CARBON MAPPER

Algorithm Theoretical Basis Document

L3/L4: Methane and Carbon Dioxide Emission Quantification for Satellites

Version 1.0.0
Release date October 24th, 2024

Carbon Mapper, Inc.
info@carbonmapper.org

Table of Contents

1 Background

Methane (CH4) and Carbon Dioxide (CO2) are greenhouse gasses that have primarily driven anthropogenic warming since the pre-industrial era. High emission CH4 and CO2 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 CH4 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 CH4 and CO2 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 CH4 and CO2 data that is transparent, trusted and actionable. Here we provide an overview of our methods and procedures to quantify CH4 and plumes, along with relevant QC review. Other Carbon Mapper Coalition documentation will describe the theoretical basis for other key retrieval and detection processes.

2 Overview of data products and data processing

Figure 1 below lists Carbon Mapper Coalition data products and processing levels. In brief, the L2b data product is an estimate of column CH4 concentrations that are derived from L1b top of the atmospheric calibrated radiance using CH4 and CO2 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 quantification processes (L3 and L4 products).

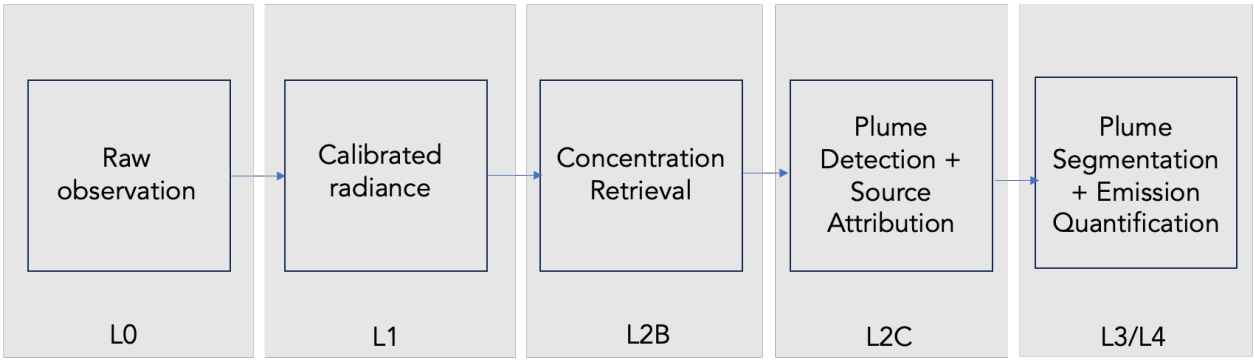

***Figure 1**: Simplified data flow indicating the Carbon Mapper data processing pipeline and product levels.*

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 Plume segmentation and emission quantification

Carbon Mapper implements an automated plume segmentation and delineation process on plumes identified and geolocated during the L2c data processing step. The L2c processing step output is an origin location (latitude, longitude) for a plume given a unique L2b concentration map (units in column enhancement CH4 or CO2 (units ppm-m)). The L3 process then segments a plume around this origin point to create a masked plume boundary that is used for mass and emission quantification. The segmentation algorithm proceeds as follows (visual example in Figure 2):

1. The L2b concentration map is cropped around the origin of a plume: +/- 2500 meters in both directions
2. A concentration threshold is dynamically determined to separate lingering background enhancements from plume enhancements. This threshold is subtracted off the cropped concentration map around the plume origin.
3. Connected pixels of enhanced concentration (> ppm-m threshold) are grouped together. A cluster must contain 5 pixels to be considered part of the plume.
4. A proximity metric is enforced on each cluster group. Separated clusters that exceed 15 pixels from the plume origin are excluded from the plume

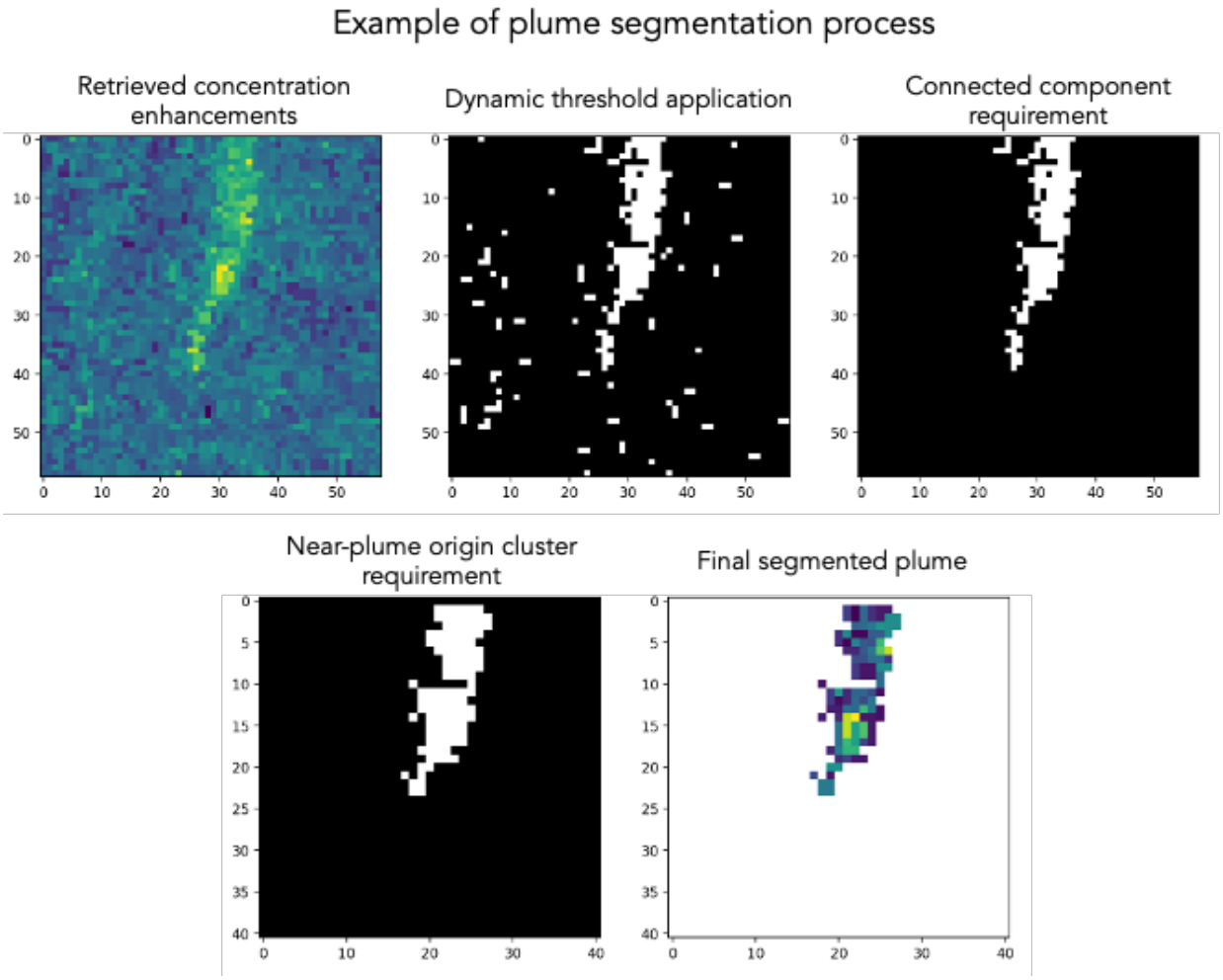

***Figure 2**. Visual example of plume segmentation processed applied to a CH4 detected from EMIT observation.*

Step 2 - determining a concentration threshold - has the largest impact on plume segmentation. For each plume, we seek a dynamically estimated threshold, based on the noise/background concentration information in the immediate vicinity around a plume. Thresholds are parameterized as a percentile of all pixels within a cropped distance around the origin of the plume (example shown in Figure 3a). Selection of a proper percentile/crop threshold is estimated by comparing to independent releases (e.g., airborne underflights or controlled releases). Figure 3b demonstrates how concentration threshold varies as a function of crop lengths and percentile applied to retrieved concentrations. There is a consistent, strong relationship between crop and percentile: concentration threshold stay roughly fixed against a set of crop/percentile parameters - for example, in Figure 3, a crop of 1500 m and the 90th percentile results in a similar concentration threshold (800-1000 ppm) as a 500 m crop and 80th percentile. Therefore we seek a segmentation solution across a set of crop/percentile pairs that result in consistent concentration thresholds.

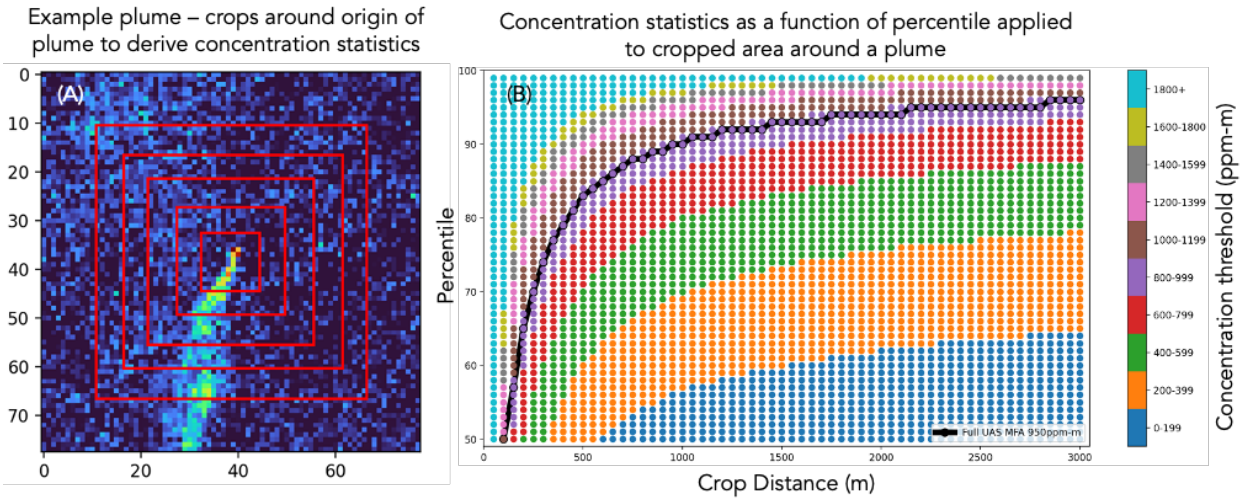

Figure 3. *The process of cropping an image of retrieved concentrations around the origin of a plume (panel A). Concentration thresholds (units ppm-m) as a function of spatial crop around a plume and percentile of retrieved concentrations within that crop for a benchmark of plumes detected using the EMIT satellite instrument (panel B). The black line represents the optimized set of crop/percentile pairs we apply to detected plumes.*

To find the best set of crop/percentile pairs, we perform an optimization procedure where emission rates are calculated for a benchmark of detected satellite plumes under a variety of crop/percentile threshold candidates. To compare against independent data, segmented plumes must be processed to emission rate estimates. This is done by incorporating the Integrated Mass Enhancement (IME; units kg; Thompson et al., 2016) approach, which calculates the excess mass emitted to the atmosphere from a source:

$$IME = \alpha \sum_{i=1} \Omega_i A_i \quad (1)$$

Where i refers to a single plume pixel, Ω is the concentration enhancement of that pixel, α is a unit conversion scalar (from ppm-m to kg m^{-2}), and A is the area of that pixel (m^2). We calculate an emission rate Q using the following relationship (Ayasse et al., 2023):

$$Q = \frac{IME}{L} U \quad (2)$$

Where U is the 10-m wind speed (m s^{-1}) and L is the plume length (m). Here U is taken from the High Resolution Rapid Refresh (HRRR) 3km, 60 minute reanalysis product in the U.S. and the ECMWF IFS 9km outside the U.S. Comparisons (Figure 4) of these gridded forecast products to

10 m weather station data in the U.S. (via the Synoptic Weather Data API: synopticdata.com) show high scatter for any data point, and roughly 1 m/s mean absolute bias.

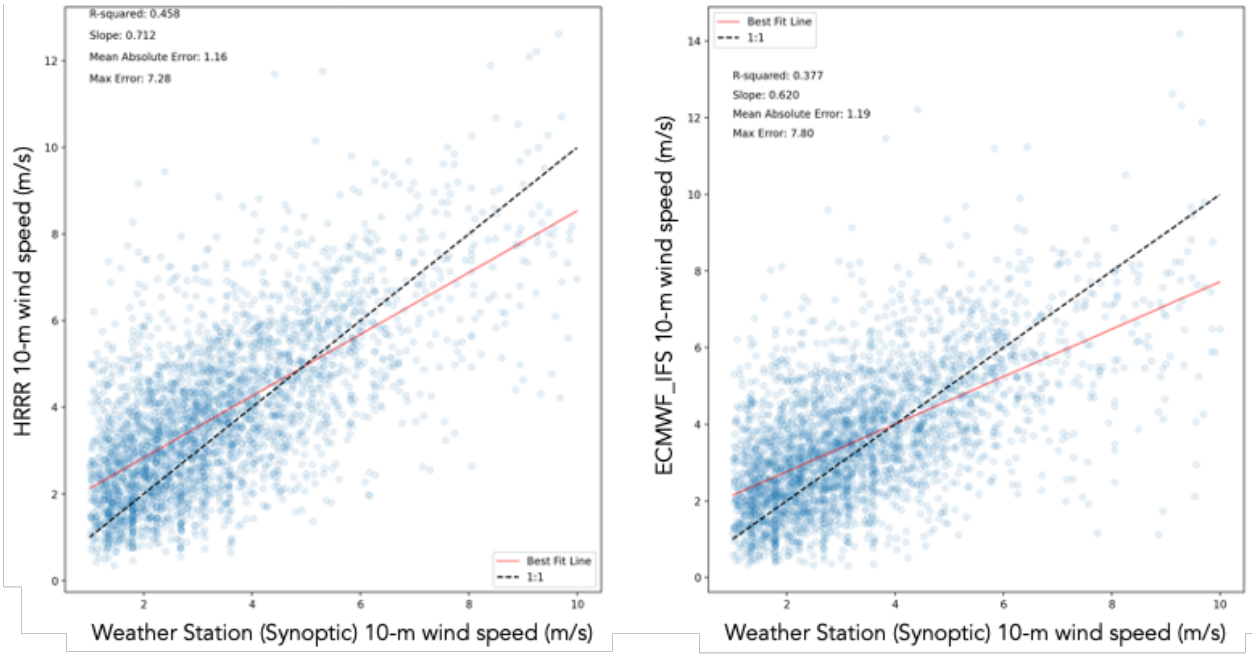

Figure 4*. Comparison of U10 wind speeds from the HRRR and ECMWF_IFS global data products to 10-m weather data pulled from the Synoptic Weather Data API (synopticdata.com).*

In Equation 2, L is estimated as the max distance from the origin point of the plume to another point along the segmented plume's convex hull. For plumes covering large spatial distances, we impose a distance constraint such that the segmented plume mask is clipped to not exceed a 2500 m radial extent from the origin of the plume. Therefore, $L = min\{max(hulldist), 2500\ m\}$. The IME (Equation 1) is also only calculated within this clipped plume mask. This clipping procedure is employed to reduce bias that may enter into IME quantification due to differing surface and meteorological conditions across large plumes, intermittency of the emission rate of the source, and to limit potential merging of multiple plumes downwind of their sources.

To optimize the percentile/crop, we seek to find the best candidate crop/percentile curve from Figure 3 that produces the closest emission estimates against independent or validation datasets. This is best achieved through controlled releases, which are planned to start for EMIT and Tanager satellites in 2024-2025. Aircraft underflights, though not absolute truth, are also suitable for constraining crop/percentile curves. In 2023, Carbon Mapper coordinated aircraft underflights with the Global Airborne Observatory and EMIT in the Permian Basin to estimate EMIT detection capabilities (Ayasse et al., 2024). Three near-simultaneous observations were made between GAO and EMIT (other asynchronous observations also occurred during the

campaign) (Figure 5). We find that we can best match GAO emission rates from EMIT-observed plumes when the dynamic threshold for EMIT is set to the black curve in Figure 3 which would correspond to roughly a 90th percentile applied at a 1000 m crop.

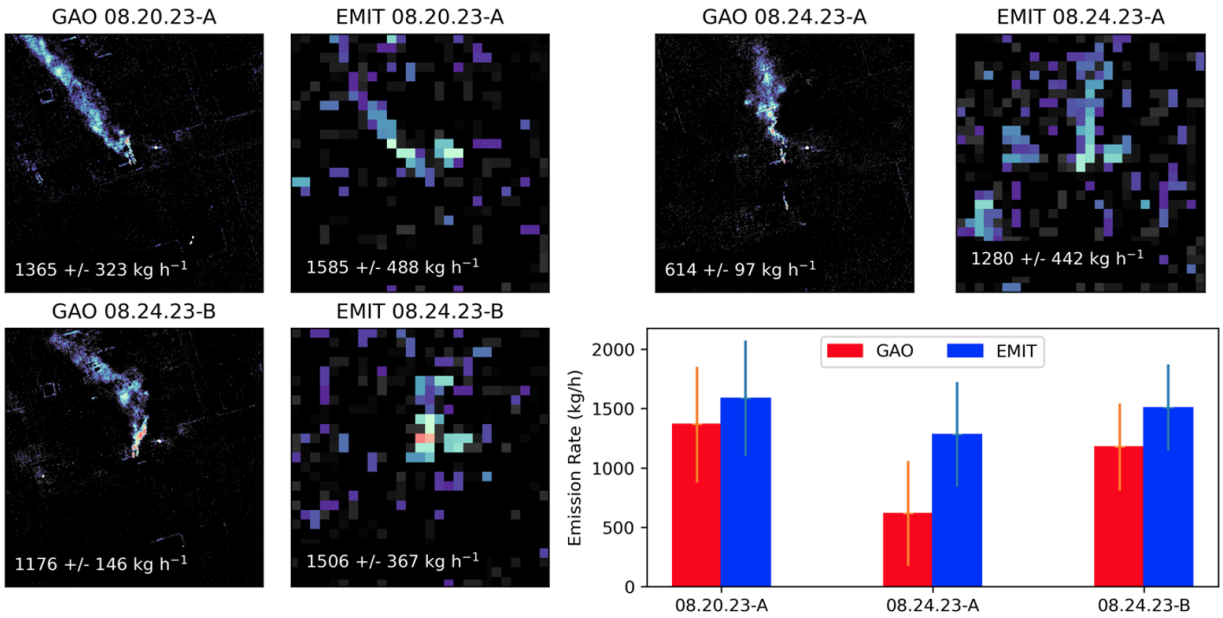

***Figure 5**. Comparison of simultaneously observed plumes by EMIT and Global Airborne observatory (Ayasse et al., 2024).*

Given the relatively sparse validation for EMIT to date, we also test optimized EMIT algorithms against simulated data from the Weather and Research Forecasting Model - Large Eddy Simulation (Varon et al., 2018). These simulations were performed at 50m spatial resolution, smaller, but similar in magnitude to EMIT's pixel resolution (60m) and coarser than Tanager's spatial resolution (30m). Figure 6 shows the result - though the segmentation and quantification procedures result in some scatter, especially at large (>10,000 kg/h) emission rates, the results are unbiased in aggregate; an ordinary least squares (OLS) fit to the data without an intercept (Sherwin et al., 2023) results in R^2 = 0.83 and y = 1.08x.

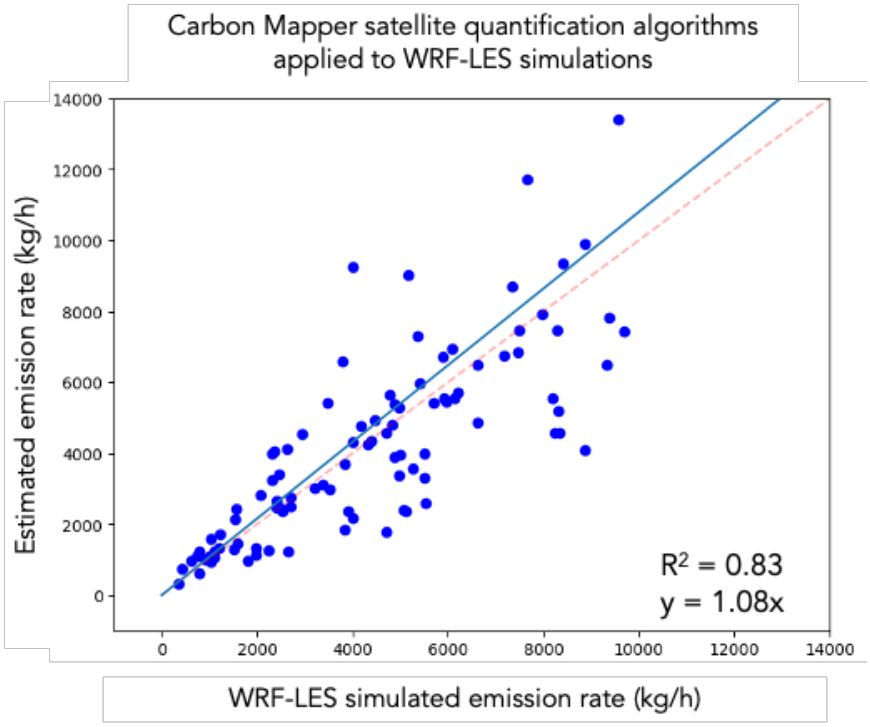

***Figure 6**. Performance of quantification algorithms optimized for EMIT applied to 50-m WRF-LES simulations.*

As an additional ancillary validation procedure, we apply optimized EMIT segmentation and quantification algorithms to PRISMA (Hyperspectral Precursor of the Application Mission, Italian Space Agency) satellite observations. PRISMA observed two controlled release sites in Arizona in 2021 and 2022 (Sherwin et al., 2023; Sherwin et al., 2024). PRISMA is a pushbroom imaging spectrometer that acquires approximately 200 30 x 30 km^2 images per day with roughly 30 m spatial resolution, 10 nm spectral resolution, and 180 signal-to-noise ratio at 2200 nm (Loizzo et al, 2018). Application of Carbon Mapper L2b and L3/L4 algorithms, as described in this and other theoretical basis documents, without any additional tuning for PRISMA, at the controlled release sites, are shown in Figure 7. Though only 5 points, Carbon Mapper L2-L4 algorithms show little bias and high correlation with metered data.

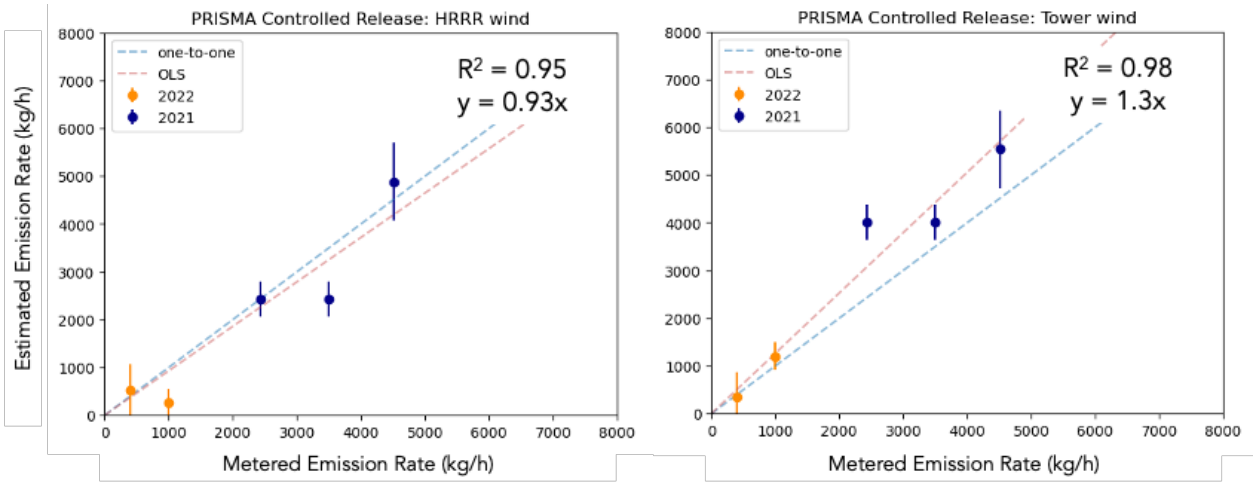

***Figure** 7. Comparison of controlled release data with estimated emission rates from Carbon Mapper application of its L2-L4 algorithms to PRISMA satellite observations, using both reanalysis wind datasets (HRRR) and provided wind datasets recorded from a tower sonic anemometer onsite.*

Finally, we assess the distribution of quantified emissions from EMIT plumes detected and attributed to the oil&gas sector. Multiple previous studies have shown that oil&gas emissions above super-emitter thresholds (100 kg/h) generally follow a power-law distribution (Sherwin et al., 2024). Under this assumption, the peak of an emission rate histogram represents a reliable detection rate (e.g., 90% probability of detection (POD)) for an instrument, and all emissions left of the peak are not representative of the true distribution of emissions but rather the partial detection limit of the instrument. Theoretical and empirical assessments of EMIT's reliable detection range result in estimates of the 90% POD between 900-1200 kg/h (Ayasse et al., 2024). Figure 8 shows the distribution of oil&gas plumes quantified for EMIT using the algorithms described above and shows a peak near the theoretical and empirical derived limits. Between this distributional assessment, controlled release assessment with PRISMA, comparison against WRF-LES simulations, and EMIT aircraft underflight validation, Carbon Mapper L3/L4 algorithms do not show evidence of overwhelming bias in aggregate. Further controlled testing is planned to validate, confirm, and refine as necessary.

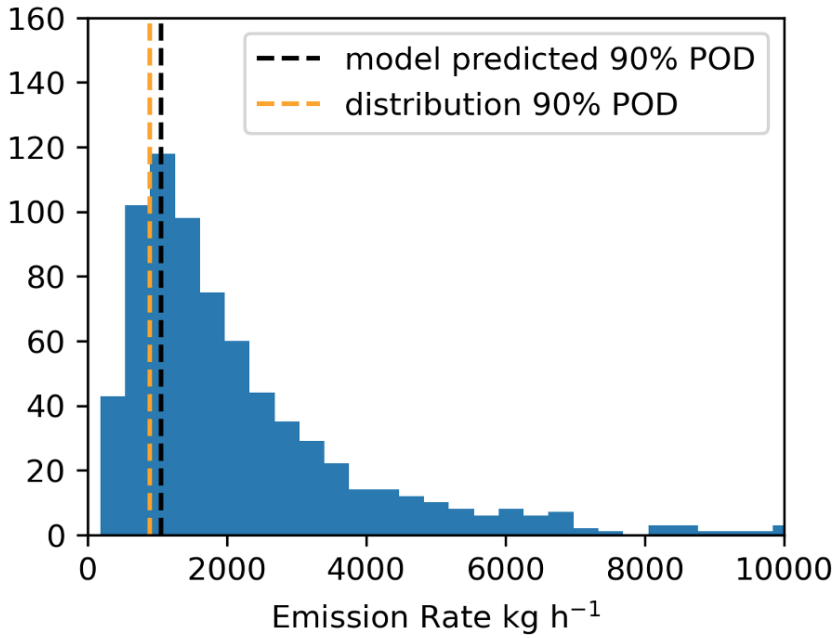

***Figure 8**. Distribution of quantified EMIT oil & gas plumes. The peak of the distribution aligns closely with the theoretically and empirically derived EMIT 90% probability of detection.*

4 Uncertainty Quantification

Uncertainties in emission estimates are calculated by summing in quadrature elements that contribute to variability in emissions:

$$\sigma_q = \sqrt{\left(\frac{\partial Q}{\partial U}\sigma_U\right)^2 + \left(\frac{\partial Q}{\partial IME}\sigma_{IME}\right)^2 + \left(\frac{\partial Q}{\partial L}\sigma_L\right)^2} \quad (4)$$

Where

$$\sigma_{IME} = \frac{\partial Q}{\partial IME}\sigma_N + \frac{\partial Q}{\partial \Omega}\sigma_\Omega \quad (5)$$

In Equation 4 - the $\left(\frac{\partial Q}{\partial U}\sigma_U\right)$ term represents the uncertainty due to wind speed, which we estimate by computing the standard deviation of 10-m wind estimates across the hour before and after the plume detection. The $\left(\frac{\partial Q}{\partial IME}\sigma_{IME}\right)$ term is decomposed into two components, first uncertainty due to masking, which we parameterize as the standard deviation of IME estimates across all segmented plume masks calculated for optimal candidate crop/percentile masks (black curve in Figure 3), and second uncertainty due to the retrieval, which was estimate as the standard deviation of concentration enhancements outside of the segmented plume mask, but

within a 2500 m crop of the plume. Finally, the $(\frac{\partial Q}{\partial L}\sigma_L)$ represents an irreducible uncertainty term due to the pixel resolution of the satellite instrument and how it affects the estimate of plume length *L*. Figure 8 shows the effect of this term: the edge of a plume may manifest as concentration enhancement in a single or multiple pixels depending on the true geolocation of the plume and the spatial resolution of the instrument.

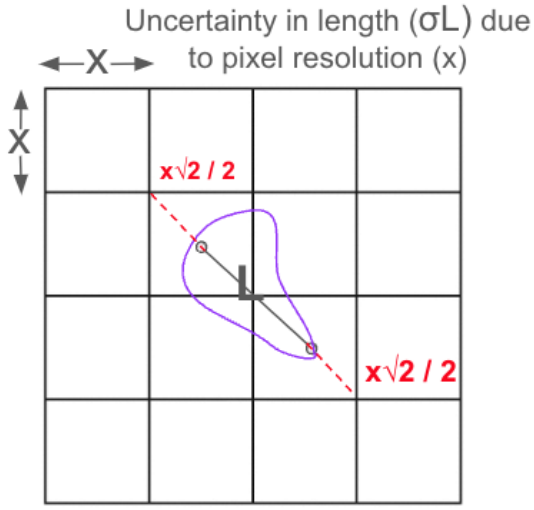

***Figure 9**. Irreducible uncertainty due to pixel resolution for IME quantification approaches.*

As of August 2024, the Carbon Mapper Data Platform only routinely processes EMIT observations for detection and quantification. We anticipate the Tanager algorithms to follow a similar processing chain, with slight adaptation according to differences in instrument performance. The Carbon Mapper Data Platform also processes EMIT (and in the future Tanager) data for CO2 point sources. The quantification algorithms proceed in a similar fashion as described above.

5 Emission Quantification Issues and Quality Checks

Not all instances of plume detections are quantified for emissions. A quality control procedure rules out publishing emission rates if there are issues severely affect quantification, including (1) the overwhelming presence of artifacts in the retrieval; (2) non-standard plume shapes that may violate mass balance assumptions of IME (high wind shear, concentration pooling, etc) or make plume segmentation difficult (e.g. large gaps in plume); (3) plumes that appear at the boundaries

of images; (4) overlapping plumes where emission rates from distinct sources can not be partitioned. In these cases, a detection with no emission rate may be published.

6 Outputs

6.1 Plume image: segmented plume image (GeoTIFF) that provides the boundaries and concentrations (units ppm-m) of pixels that were used for quantification

6.2 ***Plume information***: Provided in Table 1: each plume that has passed quality control procedures may include the following information: geographic coordinate of plume origin (process described in Carbon Mapper Plume Detection Quality Control Protocol), emission rate, emission rate uncertainty, IPCC sector attribution, IME, wind speed, quality flags.

7 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.

Ayasse, A., Cusworth, D.H., Howell, K., O'Neill, K., Conrad, B., Johnson, M.R., Asner, G. and Duren, R., 2024. Probability of Detection and Multi-Sensor Persistence of Methane Emissions from Coincident Airborne and Satellite Observations.

Loizzo, R., Guarini, R., Longo, F., Scopa, T., Formaro, R., Facchinetti, C. and Varacalli, G., 2018, July. PRISMA: The Italian hyperspectral mission. In *IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium* (pp. 175-178). IEEE.

Sherwin, E.D., Rutherford, J.S., Zhang, Z., Chen, Y., Wetherley, E.B., Yakovlev, P.V., Berman, E.S., Jones, B.B., Cusworth, D.H., Thorpe, A.K. and Ayasse, A.K., 2024. US oil and gas system emissions from nearly one million aerial site measurements. *Nature*, *627*(8003), pp.328-334.

Sherwin, E.D., Rutherford, J.S., Chen, Y., Aminfard, S., Kort, E.A., Jackson, R.B. and Brandt, A.R., 2023. Single-blind validation of space-based point-source detection and quantification of onshore methane emissions. *Scientific Reports*, *13*(1), p.3836.

Sherwin, E.D., El Abbadi, S.H., Burdeau, P.M., Zhang, Z., Chen, Z., Rutherford, J.S., Chen, Y. and Brandt, A.R., 2024. Single-blind test of nine methane-sensing satellite systems from three continents. Atmospheric Measurement Techniques, 17(2), pp.765-782.

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.