Algorithm Theoretical Basis Document

L2b: Methane and Carbon Dioxide Concentration Retrievals for Satellites

---

Version	1.0.0
Document Status	Approved
Document Owner	Daniel Cusworth
Contributors	Daniel Cusworth, Alana Ayasse

Revision Log

Version	Description	Date
0.1.0	Initial draft	5/1/2022
0.1.1	Draft for Carbon Mapper pre-CDR	12/8/2022
1.0.0	Updates post ORE	8/7/2024

Carbon Mapper, Inc.
Pasadena, CA 91105
data@carbonmapper.org

Table of Contents:

1 Background

Methane (CH_4) and Carbon Dioxide (CO_2) are greenhouse gasses that have primarily driven anthropogenic warming since the pre-industrial era. High emission CH_4 and CO_2 point sources make up a disproportionate amount of the anthropogenic budget. Carbon Mapper's mission is to detect, quantify, and publish these sources using airborne and satellite remote sensing platforms. Carbon Mapper supports policymakers and stakeholders by providing decision support tools and analyses that synthesize satellite and airborne remote sensing data into actionable insights.

The Carbon Mapper data platform is a full-scale operational implementation of a science data system that builds on 10+ years of research and development projects led by Carbon Mapper team members, initially at NASA's Jet Propulsion Laboratory supported by funding from NASA, the California Air Resources Board, and the University of Arizona. Those research projects included multiple airborne field campaigns, satellite and surface observations, and development of CH_4 retrieval algorithms, machine learning tools, multi-scale analytic frameworks, data pipelines, open data portals and synthesis analysis.

The Carbon Mapper data platform is designed to rapidly process and publish point-source CH_4 and CO_2 data from multiple satellite and airborne imaging spectrometers. The platform has been routinely processing data from airborne surveys using NASA JPL's AVIRIS-NG and the Arizona State University Global Airborne Observatory since 2022 and expanded in early 2023 to include observations from NASA's EMIT mission on the International Space Station. In 2024, the platform will begin operational processing of Planet's first two Tanager satellites which are being launched by the Carbon Mapper Coalition.

Carbon Mapper is dedicated to providing CH_4 and CO_2 data that is transparent, trusted and actionable. Here we provide an overview of our methods and procedures to quantify CH_4 and CO_2 concentration enhancements. Other Carbon Mapper Coalition documentation will describe the theoretical basis for other key detection and quantification processes.

2 Overview of data products and data processing

Figure 1 below illustrates the Carbon Mapper Coalition data products and processing levels. In brief, the L2b data product is an estimate of column CH_4 concentrations that are derived from L1b top of the atmospheric calibrated radiance using CH_4 and CO_2 absorption features at shortwave infrared (SWIR) wavelengths. L2b data then undergo plume detection and attribution procedure (L2c). Each plume is individually segmented (L3) and emission rates are quantified (L4). In this document we describe the concentration enhancement retrieval process (L2B).

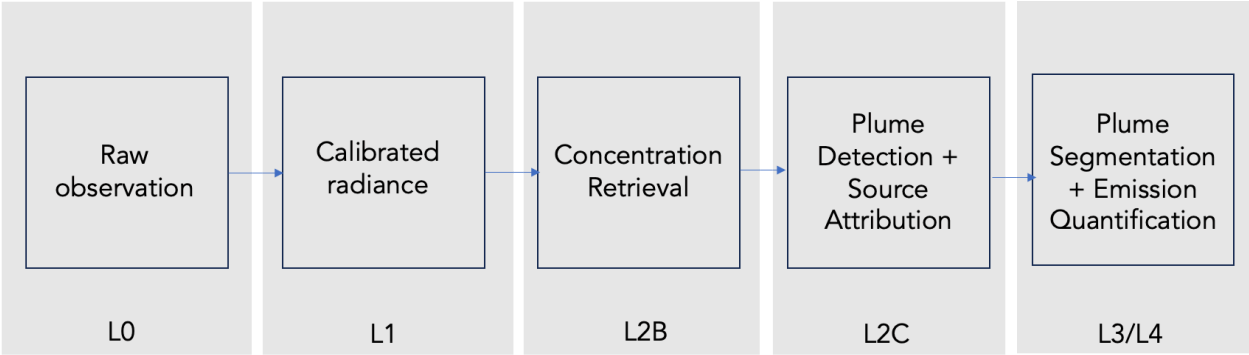

Figure 1*: Simplified data flow indicating the Carbon Mapper data processing pipeline and product levels.*

Instrument specifications for satellites that Carbon Mapper routinely processes for CH4 and CO2, Carbon Mapper Coalition's Tanager and NASA Earth Mineral Dust Investigation (EMIT), are listed in Table 1.

Table 1*: Instrument specifications for satellites that Carbon Mapper routinely processes for CH4 and CO2.*

Instrument Name	**Carbon Mapper Coalition Tanager**	**NASA EMIT**[†]
Swath width	18.6-24.2 km (varies with look angle)	75 km
Off-nadir pointing ability ("look angle")	30 degrees	None
Ground Sample Distance (GSD)	30-43 meters (varies with look angle)	60 m
Spectral response (FWHM)	5.5 nm	8.5 nm
Spectral sampling	5 nm	7.5 nm
Spectral range	400 - 2500 nm	381-2493 nm
Signal-to-noise @ 2200 nm	310 – 655 (varies with imaging mode)*	450

*35 deg Solar Zenith Angle, 25% albedo
[†]Values taken or extrapolated from Thompson et al., 2024

3 Methane (CH4) and Carbon Dioxide (CO2) retrievals

Carbon Mapper operationally implements a columnwise matched filter (CMF) to estimate pathlength enhancements of CH4 (units ppm-m) relative to the background. This retrieved quantity has an equivalent interpretation as dry air column average enhanced mole fraction (unit ppm) or column density of enhanced methane (kg m^{-2}) as reported by other passive remote sensing data providers. Each retrieval takes as input L1B non-orthorectified radiance data and outputs column CH4/CO2 concentration enhancements. The benefits of a CMF algorithm include fast computation and normalization of nonuniformities across sensor elements across the focal plane array. The retrieval utilizes SWIR windows where CH4 and CO2 exhibit strong absorption: 2100-2480 nm for CH4 and 1860-2190 nm for CO2.

3.1 Columnwise Matched Filter

The CMF algorithm seeks an estimate for concentration length of methane or CO2 ($\hat{\alpha}$) in units in parts per million meter (ppm-m) for each observed spectrum (Thompson et al., 2016; Thompson et al., 2015). This is done by testing each observed spectrum against a target signature ($\vec{t}$), accounting for noise and background covariance **Σ**. At sensor radiance $\vec{L_m}$ (unit μW cm^{-2} sr^{-1} nm^{-1}) for a pixel affected by enhanced gas concentration is modeled through Beer-Lambert's Law:

$$\vec{L_m} = \vec{L_0} e^{-\vec{k}\alpha} \quad (1)$$

Where $\vec{L_0}$ represents at-sensor radiance in presence of background levels of a gas, and $\vec{k}$ represents gas absorption. This model can be further simplified using Taylor expansion, assuming an optically thin plume:

$$\vec{L_m} \approx \vec{L_0} - \alpha \vec{t}(\vec{L_0}) \quad (2)$$

Where $\vec{t}(\vec{L_0}) = \vec{k} * \vec{L_0}$, the unit absorption spectrum, as is calculated (simulations described in section 3.3) through radiative transfer simulations of transmittance. To estimate $\vec{L_0}$, we use the mean spectrum $\vec{\mu}$ for all pixels in a "column" (i.e., all pixels in the flight direction of a single cross-track element; Figure 1) of observed data:

$$\vec{L_m} \approx \vec{\mu} - \alpha \vec{t}(\vec{\mu}) \quad (3)$$

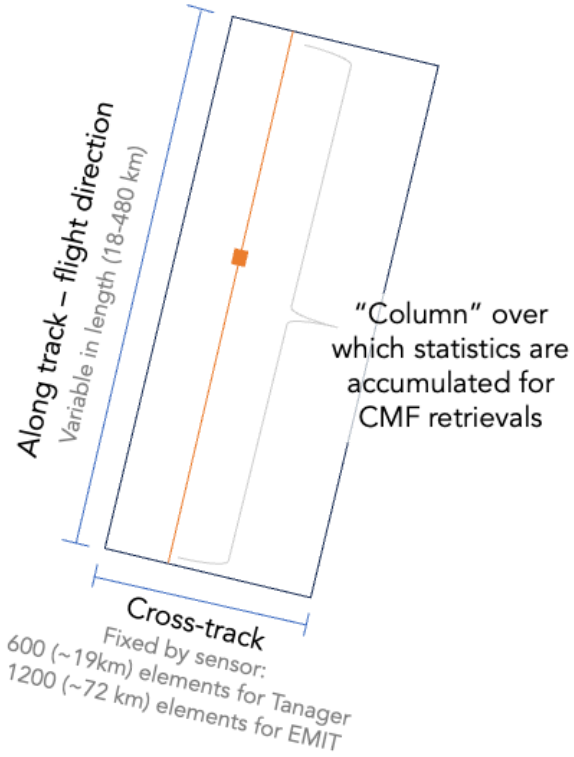

Figure 2. *Simplified diagram of an image collection as it pertains to along and cross tracks - including which pixels are used in CMF algorithms.*

The optimal value $\hat{\alpha}$ is found through optimization of log-likelihood - e.g.., for the *i*th spectrum, $\hat{\alpha}_i$ is the solution that minimizes the residual between observed and modeled spectra while accounting for covariance. The maximum likelihood solution takes the following form (Foote et al., 2021):

$$\hat{\alpha} = \frac{\vec{t}^{T}\,\Sigma^{-1}(\vec{L_m} - \vec{\mu})}{\vec{t}^{T}\,\Sigma^{-1}\vec{t}} \qquad (4)$$

3.2 Covariance Estimation

The covariance matrix **Σ** is estimated from these same pixels in the across-track column. We employ a low-rank approximation for covariance in order to stabilize the solution in Equation 4, especially under regimes of few pixels per column. This approach is explained in more detail in

Thompson et al. (2015) and Manolakis et al. (2009), but briefly we decompose the covariance matrix into *p* eigenvalues (ϕ) and eigenvectors ($\vec{q}$):

$$\Sigma = \sum_{i=1}^{p} \phi_i \vec{q}_i \vec{q}_i^{T} \quad (5)$$

We approximate the inverse using the top 30 eigenvectors (*d*=30):

$$\Sigma^{-1} = \frac{1}{\beta}\left(I - \sum_{i=1}^{d} \left[\frac{\phi_i - \beta}{\phi_i} \right] \vec{q}_i \vec{q}_i^{T} \right) \quad (6)$$

$$\beta = \frac{1}{p-d}\left(tr(\Sigma) - \sum_{i=1}^{d} \phi_i \right) \quad (7)$$

3.3 Unit Absorption Spectrum

The unit absorption spectrum is defined as the change in transmittance that relates to a perturbing signal $\vec{t}$. Recent work using airborne campaigns explored the sensitivity of CH4 and CO2 matched filter retrievals to solar zenith angle, water vapor, and ground elevation (Foote et al., 2021). For example, larger SZAs mean longer path lengths from the sun to the sensor, resulting in deeper absorption per unit enhancement of CH4 or CO2 (Figure 3). The key findings from this work were that CO2 retrievals were more sensitive than CH4 retrievals to changes in water vapor and surface elevation, and that retrievals were most sensitive to solar zenith angle variability. For CO2, using scene specific unit absorption spectra changed the integrated mass enhancement (IME) estimates (sum of excess methane in a plume of gas) by 50 to 75%, and the emissions estimates by factors of 2.4 to 6. The impact on CH4 retrievals and flux estimates was smaller, with IME changing by ~ +/- 25%.

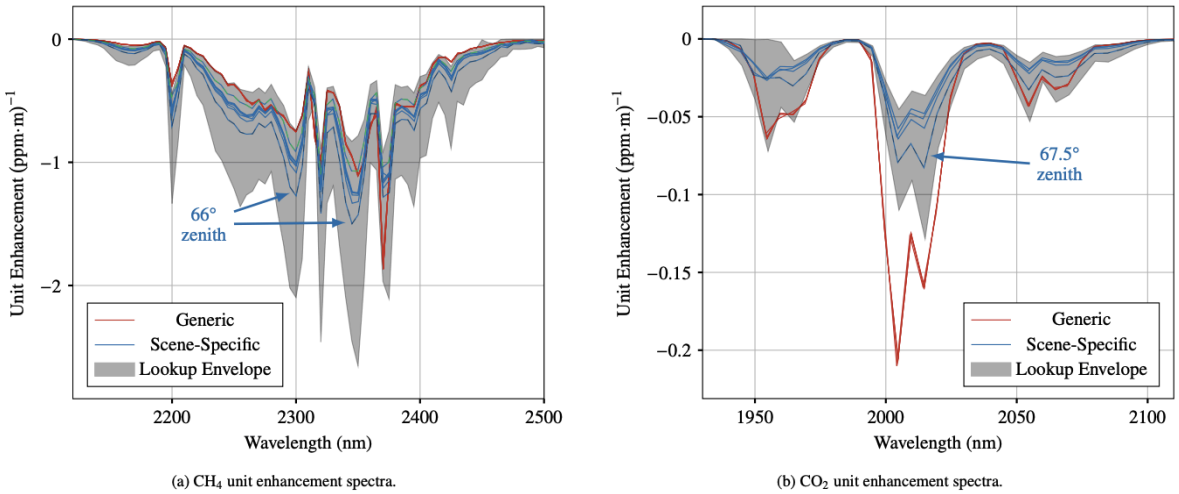

Figure 3. *Dynamic unit absorption spectra ($\vec{t}$) for CH4 and CO2. Blue lines represent unit absorption spectra for different solar zenith angles. Figure taken from Foote et al. (2021).*

We follow the approach described in Foote et al., (2021) and select a scene-specific unit absorption for each observation for both retrievals of CH4 and CO2. For column water vapor content, for each image acquisition, we query either the High Resolution Rapid Refresh meteorological product in the U.S. or the ECMWF IFS meteorological product outside the U.S. For ground elevation, for the EMIT satellite instrument, we query values provided by that data source. For Tanager, we query the USGS GTOP030 Global Digital Elevation Map. For these dynamically queried or calculated parameters, we query a lookup-table database of unit absorption spectra that were precompiled and interpolated via MODTRAN simulations of various SZAs, water vapor concentration, CH4/CO2 background concentration, and surface heights (Foote et al., 2021).

Against a benchmark of plumes detected from EMIT, the effect of environmental conditions on the dynamically calculated unit absorption spectra, and therefore emissions rates (Refer to Carbon Mapper's L3/L4 ATBD for reference on emission rate calculation), is shown in Figure 4. We find that for CH4, changes in elevation and water vapor lead to as much as 50% relative change in emission rate, when used in lieu of a generic unit absorption spectrum.

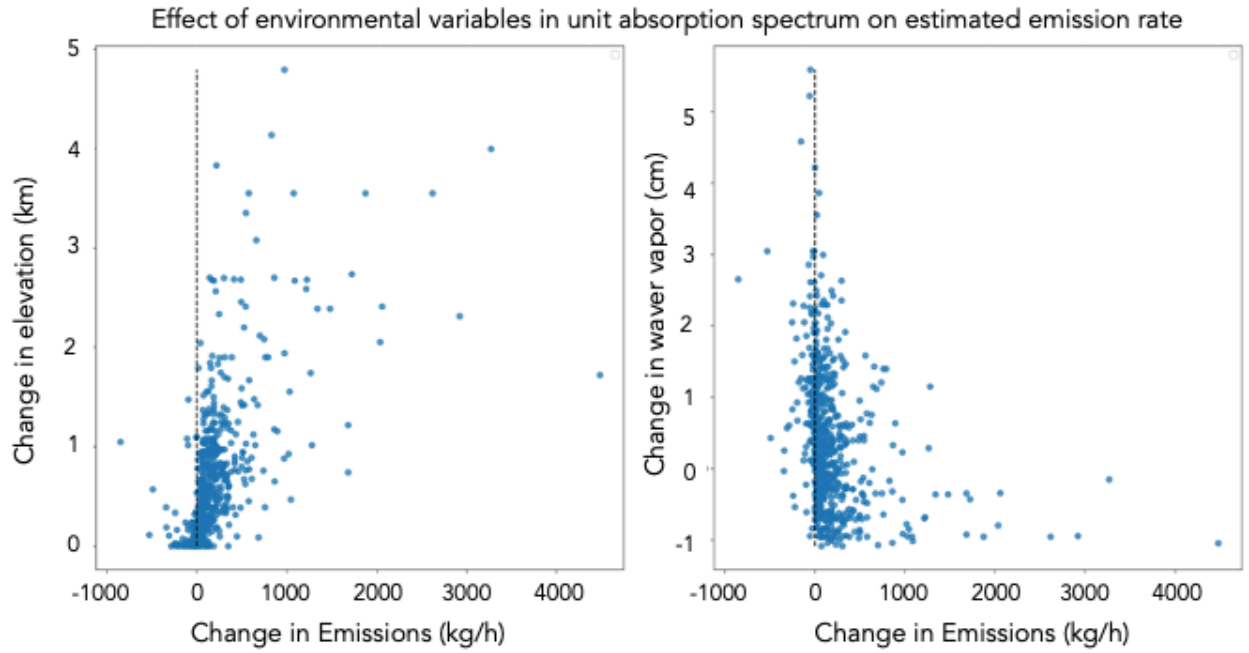

Figure 4. *Influence of environmental assumptions - ground elevation and water vapor - of unit absorption spectrum on CH4 emissions for a benchmark of EMIT observed plumes. The change in these plots is relative to a generic unit absorption spectrum.*

The final CMF product is a one band image with values in ppm-m concentration. Figure 5 shows an example of red-green-blue (RGB) channels from EMIT with CMF output in Bordj Omar Driss, Illizi, Algeria. In this scene, the bright, relatively homogeneous surface allows for clear identification of a plume signature (white pixels in the left panel of Figure 5), likely due to oil&gas operations. These concentration maps are then used for downstream processing, including plume identification, source attribution to geographic origin, attribution to emission sector and infrastructure, and emission quantification. These downstream steps are described in L2c to L4 theoretical basis documents.

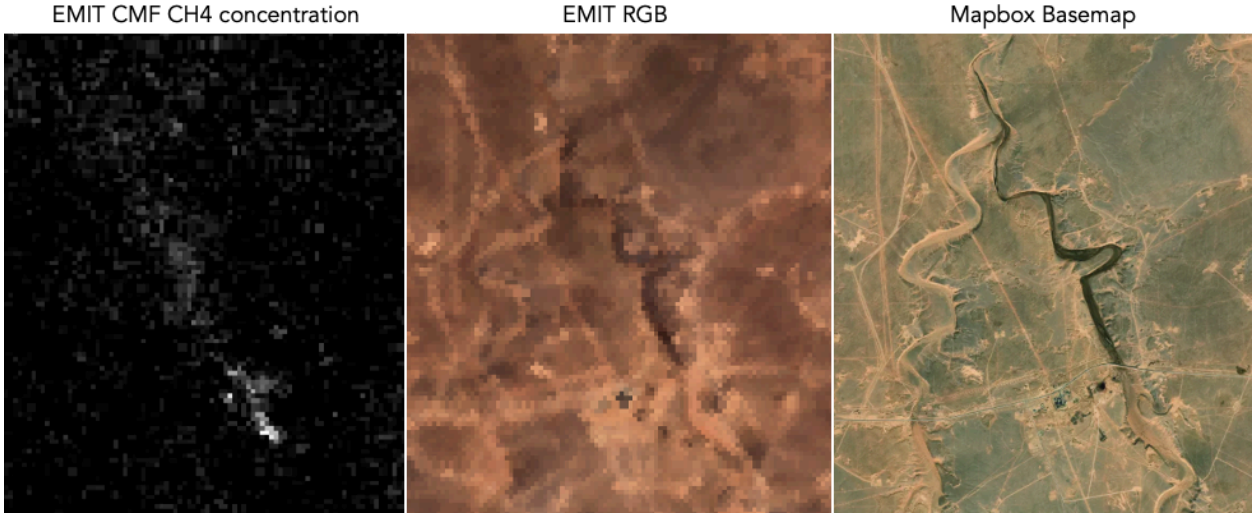

***Figure 5**. Example RGB radiance (top panel) and CMF concentrations (bottom panel). White regions represent areas of high methane concentration.*

4 Bad pixel masking

Confusing spectral surfaces that are dark, specular reflectors, or have hydrocarbons present can produce artifacts in matched filter results, as these surfaces may produce absorption signals similar to CH4 and CO2, but which are not actually associated with CH4 or CO2 enhancements. Additionally, emitted radiance from oil/gas flaring can cause strong retrieval artifacts and may contaminate the background covariance estimate, which can produce column-wise artifacts that persist over many pixels across an image (Figure 6). Therefore, to not expose potentially uncharacteristic pixels to the mean and covariance estimates of the CMF retrieval, we implement a pixel filter to radiance imagery beforehand: pixels that exceed 1.5 $\mu W\ cm^{-2}\ nm^{-1}\ sr^{-1}$ (L1b radiance units) at 2390 nm, which is typical of pixels due to flaring or other bright or specular reflecting surfaces (Cusworth et al., 2021), are removed from the analysis. Carbon Mapper has future plans to implement other pixel based thresholds to remove dark pixels (water bodies, cloud shadows, etc) from retrievals.

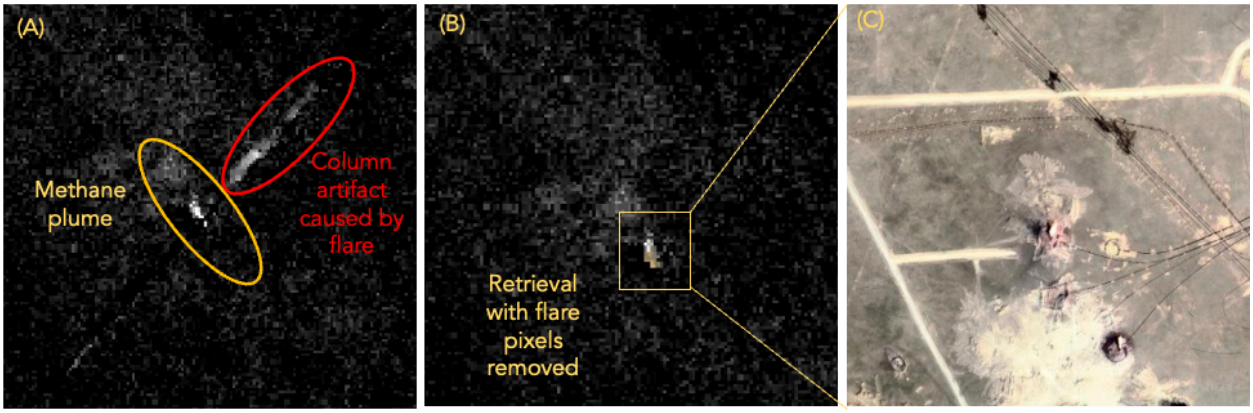

***Figure 6**. Example of column artifact due to flaring. Panel A shows an example of partially combusting flare: a clear methane plume is visible, but emitted radiance from the flare leads to column artifacts (red circle). After removal of these flare pixels (panel B), the column artifact is diminished while retaining the methane plume signature. Panel C shows a close up image of the combustion source.*

Clouds also confuse the mean and covariance used for matched filter retrievals. Carbon Mapper pre-screens each satellite acquisition using independently derived cloud masks as part of the L1b data collections provided by NASA JPL for EMIT and Planet for Tanager. For scenes with minimal corruption from clouds (less than 20%), we allow the CMF to proceed and all downstream data processing (L2c to L4) is performed. For scenes with excess cloud

concentration (excess of 60%), we classify these as unusable for reliable retrieval and therefore limit any subsequent data processing. For scenes in between these cloud thresholds, we perform manual visual quality inspection to check how much cloud corruption impacts mean and covariance estimates in certain regions of the scene. For EMIT, some columns are dropped entirely before delivery of L1b datasets. For processing of CMF and subsequent downstream products, Carbon Mapper only processes EMIT data where at least 30% of columns are fully intact.

5 Validation

Traditional approaches for calibration of column average gas mixing ratios (e.g., using a ground network like the Total Carbon Column Observing Network (TCCON)) are not readily applicable for application of validating plume enhancements. Generally, validation of plume products from L2b to L4 products occur at the L4 level through controlled release experiments or comparison with independent evaluation of emissions (e.g., simultaneous mass-balance observations). Starting in 2024, controlled release experiments for EMIT and Tanager are planned, though experience with application of CMF algorithms in blinded controlled releases for similarly designed aircraft instruments have shown low bias (El Abbadi et al., 2023). Still, validation against controlled releases complicates the separation of error or biases due to retrievals from other downstream processing steps. Therefore, sensitivity experiments to understand potential error sources in CMF algorithms can be used to inform best practices for retrieval. For example, using aircraft data in the Permian Basin, Texas, Ayasse et al. (2023) compared the effect of artificially cropping L1b radiance data in the along track / column direction and found that too-few along track pixels can lead to a low-bias in concentration retrievals and subsequently, emission quantification (Figure 7).

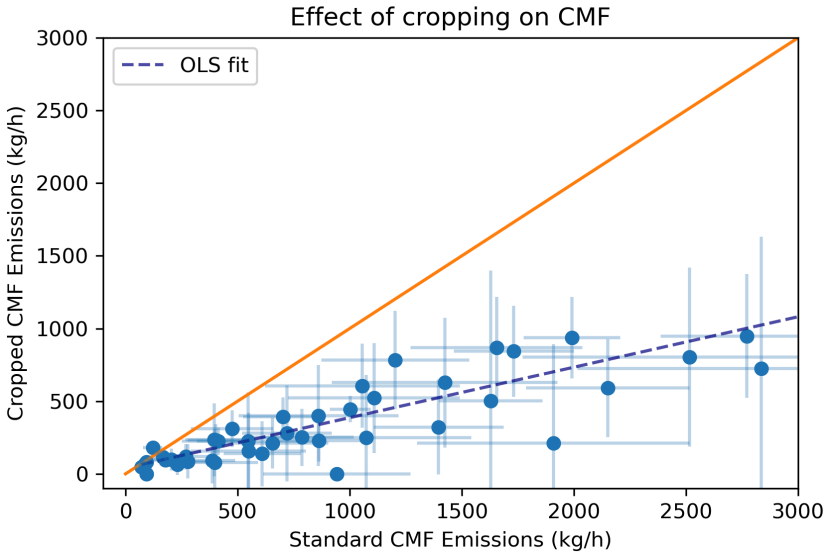

Figure 7*. Example RGB radiance (top panel) and CMF concentrations (bottom panel). White regions represent areas of high methane concentration. Figure taken from Ayasse et al., 2023.*

The study further found that to reproduce original IME values derived from fully flown flight lines, the CMF algorithm generally requires at least 15 km distance in the along-track direction (Figure 8), though optimal flight lengths will necessarily depend on the surface reflectance and atmospheric conditions present at the time of observation. EMIT acquires data that generally spans 80 km in length, and Tanager's length varies by imaging mode with variable lengths requested during tasking. However, the satellite will typically collect imagery between 18 - 480 km in length. Therefore, generally we anticipate sufficient pixels for each column to constrain mean and covariance for CMF application. In cases where insufficient pixels are available for robust CMF quantification (e.g., 18 km line where most data are corrupted by dark pixels), other more computationally expensive pixelwise retrieval processing algorithms (i.e., IMAP-DOAS (Cusworth et al., 2019)) may be deployed. Carbon Mapper originally plans to release concentration and emission estimates based on CMF algorithms, but plans to include IMAP-DOAS or other pixelwise retrieval concentration estimates in future data offerings.

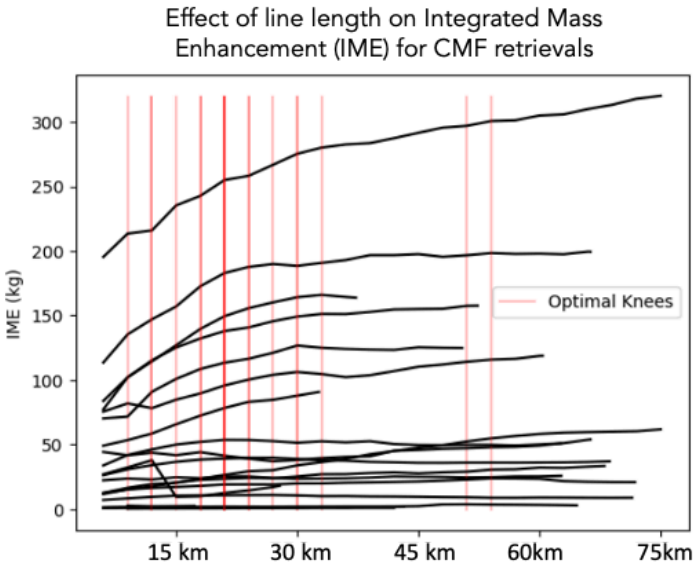

Figure 8*. The dependence of IME on the along-track length of a scene for CMF algorithms. Most of the full-length IME is reproduced between 15-45 km.*

4 Outputs

4.1 ***CMF-derived CH4 and CO2 concentrations***: provided as an *n* x *m* x 1 geotiff where *n* corresponds to number of along-track pixels in a scene and *m* is the number of cross-track elements in a scene. Concentrations are reported in ppm-m CH4 or CO2 column enhancements.

5 References

Ayasse, A., Cusworth, D., O'Neill, K., Thorpe, A. and Duren, R., 2023. Performance and sensitivity of column-wise and pixel-wise methane retrievals for imaging spectrometers.

Cusworth, D.H., Duren, R.M., Thorpe, A.K., Olson-Duvall, W., Heckler, J., Chapman, J.W., Eastwood, M.L., Helmlinger, M.C., Green, R.O., Asner, G.P. and Dennison, P.E., 2021. Intermittency of large methane emitters in the Permian Basin. *Environmental Science & Technology Letters*, *8*(7), pp.567-573.

Cusworth, D.H., Jacob, D.J., Varon, D.J., Chan Miller, C., Liu, X., Chance, K., Thorpe, A.K., Duren, R.M., Miller, C.E., Thompson, D.R. and Frankenberg, C., 2019. Potential of next-generation imaging spectrometers to detect and quantify methane point sources from space. *Atmospheric Measurement Techniques*, *12*(10), pp.5655-5668.

El Abbadi, S.H., Chen, Z., Burdeau, P.M., Rutherford, J.S., Chen, Y., Zhang, Z., Sherwin, E.D. and Brandt, A.R., 2023. Comprehensive evaluation of aircraft-based methane sensing for greenhouse gas mitigation.

Foote, M.D., Dennison, P.E., Sullivan, P.R., O'Neill, K.B., Thorpe, A.K., Thompson, D.R., Cusworth, D.H., Duren, R. and Joshi, S.C., 2021. Impact of scene-specific enhancement spectra on matched filter greenhouse gas retrievals from imaging spectroscopy. *Remote Sensing of Environment*, *264*, p.112574.

Foote, M.D., Dennison, P.E., Thorpe, A.K., Thompson, D.R., Jongaramrungruang, S., Frankenberg, C. and Joshi, S.C., 2020. Fast and accurate retrieval of methane concentration from imaging spectrometer data using sparsity prior. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(9), pp.6480-6492.

Manolakis, D., Lockwood, R., Cooley, T. and Jacobson, J., 2009, August. Hyperspectral detection algorithms: Use covariances or subspaces?. In *Imaging Spectrometry XIV* (Vol. 7457, pp. 200-207). SPIE.

Thompson, D.R., Leifer, I., Bovensmann, H., Eastwood, M., Fladeland, M., Frankenberg, C., Gerilowski, K., Green, R.O., Kratwurst, S., Krings, T. and Luna, B., 2015. Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane. *Atmospheric Measurement Techniques*, *8*(10), pp.4383-4397.

Thompson, D.R., Thorpe, A.K., Frankenberg, C., Green, R.O., Duren, R., Guanter, L., Hollstein, A., Middleton, E., Ong, L. and Ungar, S., 2016. Space-based remote imaging spectroscopy of the Aliso Canyon CH4 superemitter. *Geophysical Research Letters*, *43*(12), pp.6571-6578.

Thompson, D.R., Green, R.O., Bradley, C., Brodrick, P.G., Mahowald, N., Dor, E.B., Bennett, M., Bernas, M., Carmon, N., Chadwick, K.D. and Clark, R.N., 2024. On-orbit calibration and performance of the EMIT imaging spectrometer. *Remote Sensing of Environment*, *303*, p.113986.