

Response to Referee #1:

We appreciate the very helpful feedback from the referee. The referee's comments are listed in *italics*, followed by our response in **blue**. New/modified text in the manuscript is in **bold**.

*Major points:*

*1. I would like to see more relevance for MethaneAIR w.r.t. MethaneSAT. Is the method presented here is only briefly revisited in the last few paragraphs of the conclusions. More and/or an expanded discussion on this point likely improves the quality of the paper.*

Information on the similarities and differences between the MethaneAIR and MethaneSAT instruments has been added to Section 2, at line 80 of the original manuscript.

**“The airborne instrument provides similar spectroscopy to MethaneSAT with higher spatial resolution, although with a significantly smaller swath width (5.05 km at a flight altitude of 12 km for MethaneAIR vs. ~260 km for MethaneSAT) due to the difference in operating altitude. The MethaneAIR point spread function (PSF) is roughly 2.5 pixels wide across-track, estimated from the spatial stray light data. The swath angles of MethaneAIR and MethaneSAT are similar, at 23.7° and 21.3°, respectively. While MethaneSAT has a larger FPA of 2048 × 2048 pixels compared to the 1024 × 1280 pixels of MethaneAIR, the spectral range of the satellite instrument is reduced due to illumination of only 1000 spectral pixels at most. The spectral resolution of MethaneSAT, however, is 20%-30% higher than that of MethaneAIR. One of the most significant differences between the instruments is the detector material; the satellite detector is Mercury-Cadmium-Telluride (MCT), unlike that of MethaneAIR.”**

The stray light correction for MethaneSAT is under development and will be similar to (but more sophisticated than) the method presented for MethaneAIR. The following sentence has been added to line 151 to tie in MethaneSAT:

**“This method will also be applied to MethaneSAT, incorporating lessons learned from MethaneAIR.”**

The following sentence has been added to line 225 of the original manuscript regarding the connection between MethaneAIR and MethaneSAT for ISRF characterization:

**“The spectral calibration of MethaneSAT will follow and build upon the method of ISRF determination presented here for MethaneAIR.”**

In addition, the retrieval algorithm used to evaluate the calibration in Section 7 is being developed for both MethaneAIR and MethaneSAT, as pointed out in the first sentence of Section 7.

*2. Section 2: Although less important than the ISRF, the PSF size should be mentioned.*

The PSF may be estimated from the spatial stray light data to be roughly 2.5 pixels wide across-track. This approximation is limited given that data were collected at only three spatial positions. The following sentence is added to the first paragraph in Section 2:

**“The MethaneAIR point spread function (PSF) is roughly 2.5 pixels wide across-track, estimated from the spatial stray light data.”**

*3. Section 3: I am missing information on source (i.e. laser) stability for the measurements as well as expected noise levels.*

The following sentences were inserted at line 127 of the original manuscript to quantify the laser stability:

**“The wavelength stability of the CH<sub>4</sub> laser is  $\pm 5$  pm, and the optical power is  $+1.2\%/-0.9\%$ . For the O<sub>2</sub> laser, the wavelength stability is  $+5/-6$  pm, and the optical power is  $+1.1\%/-0.5\%$ .”**

*4. Figures: Some figures are relatively small, with important details being referenced at times difficult to discern. Enlarging some figures (notably Fig. 4, 5, 6, 7, 12, 13, 14) greatly enhances their effect.*

The referenced figures have been enlarged.

*5. Figure 6: Straylight kernels are often four times bigger than the detector itself to account for far-field. See e.g. Fig. 11a in Tol et al., 2018. This is not given here. Even if these far-field straylight are within the noise, this is a key part of the kernel. Please give the full kernel.*

We don't have comparable spatial coverage on the FPA to TROPOMI (3 swath angles in MethaneAIR vs. 99 swath angles in TROPOMI SWIR), so we have to focus on the relative near field, where more than one coverage can be achieved. The spectral coverage in the CH<sub>4</sub> band is further limited by the QE drop beyond 1656 nm, where the noise is too large. The following statement is added to the end of the second paragraph in Section 4 to clarify this point:

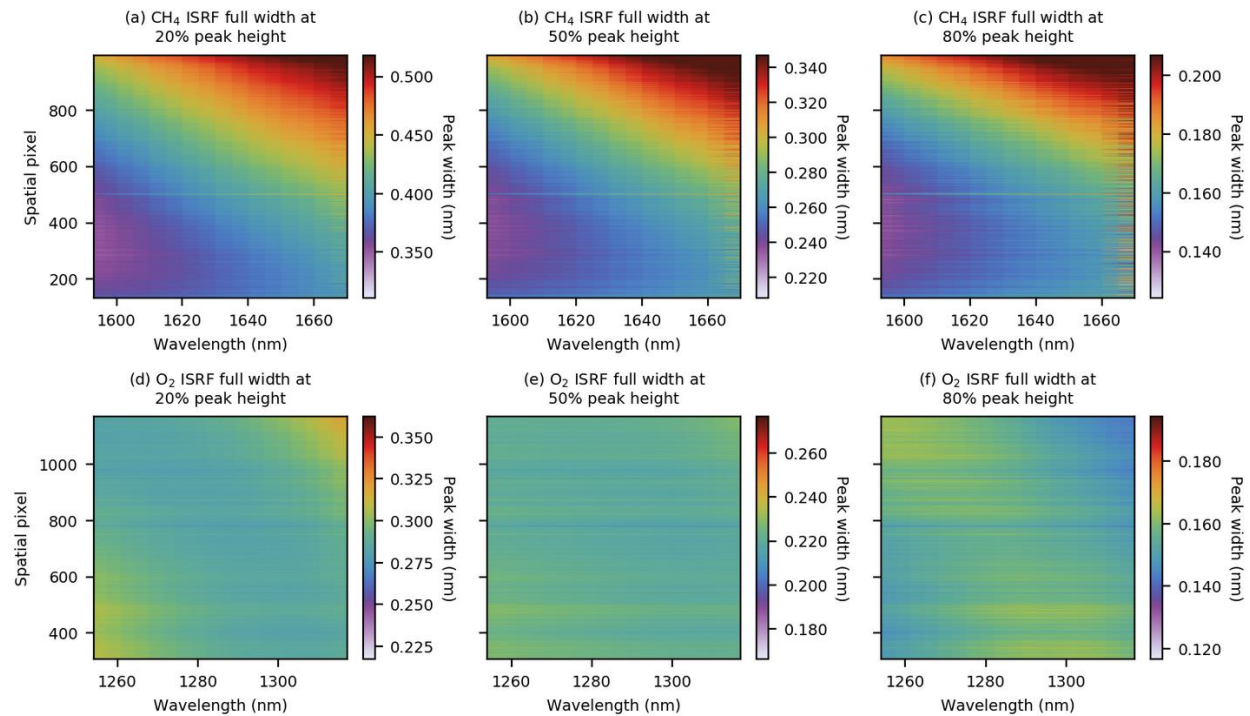
**“This process ideally gives a stray light kernel that is up to four times larger than the FPA. However, due to the limited spatial swath angles, the region near the spatial dimension edges are often subject to excessive noise. The spectral coverage in the CH<sub>4</sub> band is further**

limited by the QE drop towards the long wavelength. As such, we limit the stray light kernel to be within +/-400 pixels for both spatial and spectral dimensions.”

6. Figure 14: For the total ‘smoothed’ ISRF there are discrete transitions. It appears there is no truly smooth ISRF. Why was this done? Please explain.

The transitions correspond to the discrete wavelengths where ISRF measurements were collected. Figure 14 has been updated using Gouraud shading in order to show the smooth ISRF variation across the FPA. The updated Figure 14 and caption is included here:

“



**Figure 14. Full width of the smoothed ISRF at 20% (a,d), 50% (b,e), and 80% (c,f) of maximum peak height for all spatial and spectral pixels. Gouraud shading is applied to render smooth ISRF variation across the FPA. Wavelength labels have been projected to the abscissa in order to provide more context compared to the spectral pixel index. The top three panels correspond to the CH<sub>4</sub> band, which shows a general broadening of the ISRF with increasing wavelength and spatial pixel index. In contrast, the O<sub>2</sub> band ISRF is more homogeneous across different spatial and spectral pixels, as reinforced by the relatively narrower scales for the bottom three panels. The color limit in each panel is fixed at +/- 25% from the FPA-mean value.”**

7. Section 7: Only the ISRF was validated in flight. Although in that derivation a straylight correction was done, this is not presented. Was there some activity on straylight validation?

The stray light correction is demonstrated for laboratory ISRF measures as shown in Figure 7. Currently we don't have sufficient resources for in-flight stray light validation given the limited flight hours. Potential validation efforts include retrieval sensitivity test over high-contrast scenes and comparison with ground-based measurements (e.g., from EM-27), but those are not deemed to be within the scope of this paper.

*Minor points:*

*1. Line 2: ...anthropogenic CH<sub>4</sub> point sources...*

The words “CH<sub>4</sub>” and “point” have been swapped, as suggested.

*2. Line 2 (as well as later): Define scale of ‘basin’*

The first sentence of the abstract has been revised.

**“MethaneAIR is the airborne simulator of MethaneSAT, an area-mapping satellite currently under development with the goal of locating and quantifying large anthropogenic CH<sub>4</sub> point sources as well as diffuse emissions at the spatial scale of an oil and gas basin.”**

Clarification has been added to the sentence referring to ‘basin-scale’ at line 32:

**“MethaneSAT aims to characterize oil and gas basin-scale, diffuse CH<sub>4</sub> emissions through a wide swath of 260 km, and at the same time locate and quantify large point sources within each target area that is typically 200 × 140 km<sup>2</sup> (Benmergui et al., 2020; MethaneSAT, LLC, 2020).**

*3. Line 22: minor, but GOSAT is a satellite, while TROPOMI is an instrument. They are treated equally here. Sentinel-5P is the satellite carrying TROPOMI.*

This sentence at line 22 has been revised to provide TROPOMI satellite information.

**“Currently, the Greenhouse gases Observing SATellite (GOSAT) and TROPOspheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5P satellite collect CH<sub>4</sub> abundance data at a global scale (Yoshida et al., 2011; Hu et al., 2016).”**

*4. Line 30/31: Define ‘intermediate’ scales*

The use of ‘intermediate scale’ references the previous sentence (line 28) to be loosely defined as a scale larger than that of GHGSat-D and smaller than that of TROPOMI. The sentence at line 30/31 has been revised:

**“MethaneSAT is a push broom imaging satellite under development that is designed to operate at a scale in between that of current target-mode satellites and global mappers.”**

5. Line 32/33: I assume the swath is 200 km, not 200 km<sup>2</sup>?

Upon re-checking, the swath at nadir is 260 km (not km<sup>2</sup>). The sentence at line 32/33 in the original manuscript has been updated:

**“MethaneSAT aims to characterize oil and gas basin-scale, diffuse CH<sub>4</sub> emissions through a wide swath of 260 km and at the same time locate and quantify large point sources within each target area that is typically 200 × 140 km<sup>2</sup> (Benmergui et al., 2020; MethaneSAT, LLC, 2020).**

6. Line 36: Please motivate the choice for the 1.27 O<sub>2</sub> band more (e.g. explain recent advances, why proximity is better)

The text at line 36 in the original manuscript has been revised to include more information about the O<sub>2</sub>  $a^1 \Delta_g$  band.

**“Although the O<sub>2</sub> A band (~0.76 μm) has been commonly used for this purpose, the O<sub>2</sub>  $a^1 \Delta_g$  band may be more advantageous. Recent advances in separating emitted airglow from backscattered light enable the use of the O<sub>2</sub>  $a^1 \Delta_g$  band in remote sensing applications (Sun et al., 2018; Bertaux et al., 2020). Further, the close spectral proximity of the O<sub>2</sub>  $a^1 \Delta_g$  band to the CH<sub>4</sub>/CO<sub>2</sub> bands is favorable, as any differences in aerosol and cloud optical properties between the bands are reduced (Sun et al., 2018).”**

7. Line 64: Remove ‘only’

Removed.

8. Line 64 - 69 : Only TROPOMI is mentioned here. Are there other stray light treatments, e.g. GOSAT, GOSAT2, that are worth mentioning?

GOSAT1/2 are scanning Fourier transform spectrometers where stray light behaves differently than imaging grating spectrometers. OCO-2/3 are similar grating spectrometers with many fewer across-track footprints (8), and the stray light is simply treated as a zero-level offset to our knowledge. The TROPOMI SWIR band is the closest reference. We have added the sentence at line 66:

**“The MethaneAIR stray light is measured and corrected for following the TROPOMI SWIR band, which is the closest reference to MethaneAIR with a wide across-track swath.”**

9. Line 71 or 80: similar(ity) – At times it must be guessed how similar or dissimilar MethaneSAT is from MethaneAIR (see major point). Please quantify as much as possible.

As stated in response to major point #1, the sentence at lines 80-81 of the original manuscript has been revised to include swath widths.

The following text has also been added to provide detail regarding the similarities and differences between MethaneAIR and MethaneSAT.

**“The airborne instrument provides similar spectroscopy to MethaneSAT with higher spatial resolution, although with a significantly smaller swath width (5.05 km at a flight altitude of 12 km for MethaneAIR vs. ~260 km for MethaneSAT) due to the difference in operating altitude. The swath angles of MethaneAIR and MethaneSAT are similar, at 23.7° and 21.3°, respectively. While MethaneSAT has a larger FPA of 2048 × 2048 pixels compared to the 1024 × 1280 pixels of MethaneAIR, the spectral range of the satellite instrument is reduced due to illumination of only 1000 spectral pixels at most. The spectral resolution of MethaneSAT, however, is 20%-30% higher than that of MethaneAIR. One of the most significant differences between the instruments is the detector material; the satellite detector is Mercury-Cadmium-Telluride (MCT), unlike that of MethaneAIR.”**

*10. Line 81: quantify swath width in text. One important part is the swath \*angle\* difference between MethaneSAT and MethaneAIR.*

The sentence at line 81 in the original manuscript has been edited to include swath widths, and an additional sentence has been added to provide the swath angles:

**“The airborne instrument provides similar spectroscopy to MethaneSAT with higher spatial resolution, although with a significantly smaller swath width (5.05 km at a flight altitude of 12 km for MethaneAIR vs. ~260 km for MethaneSAT). The swath angles of MethaneAIR and MethaneSAT are similar, at 23.7° and 21.3°, respectively.”**

*11. Line 86: sub-pixel \*spectral\* smile*

The word “spectral” has been inserted into the sentence at line 86.

**“The optical design provides sub-pixel spectral smile and keystone distortion and relatively uniform focus across wavelength and field angle.”**

*12. Line 83: Figure 5 referenced before Figures 2,3 and 4.*

The figure order has been rearranged such that the original Figure 5 is now Figure 2.

*13. Line 118: I do not understand the motivation to force the intercept to be zero.*

The dark current has been removed so zero digital number should correspond to zero radiance. If an intercept is included in the fitting, errors in the upper range of the calibration curve will



propagate to the intercept, potentially leading to very large relative errors in the low radiance range. Lines 116-118 are updated:

**“Flat field data were taken at exposure times of 50, 100, and 150 ms, matching the exposure times used in flight and in the ISRF calibration, and the dark frames were subtracted. This resulted in curves of spectral radiance ( $\text{phot s}^{-1} \text{cm}^{-2} \text{sr}^{-1} \text{nm}^{-1}$ ) vs. dark-subtracted focal plane intensity ( $\text{DN s}^{-1}$ ) for each exposure time and every active pixel. These radiometric calibration curves were fitted by fifth-order polynomials with the intercept forced to be zero. The zero intercept is necessary to guarantee zero radiance at zero DN.”**

*14. Line 136-139: What is the relevance of these statements?*

This paragraph has been revised to provide more context.

**“The measurements described above are preliminary and were used primarily to develop the stray light correction algorithm. In the near future, the stray light will be measured with higher precision and an updated correction will be derived. Improvements to the measurement setup are currently underway and include automated tilt and translation stages to address many more field angles and an all-reflective collimator to avoid stray reflections from the refractive corrector plate. In addition, the pinhole will be replaced with a  $100 \mu\text{m}$  slit oriented perpendicular to the spectrometer entrance slit, in order to fill the width of the spectrometer slit while providing a point source in the across-track dimension.”**

*15. Line 149: Quantify the decrease in QE*

Lines 90-91 address the QE decrease. The sentence in the original manuscript has been revised to include additional data points:

**“The QE begins to roll off above this wavelength, decreasing to 0.52 at 1660 nm, then 0.26 at 1670 nm before reaching a minimum of about 0.15 at the 1680 nm end of the  $\text{CH}_4$  passband.”**

*16. Figure 5 : Please mark which laser responses were done with different power to account for the decrease in QE*

The sentence at lines 147-148 of the original manuscript has been revised to provide the wavelengths and laser power range, which may help to further address comment #15 above.

**“In the  $\text{CH}_4$  band, the laser power was increased progressively from -3.0 dBm at wavelengths  $\leq 1640 \text{ nm}$ , up to +2.5 dBm at 1670 nm to compensate for the decrease in QE at longer wavelengths.”**

*17. Line 186 : Is the 2 % for both? Please quantify for either spectrometer.*

Both are approximately 2%, if a  $\pm 400$  spectral and spatial pixel window is defined. The sentence at line 186 has been removed, and the following sentence is added to line 177:

**“The sum of the far-field kernel is 2.4% for the CH<sub>4</sub> band and 2.1% for the O<sub>2</sub> band. This indicates that the stray light is small relative to the useful signals at each spatial and spectral position.”**

*18. Figure 10: Are these normalized? Please give Y-axis.*

The first unscaled, unshifted panel is not normalized. The y-axis has been added to Figures 9 and 10.

*19. Line 242: Please give ISRF beyond 7.5 pixels in nm as well.*

Line 242 in the original manuscript has been revised, and a sentence has been added to include the quantity in nm:

**“Since values in the smoothed ISRF beyond  $\pm 7.5$  pixels from the center should be taken care of by the spectral stray light correction as described in Section 4, these values are set equal to zero. This area corresponds to pixels outside of the central wavelength  $\pm 0.75$  nm and  $\pm 0.6$  nm in the CH<sub>4</sub> and O<sub>2</sub> bands, respectively.”**

*20. Line 246: Confusing. These are individual pixels within the full spatial illumination of a laser? And are these not correlated with the bad pixels mentioned earlier?*

Lines 245-250 of the original manuscript have been revised to clarify the observed anomalous behavior of the ISRFs.

**“After applying the Savitzky-Golay filter to the ISRF across all spatial and spectral pixels, a small number of ISRFs (74 out of 7767 in the CH<sub>4</sub> band, 87 out of 8630 in the O<sub>2</sub> band) exhibit anomalous widths by way of sharp contrasts with their neighbors, presumably due to insufficient bad pixel removal. By nature, the ISRF shape should vary smoothly between spatial and spectral pixels. Due to the sparse ISRF measurements in the spectral dimension, it is desirable to remove those outliers to avoid inference to a broad wavelength range.”**

*21. Section 7 : Please give dates and lengths of flight. Date of flight can only be read from caption of Fig. 17.*



The following sentence specifying the date and time of flight has been inserted after the opening sentence of Section 7 (line 295 of the original manuscript).

**“The flight coverage began at 15:51:29 UTC and ended at 19:40:46 UTC on 11/8/2019.”**

*22. Line 303: Why was only a single across track position done? This could be repeated for similar across track positions at other points during the flight. Are results similar?*

The spectral fitting with varying ISRF width scaling factor has been done for all across-track positions in this research flight. Variations in both across- and along-track dimensions were observed, and the updated ISRF is used in the final level 2 retrieval. However, discussion of those results would require too much information on the retrieval algorithm implementation and is deemed to be out of scope of this paper. We are preparing manuscripts dedicated to the in-flight performance and retrieval algorithm. The sentence at line 317 is changed to:

**“The spectral fitting with varying ISRF width is applied to other across-track positions throughout the flight and reveals across-track and time dependent ISRF changes.”**

*23. Line 308 : Difference is attributed to temperature changes due to environment. Could the temperature difference be quantified?*

The internal temperature of the instrument was not monitored in flight. We did observe correlation between cabin temperature and ISRF changes. The sentence in lines 309-310 is modified to

**“The small F-number of both spectrometers makes the focal point sensitive to deformation of the mechanical structure. This may arise from a difference from the lab and flight environments, such as a change in the temperature of the optical bench or mechanical stress of the instrument that responds to cabin temperature and/or aircraft motion.”**