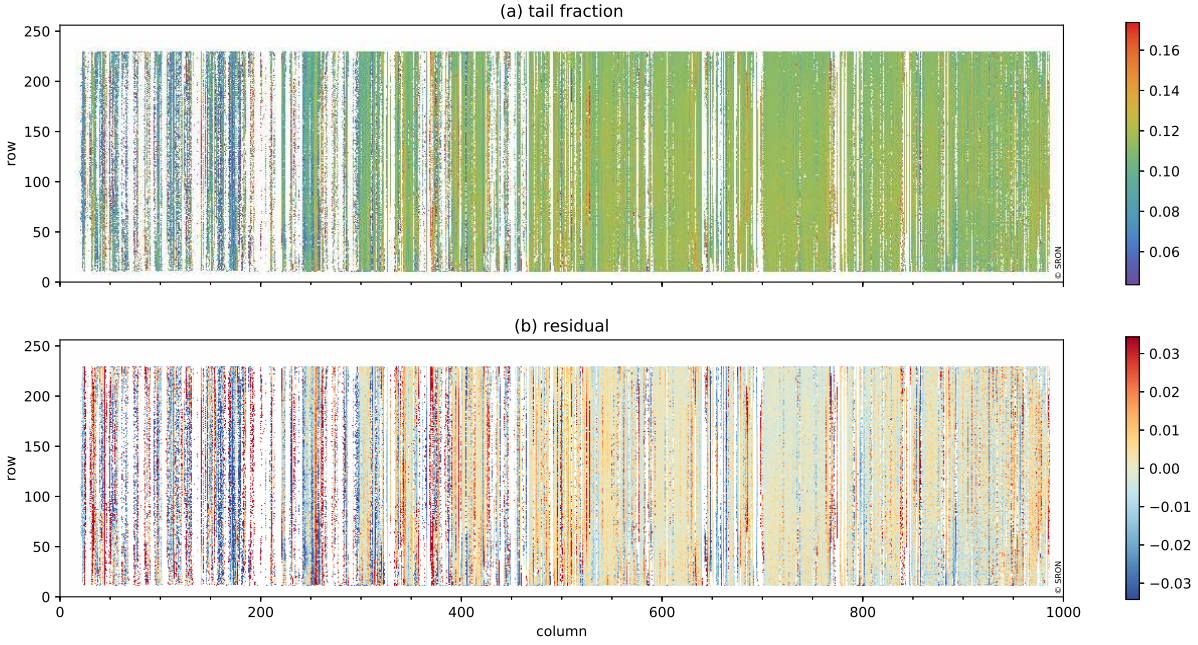

Figure 11. ~~Comparison of the five~~ ISRF ~~parameters derived from~~ parameter tail fraction η (irradiance port). Panel (b) shows the difference between η and ~~radiance~~ its bivariate poynomial fit. No measurements ~~: as a function of column~~are performed in the white edges. White vertical stripes are due to saturation in the measurements. Red vertical stripes are due to nearby saturation or laser instabilities.

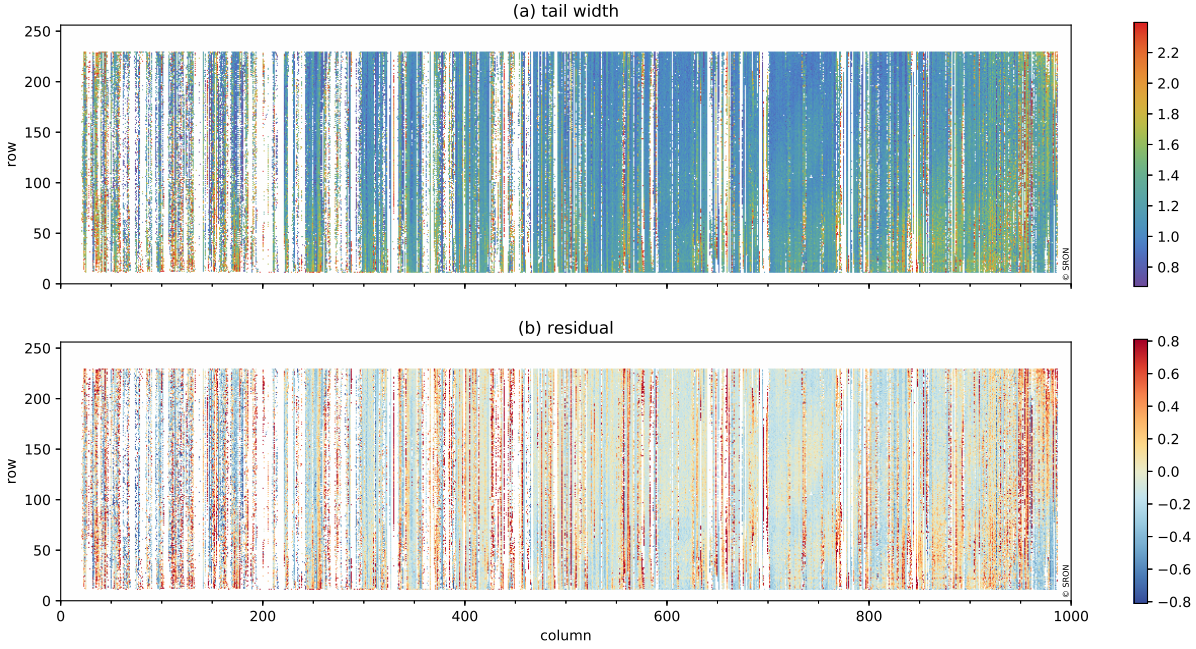

Figure 12. ISRF tail parameter γ (irradiance port). Panel (b) shows the difference between γ and its bivariate poynomial fit. No measurements are performed in the white edges. White vertical stripes are due to saturation in the measurements. Red vertical stripes are due to nearby saturation or laser instabilities.

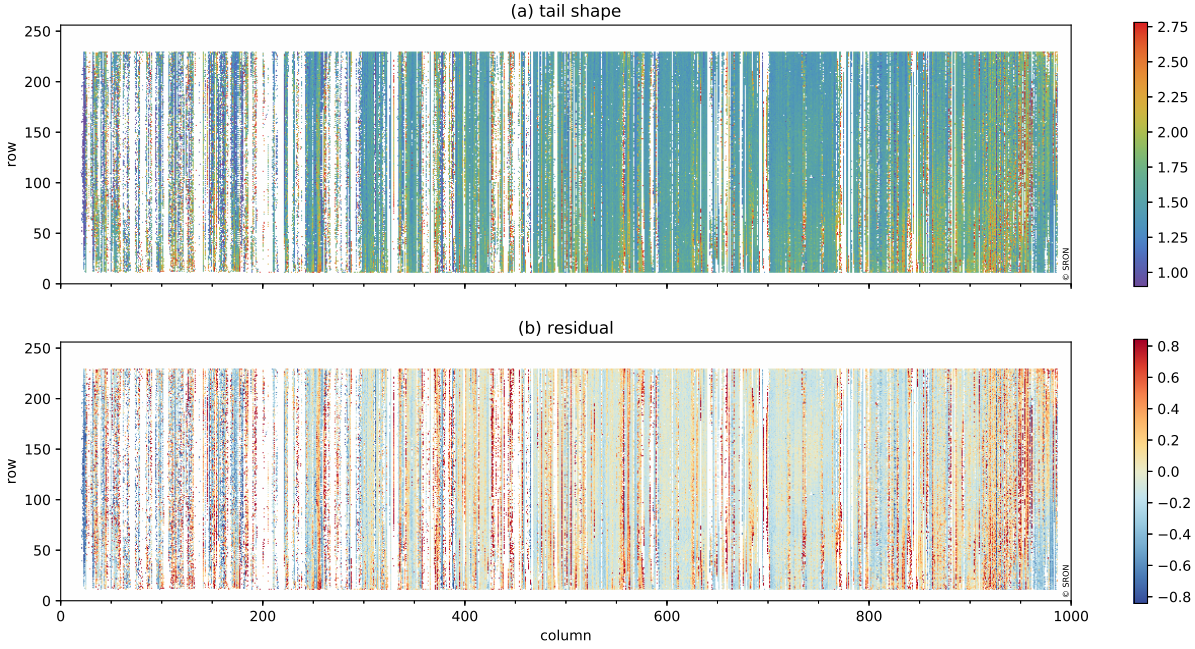

Figure 13. ISRF tail parameter m (irradiance port). Panel (b) shows the difference between m and its bivariate poynomial fit. No measurements are performed in the white edges. White vertical stripes are due to saturation in the measurements. Red vertical stripes are due to nearby saturation or laser instabilities.

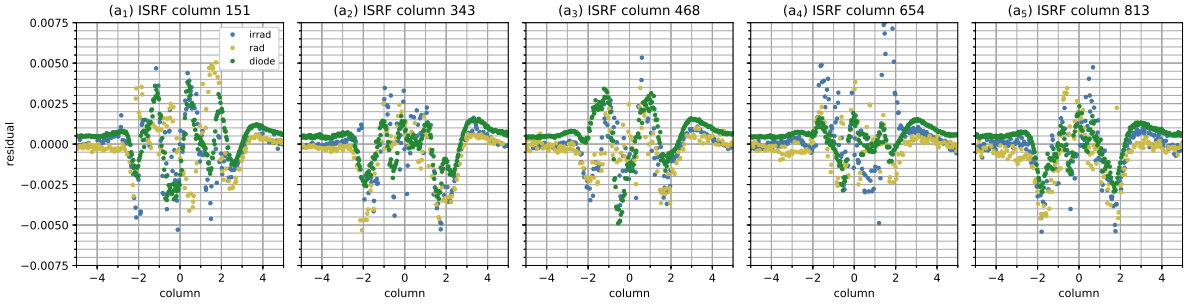

Figure 14. Differences between the local ISRF and the smoothed ISRF for irradiance, radiance and diode-laser measurements. This figure illustrates that smoothed ISRF data agree well with the local ISRF data presented in Figs. 3–5.

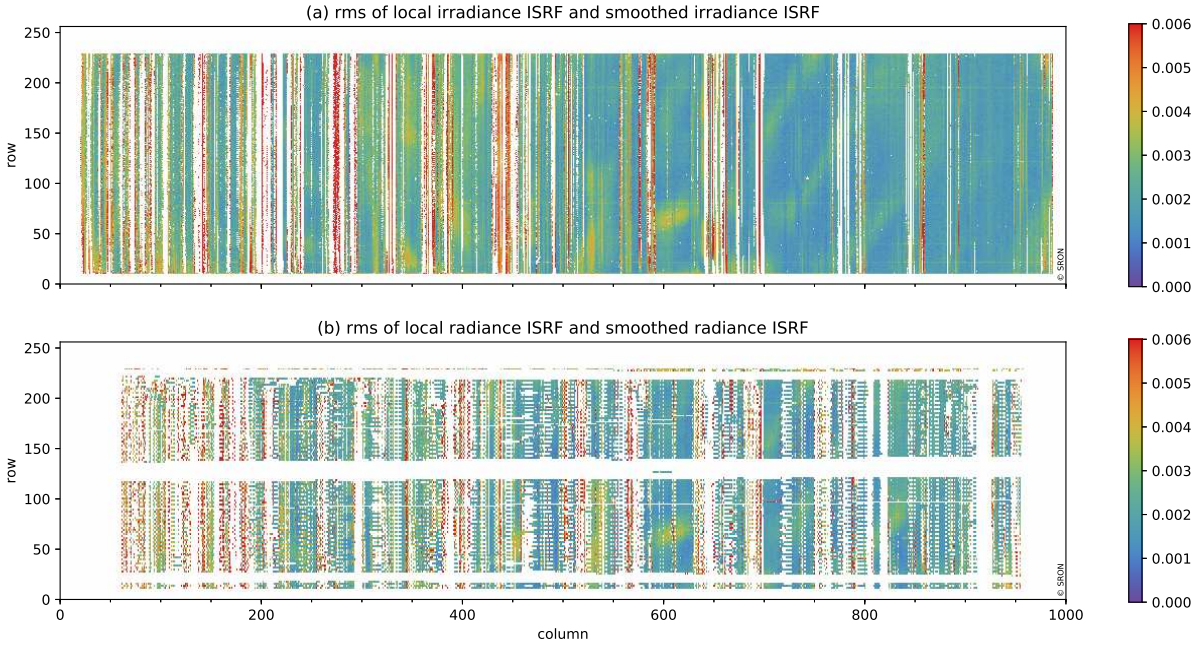

Figure 15. Fit quality of the smoothed ISRF using the (a) irradiance port and (b) radiance port. No measurements are performed in the white edges. Panel (a) white vertical stripes are due to saturation in the measurements, and red vertical stripes are due to nearby saturation or laser instabilities. Panel (b) white areas are due to partly illuminated rows.

Table 1. ~~Treatment of parameters in each stage and for each type~~ The convergence of ~~data set: either hold constant (entry is a number), interpolate~~ the ~~previous~~ ISRF ~~results (entry 'model') or include in~~ parameter iteration is shown by listing the ~~fit (entry 'fit'). In~~ determined ISRF parameters at the ~~case~~ end of ~~s~~ four stages. The simulations are performed on five synthetic ISRFs representative for ~~ISSF fits,~~ the ~~interpolated value is given the opposite sign~~determined TROPOMI-SWIR ISRF, equally spaced at positions on the SWIR detector from top left to ~~take into account that an ISSF is basically~~ bottom right. The synthetic data are generated without noise. The shape parameters for the ~~mirrored version of an~~ true ISRF are listed in italic.

~~stage data~~
(47, 15
~~2~~
(79, 34
~~4~~ (118, 4
~~Listing the intermediate parameters of the derived ISRF from two synthetic ISRF data sets: 'skew' and 'symmetric'. The ISRF parameter iteration is per~~
~~skew~~ (155
~~symmetric~~ (1

Table 2. Summary Listing of the irradiance ISRF characterizationparameters (Eq.Listed are for the 4 stages: the quality of the ISRF fit, the average and uncertainty over the complete array 13) of the rms, d, s, $w$five typical SWIR ISRFs determined from irradiance (irr), η radiance (rad) and γdiode-laser (ld) measurements, and the residuals between these parameter values and the parameter fitsrespectively. To avoid the effect The location of outliers, the biweight estimate of location and scale ISRF are used for (approximately) equally spaced at positions on the average and uncertainty, respectively (Beers et al., 1990)SWIR detector from top left to bottom right.

fit parameter pixel	type	stage 1 d	stage 2 s	stage 3 w	stage 4
quality (47, 151)	good irr	142409 0.5573	164012 3.3047	164646 2.66357	160161 0
(47, 150)	moderate rad	53253 0.5771	35639 2.2230	35118 2.67038	36560 0.
(47, 154)	bad ld	13296 0.5681	9907 3.1603	9604 2.64913	12946 0.
(79, 343)	failed irr	2617 0.4923	2015 1.4446	2206 2.57038	1907 0.1
rms 10^{-3} (79, 338)	median rad	3.15 ± 1.68 0.4941	2.44 ± 1.41 1.5636	2.41 ± 1.40 2.58129	2.53 ± 1.49
skew-normal width d (79, 341)	median ld	0.479 ± 0.058 0.4961	0.473 ± 0.060 1.6075	0.470 ± 0.060 2.58128	0.471 ± 0.05
(118, 468)	residual irr	-0.003 ± 0.028 0.4539	-0.004 ± 0.025 1.2030	-0.004 ± 0.025 2.51046	-0.005 ± 0.0
skew s (118, 470)	median rad	0.940 ± 1.164 0.4558	1.005 ± 1.204 1.0764	1.018 ± 1.207 2.50779	1.019 ± 1.18
(118, 471)	residual ld	0.018 ± 0.490 0.4561	0.017 ± 0.442 1.2055	0.012 ± 0.439 2.52954	0.010 ± 0.41
block width w (155, 654)	median irr	2.484 ± 0.176 0.4569	2.494 ± 0.168 1.0076	2.501 ± 0.166 2.42063	—0.10
(155, 652)	residual rad	-0.003 ± 0.052 0.4341	-0.004 ± 0.049 0.5962	-0.003 ± 0.049 2.42946	—0.10
tail fraction η (155, 659)	median ld	0.075 ± 0.029 0.4330	0.087 ± 0.029 0.7617	0.092 ± 0.029 2.42022	0.097 ± 0.01
(191, 813)	residual irr	-0.001 ± 0.029 0.4122	-0.001 ± 0.029 0.7598	-0.001 ± 0.029 2.36538	-0.001 ± 0.0
tail width γ (191, 828)	median rad	1.240 ± 0.562 0.4142	1.048 ± 0.321 0.3014	1.004 ± 0.255 2.35932	0.988 ± 0.13
(191, 813)	residual ld	0.231 ± 0.558 0.4168	0.066 ± 0.298 0.6230	0.040 ± 0.218 2.36022	0.009 ± 0.10

Irradiance ISRF fit results for the peak function (final): skew s. Presented are (a) ISRF fit and (b) ISRF parameter fit, see also Fig. 8.

Same as Fig. 9, except for block width w.

Same as Fig. 9, except for tail fraction η.

Same as Fig. 9, except for tail width γ.