



# **Methane Plume Detection and Quantification Intercomparison Analysis Based on EMIT Measurements**

February 2026



Annmarie Eldering, Senior Fellow, Climate Data Collaborative, Data Foundation

Contributing authors:

Luis Guanter and Javier Roger, Universitat Politècnica de València  
Joannes D. Maasakkers and Xin Zhang, SRON, Netherlands Institute for Space Research  
Dan Cusworth, Alana Ayasse and Kate Howell, Carbon Mapper, Inc.  
Daniel J. Varon, Massachusetts Institute of Technology, AeroAstro Department  
Jason McKeever and Dylan Jervis, GHGSat, Inc.  
Andrew Thorpe, NASA, Jet Propulsion Laboratory, California Institute of Technology

Data contributors:

Marc Watine, Alba Lorente, Gonzalo Mateo Garcia and Itziar Irakulis Loitxate, International  
Methane Emissions Inventory  
Willow Coleman and Jay Fahlen, EMIT Project, NASA, Jet Propulsion Laboratory, California  
Institute of Technology  
Alexis Groshenry, Julian Akani Guery, Quentin Peyle and Dylan Ellett, Kayyros

Document review:

Israel Lopez-Coto, Stony Brook University and NIST

**Author ORCID iDs**

Annmarie Eldering: 0000-0003-1080-9922  
Luis Guanter: 0000-0002-8389-5764  
Javier Roger: 0000-0003-4851-4185  
Joannes D. Maasakkers: 0000-0001-8118-0311  
Xin Zhang: 0000-0002-1756-6620  
Dan Cusworth: 0000-0003-0158-977X  
Alana Ayasse: 0000-0001-8401-1185  
Daniel J. Varon: 0000-0002-3207-5731  
Dylan Jervis: 0009-0009-5387-635X,  
Andrew Thorpe: 0000-0001-7968-5433  
Marc Watine: 0009-0004-2797-4814  
Alba Lorente: 0000-0002-2287-4687  
Gonzalo Mateo Garcia: 0000-0002-0569-393X  
Itziar Irakulis Loitxate: 0000-0003-3646-4950  
Willow Coleman: 0000-0001-6582-4922  
Jay Fahlen: 0000-0002-1660-8104  
Alexis Groshenry: 0009-0009-8766-3061

**Contact Information**

annmarie.eldering@datafoundation.org

DOI: <https://doi.org/10.15868/socialsector.52028>

## **Abstract**

This report is about an intercomparison experiment that was performed to compare the methane plume detection and quantification based on Earth Surface Mineral Dust Source Investigation (EMIT) remote sensing measurements. The experiment consisted of eight teams analyzing four measurement scenes from EMIT and reporting back results such as plume origins, plume outlines, estimated emission rates, and uncertainties. Comparisons of results across the teams are reported, as well as the aggregate statistics such as box and whisker plots of the emission rates, scatter of the estimates of plume origins, and coefficients of variation (COV) for the results. While the small size of the datasets limits the generalizability of the results, one key outcome of this work was the development of methods for comparison of results, including application of rotation to plume origins to evaluate in the context of the wind field. We also tested the impact of holding variables such as the wind speed constant. In this report, we show that the emission rate estimates have COVs of 30 % to 50 % in the majority of cases, and using the same wind speed across all teams did not drastically change this result. The plume origin analysis showed that the cross wind spread of plume origin results was 50 m to 110 m in all but one case, and across wind spread was 60 m to 120 m in 13 of the 15 cases. The EMIT footprint is 60 m, so the variations were roughly 1 pixel to 2 pixels. In future work, we will build on the methods developed here and include measurements over controlled releases that will enable the bias of the emission estimates to be assessed.

## **Keywords**

Methane plumes; remote sensing: hyperspectral; emissions quantification



## Table of Contents

|                                                                                                             |           |
|-------------------------------------------------------------------------------------------------------------|-----------|
| <b>1. Background .....</b>                                                                                  | <b>1</b>  |
| <b>2. Overview of remote sensing of methane plumes common practices summarized from NIST IR 8575: .....</b> | <b>3</b>  |
| <b>3. Overall objectives of inter-comparison and experiment design: .....</b>                               | <b>6</b>  |
| 3.1. The participants .....                                                                                 | 6         |
| 3.2. The EMIT dataset: .....                                                                                | 7         |
| 3.3. Integrated mass enhancement approach for emission rate estimates: .....                                | 8         |
| 3.4. Experiment design .....                                                                                | 9         |
| 3.5. The data submitted by teams .....                                                                      | 10        |
| <b>4. Results .....</b>                                                                                     | <b>12</b> |
| 4.1. Plume detection results .....                                                                          | 12        |
| 4.2. Plume origin results .....                                                                             | 13        |
| 4.2.1. Origins in latitude and longitude.....                                                               | 13        |
| 4.2.2. Rotation of plumes .....                                                                             | 14        |
| 4.2.3. Origin spread in along with and cross wind direction: .....                                          | 15        |
| 4.3. Plume segmentation (delineation) results.....                                                          | 17        |
| 4.4. Emission rate estimates and key variables results.....                                                 | 18        |
| 4.4.1. Initial comparison across all cases .....                                                            | 18        |
| 4.4.2. Emissions rate estimates for two cases for two experiments.....                                      | 19        |
| 4.4.3. Formulations of U.....                                                                               | 20        |
| 4.4.4. Comparison when wind speed is held constant.....                                                     | 21        |
| 4.5. Coefficient of Variation for emission estimates.....                                                   | 22        |
| <b>5. Conclusions .....</b>                                                                                 | <b>25</b> |
| <b>6. Community Efforts and Next Steps .....</b>                                                            | <b>26</b> |
| <b>References .....</b>                                                                                     | <b>28</b> |
| <b>Appendix A. Acronyms &amp; Abbreviations.....</b>                                                        | <b>30</b> |

## List of Tables

|                                                                                          |           |
|------------------------------------------------------------------------------------------|-----------|
| <b>Table 1: Summary of participating teams .....</b>                                     | <b>6</b>  |
| <b>Table 2: Overview of EMIT scenes used in intercomparison activity .....</b>           | <b>7</b>  |
| <b>Table 3: Plumes that are in the box and whisker plots (Figures 7,8, and 11) .....</b> | <b>17</b> |

## List of Figures

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1: Typical analysis steps for methane plume detection and quantification process. The key steps are in black, and useful output products are in blue font. Additional external inputs are in the red circles. Concept credit: Dan Cusworth, Carbon Mapper, used with permission.....                                                                                                                                                                       | 4  |
| Figure 2: Steps showing the process of finding the methane enhancement and then the plume. Credit: Dan Cusworth, Carbon Mapper, used with permission. From Carbon Mapper ATBD version 1.0.0. ....                                                                                                                                                                                                                                                                 | 5  |
| Figure 3: RGB images of the 4 EMIT scenes used in the intercomparison. Provided by Andrew Thorpe, NASA JPL/Caltech.....                                                                                                                                                                                                                                                                                                                                           | 8  |
| Figure 4: Plume extent images for the 4 scenes. Note the different length scales on each panel, ranging from 3 km to 10 km. Colors are used to differentiate the plumes from different teams.....                                                                                                                                                                                                                                                                 | 13 |
| Figure 5: Map with plume outlines in a part of scene 2 (left panel). The orange arrow points to the third plume in scene 2. The scale bar in the upper left is 3 km. Map of the set of plume origins identified for the third plume in scene 2 (right panel). A scale bar (showing 40 m) is included in the lower right of the map. The point of the yellow arrow in the left panel points to the location of the northernmost red marker on the right panel..... | 14 |
| Figure 6: The coordinate rotation of UTM northing ( $y$ ) and UTM easting ( $x$ ) to along-wind direction ( $y'$ ) and cross wind ( $x'$ ) given the wind blowing into bearing angle ( $\theta$ ).....                                                                                                                                                                                                                                                            | 15 |
| Figure 7: Box and whisker plots of the plume origin data for 15 plumes measured across all 4 scenes. The data has been rotated into the along-wind and cross-wind directions, and the statistics are for differences from the mean in meters. See Table 3 for legend to connect bars to scenes and plume numbers. ....                                                                                                                                            | 16 |
| Figure 8: Box and whisker plot of the characteristic plume length. The order of the bars is listed in Table 3. ....                                                                                                                                                                                                                                                                                                                                               | 17 |
| Figure 9: Box and whisker plots for the key quantities. Panel (a): Emission rate estimates, $Q$ , Panel (b): Uncertainty on emission rate estimates (1 sigma), Panel (c): 10 m wind speed, Panel (d): effective wind parameter, Panel (e): number of data points included .....                                                                                                                                                                                   | 19 |
| Figure 10: The emission rate estimates ( $\text{kg}\cdot\text{h}^{-1}$ ) with uncertainty bars (1 sigma) for the single plumes of scene 1 (left panel) and scene 4 (right panel). The orange bars are for results on the JPL L2, and the dark blue bars are for the results of the mix and match experiment.....                                                                                                                                                  | 20 |
| Figure 11: Different formulations of the IME-method effective wind speed $U_{\text{eff}}$ as a function of 10-m wind speed $U_{10}$ used by the participating teams. Solid lines define the plume length scale $L$ for the IME method as the square root of the plume mask area. Dashed (.....                                                                                                                                                                    | 21 |
| Figure 12: The emission rate estimates ( $\text{kg}\cdot\text{h}^{-1}$ ) with uncertainty bars for the plume of scene 1 (left panel) and scene 4 (right panel) when wind speeds are constant across all teams. The orange bars are for results on the JPL L2, and the blue bars are for the resu.....                                                                                                                                                             | 22 |
| Figure 13: Coefficient of Variation (COV) for emissions estimate analysis experiments. Table 3 provides the details of the scenes and plumes for each bar. ....                                                                                                                                                                                                                                                                                                   | 23 |

## **Author Contributions**

**Annmarie Eldering:** Writing – original draft, conceptualization, data curation, project administration. **All Contributing Authors:** Investigation, data curation, formal analysis, writing – reviewing and editing. **All Data Contributor Authors:** Investigation and data curation. **Israel Lopez-Coto:** Writing – reviewing and editing



## 1. Background

In recent years, there has been rapid growth and advancement in technologies for remote sensing of methane plumes from aircraft and space. The global reach and inherent spatial sampling capabilities of on-orbit instruments make them uniquely suited for consistent, repeatable surveys across regions and borders. These systems enable the detection and quantification of methane plumes, the identification of major sources, and the attribution of emissions to specific geographic locations.

Plume-based observations are particularly valuable for rapidly identifying and quantifying the largest sources of emissions, enabling timely mitigation. While the share of total anthropogenic methane emissions attributable to such sources remains uncertain, regional estimates indicate they can be significant. Currently, airborne measurements are the only set of data with enough measurements to assess regional emissions, but the methods are expected to be applied to satellite data when larger measurement datasets are available. Based on aircraft measurements in the Permian Basin in west Texas and southeastern New Mexico, sources exceeding  $100^1 \text{ kg}\cdot\text{h}^{-1}$  were estimated to contribute 75 % of total oil and gas emissions in 2019 (Sherwin et al., 2024b), with other analyses suggesting a lower share of 23 % in 2021 (Williams et al., 2025). In basins such as the Uinta Basin in Utah (Sherwin et al., 2024b) and the San Joaquin Basin in California (Williams et al., 2025), the corresponding contributions range from 10 % to 50 %. Outside these intensively studied regions, the fraction of emissions observable with aircraft and satellite measurements remains poorly constrained. Even at the lower end of these estimates, plume-based observations represent a critical tool for emissions characterization and mitigation. Broader assessments of sector-wide or regional emissions, particularly from diffuse sources such as agriculture, will require complementary approaches such as area-based flux mapping using instruments such as MethaneSAT and TROPOMI (Schuit et al., 2023).

Despite the rapid growth in observational capacity, challenges remain. Divergent emission rate estimates, opaque methodologies, and inconsistent validation approaches can erode confidence in remote sensing-based emissions data. The emergence of non-public-sector missions using proprietary methods—often without full transparency across the data chain—further highlights the need for community-accepted practices to ensure traceability, comparability, and scientific credibility.

To address this need, the Committee on Earth Observation Satellites (CEOS) and National Metrology Institutes (NMIs), and other researchers have developed a document to articulate commonly accepted approaches for quantifying methane emission rates based on observed plumes. It provides guidance spanning from Level 0/1 radiances, to Level 2 concentrations, to Level 4 emission rate estimates, and includes current practices for validation and quality assessment. That collaborative report (NIST IR 8575, written by authors from CEOS, NIST, NPL, and LBNL, Worden et al., 2025 and CEOS, <https://doi.org/10.5281/zenodo.17047789>) focused

---

<sup>1</sup> Throughout this report, all emission rate estimates are  $\text{kg h}^{-1}$  of methane ( $\text{CH}_4$ )

on documenting current practices and the key steps in deriving emission rate estimates and plume origins from spectrally resolved measurement datasets. A brief overview of that work is presented in Section 2: [Overview of common practices for remote sensing of methane plume, summarized from NIST IR 8575.](#)

This report builds on that work, discussing an intercomparison of analyses of four EMIT measurement scenes with participation from eight independent teams, each using their own analysis workflow. The measurement scenes selected are from across the globe, and contain uncharacterized methane plumes, referred to as wild plumes. The scenes do not contain controlled releases, where a known amount of methane is emitted and which can be used to assess bias of results. Such an experiment will be the focus of a future effort. Analysis of the EMIT dataset is useful for comparing plume detection across teams, analysis of variation in emission rate estimates across teams, and assessments of the key factors driving variations. This analysis relied on the EMIT Level-1B (L1B; geolocated radiance measurements) and Level-2 (L2; retrieved methane column concentrations) datasets. The L2 data were generated using a matched filter approach.

This document begins with an overview of the NIST IR 8575 and then describes the EMIT intercomparison experiment. Results are presented for a range of analyses, and finally conclusions and a discussion of the path forward are provided.

## 2. Overview of remote sensing of methane plumes common practices summarized from NIST IR 8575:

The field of remote sensing of methane plumes is expanding rapidly and the language and methodologies of analysis are wide ranging. A more coherent and consistent approach across teams would improve the interoperability of the data and clarity for end users. This needs to be balanced with the benefits that come with a range of development paths and the freedom of innovation. At a minimum, clear documentation and consistent usage of terminology is needed. In January of 2024, NIST organized a meeting to begin gathering the community working in the areas of remote sensing measurements and the detection and quantification of methane plumes in remote sensing measurements. Following the January 2024 meeting, a document was created to capture current practices for data analysis and validation approaches, and to identify some of the key needs looking forward. This document has been published NIST IR 8575 (<https://doi.org/10.6028/NIST.IR.8575>) and through CEOS (<https://doi.org/10.5281/zenodo.17047789>).

Figure 1 (drawn from that report) captures the typical analysis steps in the methane plume identification and quantification process. The input data (preceding black box and brick circles) and output data (in blue) for each step are annotated. That report was focused on the analysis process, and highlighted that teams follow very similar overall processes, but at each step there are typically differences. For example, in the concentration retrieval step (box 3 of Figure 1), the details of the formulation of the matched filter for concentration retrieval vary. Similarly, the plume detection and plume segmentation (box 4 of figure 1, and as illustrated in more detail in Figure 2) are implemented in a number of different ways. Plume detection, as illustrated in Figure 2, may include applying a threshold to a concentration enhancement field and then identifying connected pixels. The team's process illustrated in Figure 2 limits the distance from the plume origin for pixels included in the plume segmentation, so we see that the panels in the bottom row of Figure 2 use a smaller scale and only include pixels within 25 pixels of the origin. This report builds from the current practices report that was focused on processes and looks at numerical results across teams.

Key output/metrics that can be derived from various stages of workflow

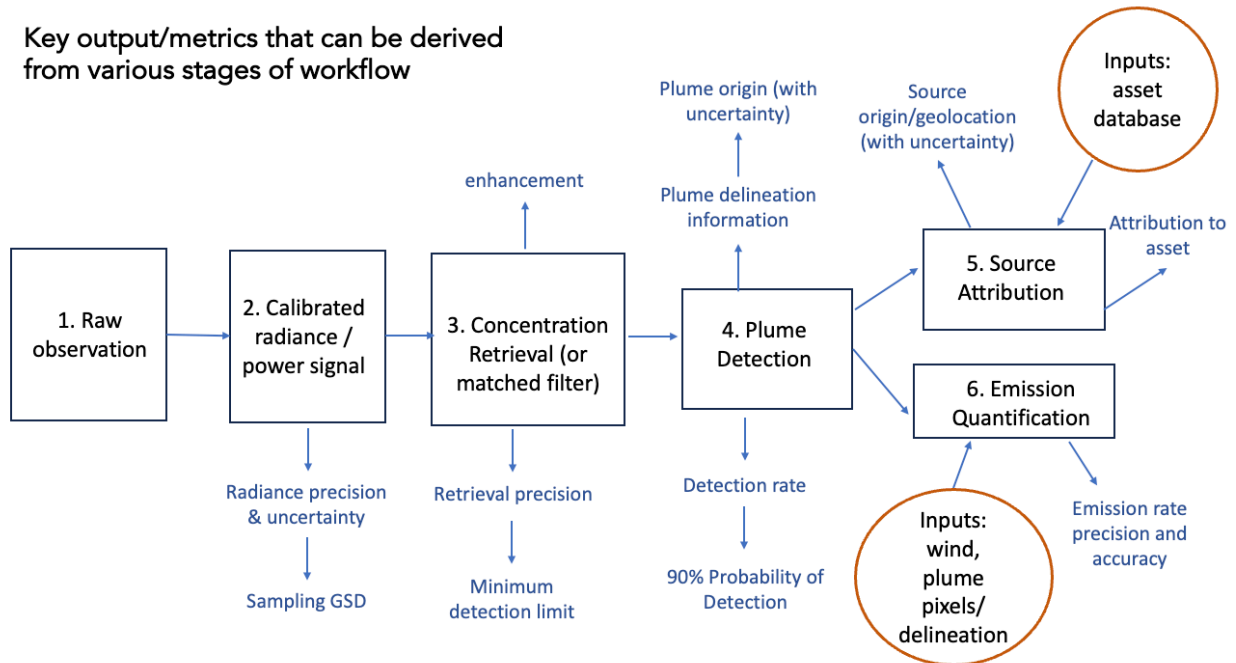


Figure 1: Typical analysis steps for methane plume detection and quantification process. The key steps are in black, and useful output products are in blue font. Additional external inputs are in the red circles. Concept credit: Dan Cusworth, Carbon Mapper, used with permission.

### Example of plume segmentation process

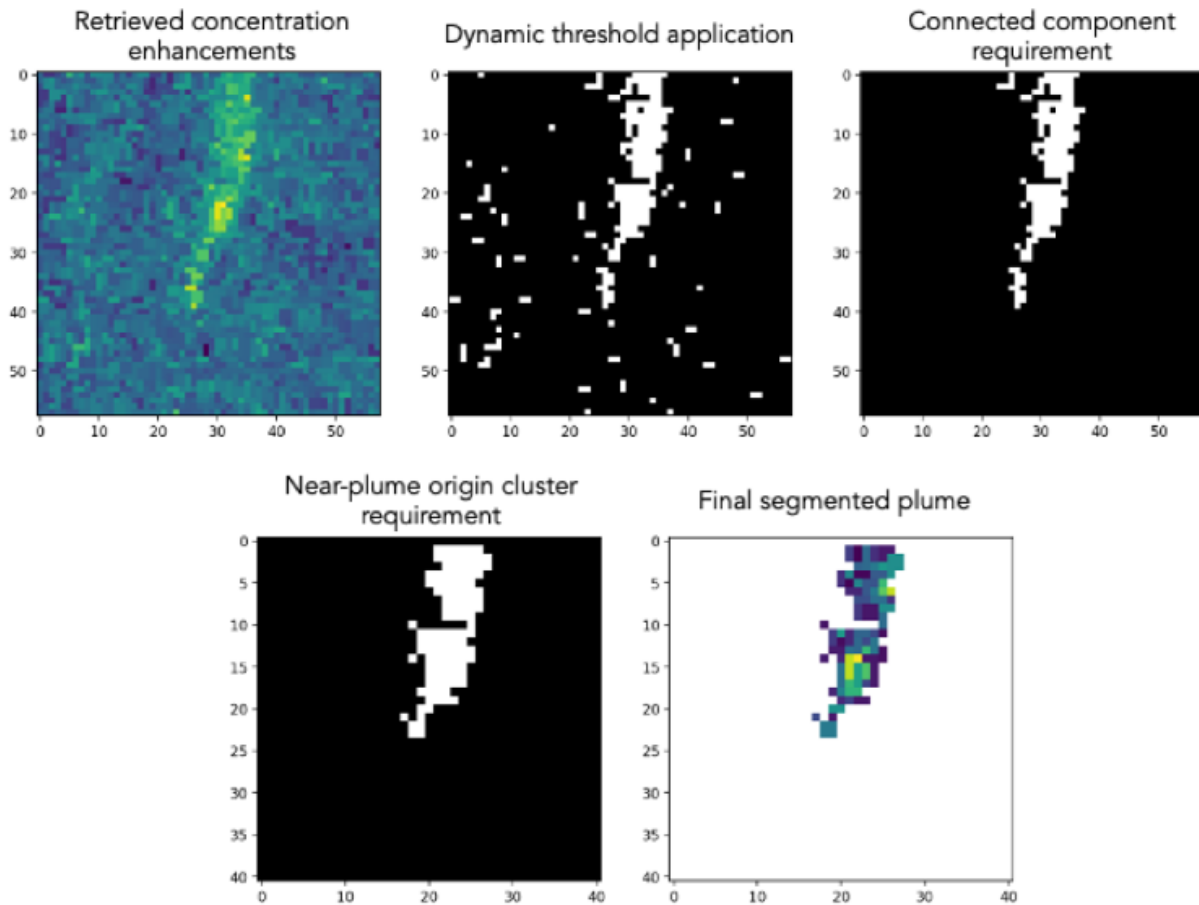


Figure 2: Steps showing the process of finding the methane enhancement and then the plume. Credit: Dan Cusworth, Carbon Mapper, used with permission. From Carbon Mapper ATBD version 1.0.0.



### 3. Overall objectives of inter-comparison and experiment design:

To move beyond documentation of the analysis process, an inter-comparison exercise was developed. The objectives of this intercomparison project are to (1) develop metrics to assess the variability in the plume origin and emission rate estimates across the teams, (2) evaluate whether the variability of emission rate estimates is consistent with the expected uncertainty, and (3) gain insights about what elements of the remote sensing analysis chain contribute to the variability across teams.

To achieve these objectives, we had many teams analyze the same set of remote sensing measurements and compared results. The data requested included plume origin location and emission rate estimates with uncertainty. We also performed analysis that tried to make terms (such as wind speed) constant across the teams to isolate the impact of a single term on the emission rate estimation results. Given the variations in analysis processes that exist across the teams, we would expect variation in the results.

The analysis requested from the teams was guided by our investigation goals, but with some limitations on how well we could control certain variables or analysis steps. The strategy for developing the experimental design is derived from the key analysis equations, discussed below. The teams that participated in this work have the capability to use the integrated mass/methane enhancement (IME) approach for estimating emission rates from remote sensing data, so the experiment focused on that formulation. Connecting that equation to the experiment design is discussed in more detail in Section 3.4.

#### 3.1. The participants

Eight teams that analyze remote sensing data participated in the EMIT intercomparison. Participating teams, whether they have developed an independent L2 processed data set for EMIT, and key citations are captured in Table 1.

**Table 1: Summary of participating teams**

| Team name     | Has developed independent L2 for EMIT | Key citations (methodology, validation etc) |
|---------------|---------------------------------------|---------------------------------------------|
| JPL/EMIT      | N/A                                   |                                             |
| Carbon Mapper | yes                                   | Carbon Mapper ATBD version 1.0.0. (2024)    |
| GHGSat        | no                                    | Varon et al. (2018)                         |
| SRON          | yes                                   | Zhang et al. (2024)                         |
| Kayrros       | yes                                   |                                             |

|          |     |                                                                                         |
|----------|-----|-----------------------------------------------------------------------------------------|
| D. Varon | no  | Varon et al. (2018)                                                                     |
| UPV      | yes | Carvalho et al. (2019), Guanter et al. (2021), Pei et al. (2023), Guanter et al. (2024) |
| IMEO     | no  |                                                                                         |

### 3.2. The EMIT dataset:

A set of four EMIT scenes was selected for the intercomparison exercise. The goal was to have a range of scenes, including relatively easy-to-analyze scenes (bright surface, isolated plume) and more complex scenes (lower surface reflectivity, multiple plumes). We also sought to have a range of wind speeds in the scenes. Table 2 provides an overview of the four scenes selected, and Figure 3 contains RGB images of each scene.

**Table 2: Overview of EMIT scenes used in intercomparison activity**

| scene | EMIT identifier     | Time of measurement        | Location and longitude and latitude         | Scene Properties and Spatial Plume characteristics                                                                         | Estimated emission rate range                           | Wind speed                                     |
|-------|---------------------|----------------------------|---------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|------------------------------------------------|
| 1     | emit20230223t063033 | Feb. 23 2023, 06:30:33 UTC | Turkmenistan<br>Lon: 61.5°E,<br>Lat: 40°N   | Bright, homogenous surface. One narrow plume, ~ 30 km in length.                                                           | ~ 4000 kg·h <sup>-1</sup> to 6000 kg·h <sup>-1</sup>    | 3.0 m·s <sup>-1</sup> to 3.5 m·s <sup>-1</sup> |
| 2     | emit20220827t060753 | Aug. 27 2022, 06:07:53 UTC | Turkmenistan<br>Lon: 61.5°E,<br>Lat: 36.5°N | Bright scene, some surface inhomogeneity. Many (12 to 16 identified by teams), ranging from a few kms to ~20 km in length. | < 1000 kg·h <sup>-1</sup> to > 10000 kg·h <sup>-1</sup> | 3.5 m·s <sup>-1</sup> to 4.5 m·s <sup>-1</sup> |
| 3     | emit20230612t162103 | Jun. 12 2023, 16:21:03 UTC | Wyoming<br>Lon: 111°W,<br>Lat: 41.5°N       | Heterogeneous surface. Two broad plumes, one ~2 km in length, one ~20 km in length.                                        | ~10000 kg·h <sup>-1</sup> to ~15000 kg·h <sup>-1</sup>  | 1.5 m·s <sup>-1</sup> to 2.0 m·s <sup>-1</sup> |

|   |                     |                              |                                                                   |                                                                                              |                                                         |                                                      |
|---|---------------------|------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------------------------------------|---------------------------------------------------------|------------------------------------------------------|
| 4 | emit20240403t165815 | Apr. 3 2024,<br>16:58:15 UTC | Simeprodeso,<br>Nuevo León,<br>México<br>Lon: 100°W,<br>Lat: 26°N | Urban,<br>heterogeneous<br>surface. One<br>blob-like plume,<br>~5 km in length<br>and width. | ~6000 kg·h <sup>-1</sup> to<br>~8000 kg·h <sup>-1</sup> | 1.6 m·s <sup>-1</sup><br>to 2.0<br>m·s <sup>-1</sup> |
|---|---------------------|------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------------------------------------|---------------------------------------------------------|------------------------------------------------------|

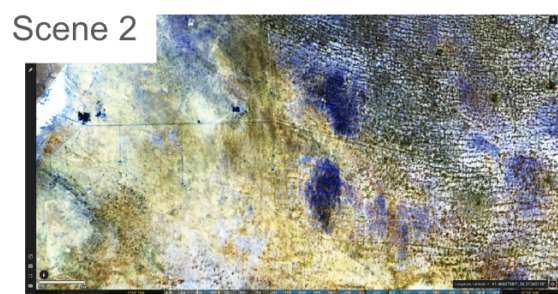
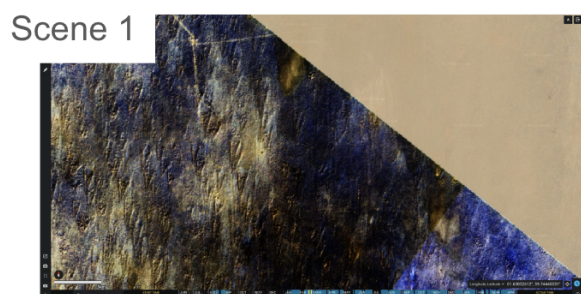


Figure 3: RGB images of the 4 EMIT scenes used in the intercomparison. Provided by Andrew Thorpe, NASA JPL/Caltech

### 3.3. Integrated mass enhancement approach for emission rate estimates:

The formulation for Integrated Mass Enhancement approach (IME) is presented briefly here, to help explain the experiment design. More discussion can be found in the report NIST IR 8575 (Figure 11 and discussion). The IME analysis begins with the L2 data field. These are concentration enhancement maps derived with the matched filter approach (Kim et al., 2016, Foote et al., 2020). A simplified explanation of the approach is that a reference spectrum is developed from background pixels, which are assumed to have no methane enhancement. A spectrum is then modeled to estimate the change in radiance if a known amount of methane was present. With small enhancements, it is assumed that the radiance changes are linear, so the enhancement of methane present in each non-background measured radiance can be

estimated by comparing to the background spectrum plus a scaled modeled methane spectrum. The scaling times the modeled methane enhancement is the estimated methane column enhancement. There are approaches that are iterative, non-linear and also ones that use more details of the background data to minimize the impact of interferants. The field of methane enhancements at each pixel that comes from the matched filter is referred to as the L2 dataset.

The IME method (Frankenberg et al., 2016; Varon et al., 2018) relates the instantaneous emission rate  $Q$  [ $\text{kg}\cdot\text{s}^{-1}$ ] of an observed plume to the total observed plume mass ( $IME$ ) [ $\text{kg}$ ] as

$$Q=(U\times IME)/L \quad (1)$$

where  $U$  [ $\text{m}\cdot\text{s}^{-1}$ ] and  $L$  [ $\text{m}$ ] are representative wind speeds and length scales for the plume, respectively. The  $IME$  is readily computed from the sum of methane column enhancements  $\Delta\Omega_i$  [ $\text{kg}\cdot\text{m}^{-2}$ ] in  $i=1\dots M$  pixels across the detectable plume extent delineated by a binary plume mask, as

$$IME = \sum_M \Delta\Omega_i A_i \quad (2)$$

Where  $A_i$  [ $\text{m}^2$ ] is the area of the  $i^{\text{th}}$  plume pixel. The methane column enhancements are from the L2 matched filter data. The length scale  $L$  and wind speed  $U$  should be viewed as operational parameters relating  $IME$  to  $Q$ .  $L$  can be computed from the plume mask in a variety of ways, including as the length of the plume or the square root of the plume mask area. The latter definition is often chosen for its straightforward application to plumes of any shape and size, and because the loss of mass from the plume reflects not only transport to a fixed distance downwind but also turbulent diffusion along the plume edges. The former definition may be considered more physically intuitive.  $U$  is often denoted as an effective wind speed  $U_{\text{eff}}$  that can be parameterized to match a chosen definition of  $L$ .  $U_{\text{eff}}$  is not directly observable but can be related to an observable wind speed such as the wind at 10 m above the surface ( $U_{10}$ ), with relationships that may vary across teams.

### 3.4. Experiment design

The first aspect investigated was the detection of plumes. Remote sensing measurements provide useful information about the presence of methane plumes and can also be used to develop estimates of the emission rates, or quantification of the plumes. We do not know the true emission locations or rates for the scenes that were studied, so for plume detection, we were limited to observations about how many plumes each team located, the location of the plume origin, and the variability across teams. We also assessed the scatter in the context of the plume and wind directions (see Section 4.2.2. [Rotation of plumes:](#) and Section 4.3.2. [Origin spread in along with and cross wind direction](#)).

To investigate whether the variability of emission rate estimates across teams is largely driven by one factor (for example, wind speed), we developed approaches to hold one factor constant across all teams and look at the impact on  $Q$ . Although the equations for  $Q$  and for IME only have a few terms, it is not always simple to hold a term constant across teams. For the  $U_{\text{eff}} = f(U_{10})$  term, this is mechanistically straightforward. If we have the equation that each team uses to transform  $U_{10}$  to  $U_{\text{eff}}$ , we can directly scale results to isolate that factor and evaluate the results applying the same  $U_{10}$  across all teams.

For other elements, this is not as easy. The IME depends on both the enhancement field (the L2 product), and on which pixels are included in the plume, or the plume segmentation. The plume segmentation also impacts the L term. We can theoretically conduct an experiment where the L2 field is held constant across teams, but which pixels are included will vary. Similarly, the term L is also entangled with the IME term – the outline of the plume, or the plume segmentation process impacts both L and the pixels (mass) included in the plume.

In this report, tests with constant  $U_{10}$  across the teams and constant L2 are presented. But it is important to note that each team might use a different parameterization and simulation dataset for  $U_{\text{eff}}$ , which is based on different criteria for masking. Using the same  $U_{\text{eff}}$  calibration formula along with a plume mask that is not appropriate for that  $U_{\text{eff}}$  calibration would not make sense.

There are no analyses in this report that have the same L or the same IME for all teams. We will aim to design such an experiment for future intercomparison exercises. Additionally, analyses of controlled releases with ancillary met data, sophisticated simulation activities, or other field experiments may provide more insight into questions of the plume segmentation, extent of plume, etc.

### **3.5. The data submitted by teams**

Each team was asked to submit their analysis results of the EMIT scenes. There were two experiments that were conducted by the teams and a third that was created from those datasets.

- All eight teams were asked to apply their workflow to the JPL L2 field. They were asked to submit numeric results as well as geotiffs for all detected emission plumes. The requested numeric results were latitude and longitude of plume origin,  $U_{10}$  wind speed,  $U_{\text{eff}}$  wind speed if different,  $L$ ,  $IME$ ,  $Q$ , and  $Q$  uncertainty. We also requested geotiffs of their enhancement fields and plume outlines, and the equation that they used to relate  $U_{10}$  and  $U_{\text{eff}}$ . All eight teams participated this experiment, which let us examine the variability of results when the underlying L2 was held constant.
- There were four teams (in addition to JPL) with the capability to develop their own L2 or enhancement fields. They were asked to apply their workflow process to analyze their

own L2 from the EMIT scenes and provide the same data as described above. This provides insight into the impact of L2 on the variability of results.

- For the third experiment, statistics were calculated for what we call the mix and match case. This reflects data that end users might obtain, so it includes results for the 4 teams that generate their own L2 data, and the other 4 teams that produce results based on the JPL L2.

As described in the sections that follow, a number of analyses were performed on the results of these experiments. These include analysis of the variation in plume origin, in a latitude/longitude framework and in an along-wind/cross-wind framework, determination of the coefficient of variation (COV) of  $Q$  for each experiment, and then reanalysis of COV for all experiments with  $U_{10}$  held constant across teams. The COV is a standardized measure of dispersion and is calculated by dividing the standard deviation by the mean of the population of data. Note that not every team submitted results for every plume, so in the analysis that follows, there will be fewer than 8 sets of results in some cases.

## 4. Results

In the results section that follows, we present summaries and discussion for a number of aspects. The first section focusses the plume detection, followed by analysis of the plume origin. There is a brief discussion of plume segmentation results across the teams. The discussion of the emission estimates is longer, including discussion of results as submitted, uncertainty estimates, and how the results compare when one element of the IME equation is held constant across the teams.

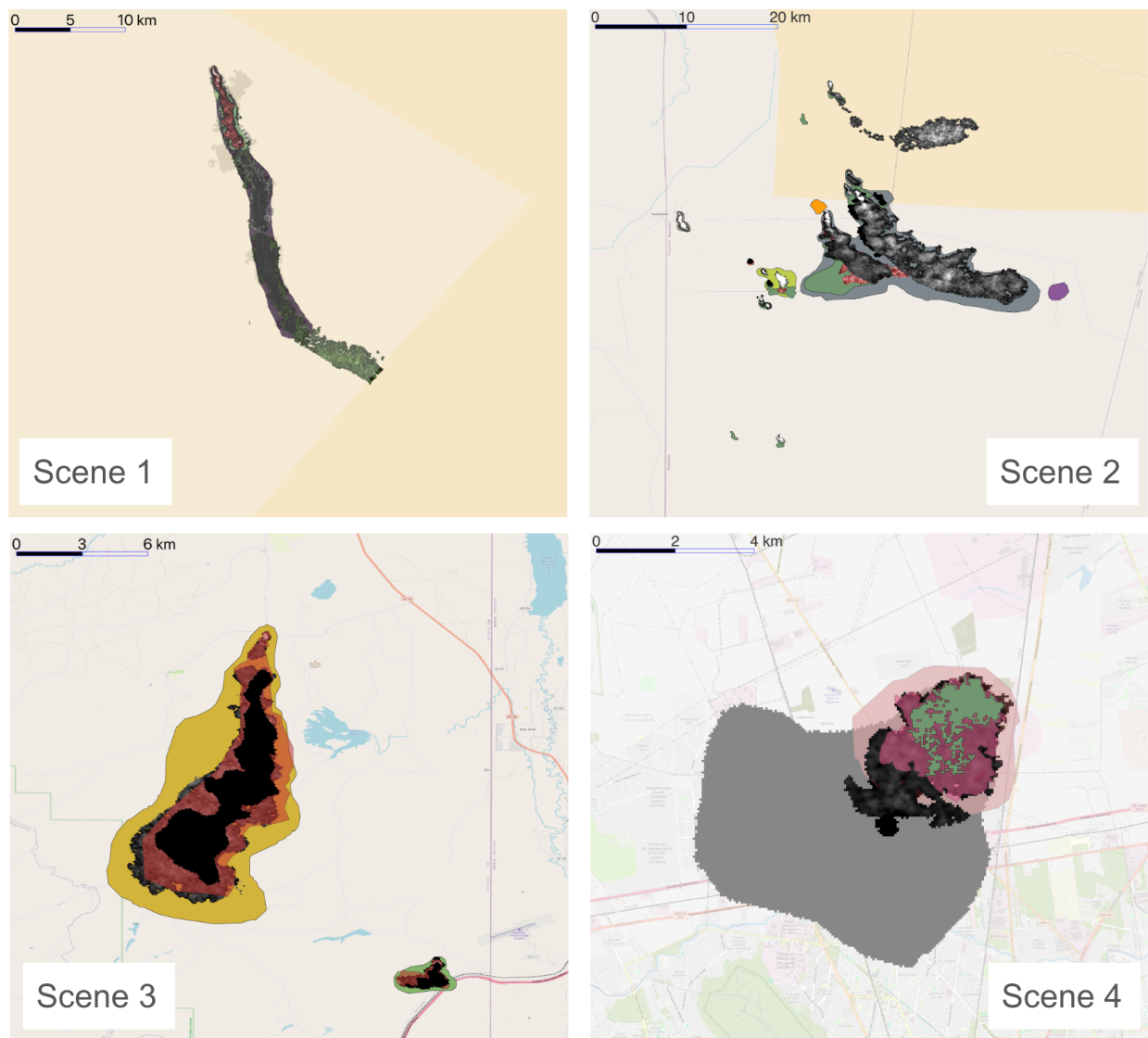
### 4.1. Plume detection results

Analysis teams evaluated the L2 (matched filter) results and identified where there are plumes. Therefore, our first step of the intercomparison exercise was to examine what plumes were detected across the eight teams when applied to the JPL L2. Figure 4 shows a collection of the plume outlines. These figures are designed to show, with overlapping colors and textures, the range of plume outlines submitted. They are not intended to connect a particular plume to a particular team, but they are useful to see the range of results.

For two of the scenes (scenes 1 and 4), where there was one well-defined plume, results were overall consistent across all of the teams, with each team detecting a single plume. The extent of the plume for scene 1 varied significantly across teams, likely because of the large size. In scene 1, some teams identified the plume extending 5 km or 10 km from the plume origin, and others showed it extending 20 km or 30 km away. For scene 3, with a very large plume, there were differences in plume origin location and in the plume extent. Some teams identified two plumes (one at the northernmost point and another a bit south), and some described this as one large, connected plume. Some teams also identified a smaller plume to the southeast of the larger plume. Similarly, in scene 4 the area impacted by the plume was very different across teams. Scene 2 was particularly complex, and some teams identified 12 plumes, while others identified 16. This made the analysis for the intercomparison exercise complex, as we had to pair the results to make sure we were comparing the same plume across all teams.

These plume images also highlight the range of visual plume images that end users may see, which can be a point of confusion. The plume segmentation and plume emission equation elements (L & IME) are interdependent. We have not yet devised an approach to quantitatively assess how the area of the plume impacts the estimate of emission rates, so we cannot advise end users on that topic yet. There are many other factors in the analysis process beyond plume delineations that also impact emission rate quantification, such as how wind speed is used and the L2 field itself. It is also important to note that some teams have two plume delineations - one for visualization and one for quantification. Also, some teams use large eddy simulation (LES) models for the determination of the  $U_{\text{eff}}$  to  $U_{10}$  relationship, and there may be limits put on the extent of the plume during the delineation process to keep it consistent with the simulation set (Varon et al., 2018). The process of plume delineation is discussed briefly again in Section 4.3. ([Plume segmentation \(delineation\) results:](#)).





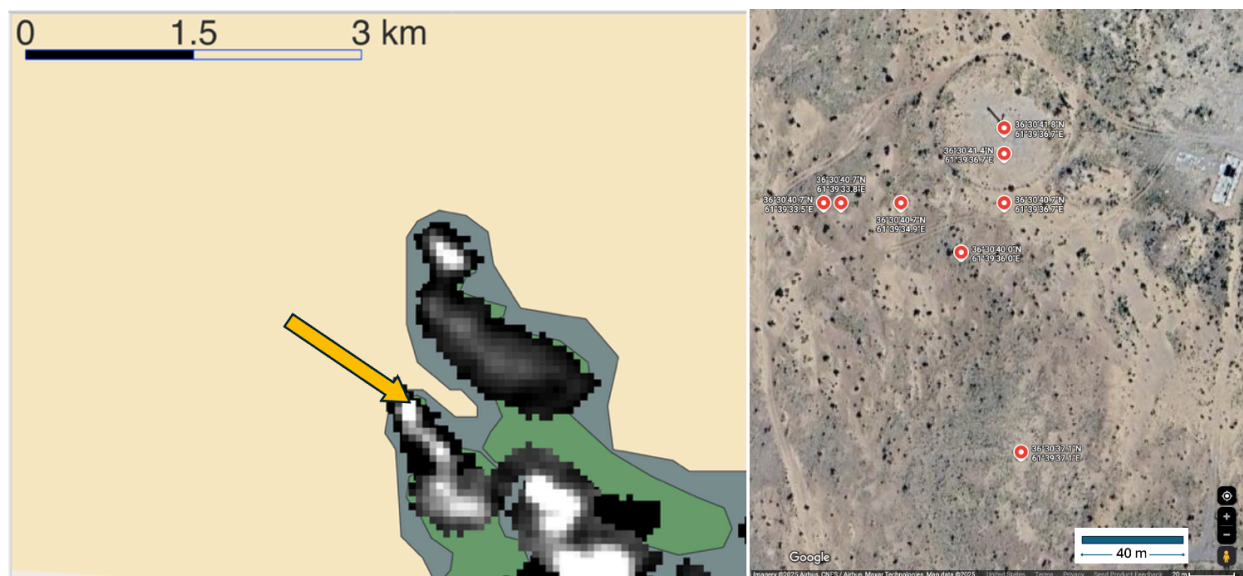
**Figure 4: Plume extent images for the 4 scenes. Note the different length scales on each panel, ranging from 3 km to 10 km. Colors are used to differentiate the plumes from different teams.**

## 4.2. Plume origin results

### 4.2.1. Origins in latitude and longitude

Plume origin refers to the latitude and longitude coordinates that have been assigned to the origin of the plume. As noted in NIST IR 8575, identifying the plume origin typically, but not always, involves the manual evaluation of various datasets, such as methane concentration fields, wind direction, and surface imagery. For the statistics discussed here, we will focus on

the results from the mix/match experiment. In general, when teams performed analysis on different L2 datasets, they did not determine different plume origins, so it is not critical to explore the statistics of results based on different L2. For this analysis, we examined the variation in latitude and longitude and quantified the variations in Universal Transverse Mercator (UTM) coordinates rotated to along-wind and cross-wind axes.



**Figure 5: Map with plume outlines in a part of scene 2 (left panel). The orange arrow points to the third plume in scene 2. The scale bar in the upper left is 3 km. Map of the set of plume origins identified for the third plume in scene 2 (right panel). A scale bar (showing 40 m) is included in the lower right of the map. The point of the yellow arrow in the left panel points to the location of the northernmost red marker on the right panel.**

Figure 5 highlights the scatter of the plume origins on a map for plume 3 in scene 2. We can see that the plume outlines extend on the order of 1 km to 2 km (left panel), whereas the scatter plot shows a scatter of the plume origin of less than 100 m for most points (right panel). While we can develop some measures of scatter of the latitude and longitude of the origin, it may be more insightful to assess the scatter in the context of the plume geometry. Specifically, assuming that the plume orientation is driven by the winds, we quantify variation of the plume origin in the along-wind and cross-wind directions.

#### 4.2.2. Rotation of plumes

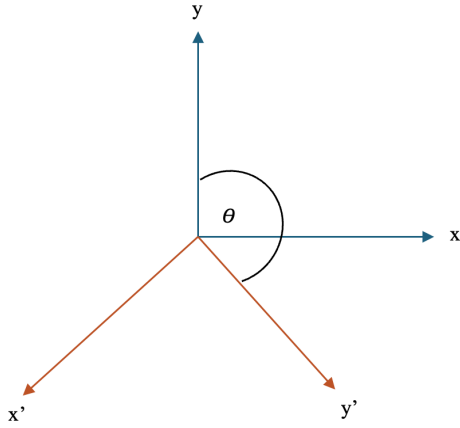
Two coordinate system transformations were applied to the plume images to develop the along-wind and cross-wind origin scatter information. The first step was to transform the data from latitude and longitude coordinates to the UTM system, which simplifies calculations. The UTM coordinates are UTM northing and UTM easting. This transformation was performed with an online conversion tool ([https://giscrack.com/download-excel-template-convert-geographic-coordinates-utm/#google\\_vignette](https://giscrack.com/download-excel-template-convert-geographic-coordinates-utm/#google_vignette)).

The next step was to determine the bearing of the plume,  $\theta$ , which is defined as the clockwise angle between the direction of plume spread (and the direction the wind is blowing into) and true north. This was determined manually using a Geographic Information System (GIS) map tool and the enhancement maps that were submitted. The rotation transformation from the UTM coordinate system, where  $y$  is the UTM northing coordinate and  $x$  is the UTM easting coordinate, to a wind coordinate system, where  $y'$  is the along-wind axis ( $y'$ ) and  $x'$  is the cross-wind axis ( $x'$ ), is given by Eq. 3. The bearing values in the cases we studied were in the range of  $131^\circ$  to  $221.5^\circ$ .

$$x' = x \cos \theta + y \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta.$$

(3)



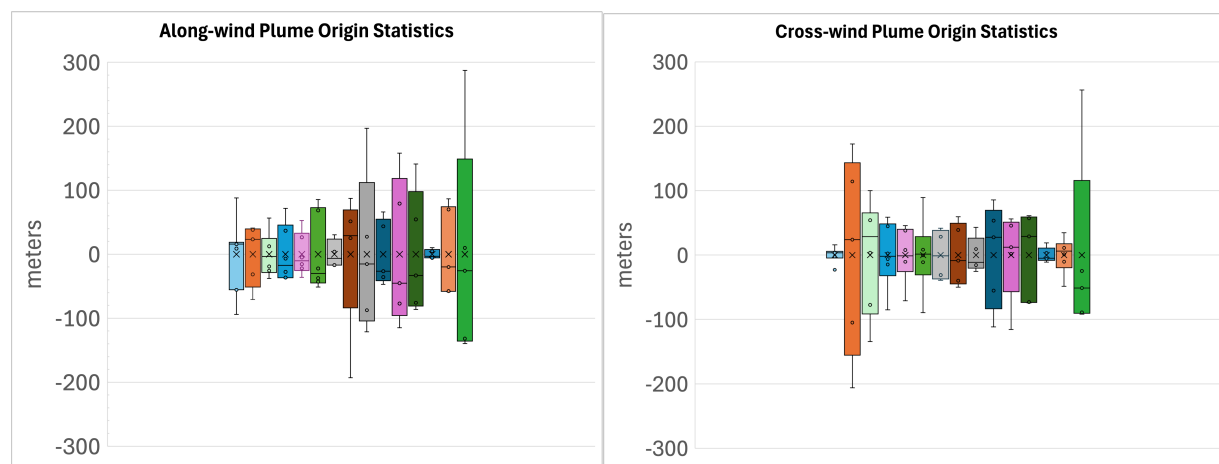
**Figure 6: The coordinate rotation of UTM northing ( $y$ ) and UTM easting ( $x$ ) to along-wind direction ( $y'$ ) and cross wind ( $x'$ ) given the wind blowing into bearing angle ( $\theta$ ).**

#### **4.2.3. Origin spread in along with and cross wind direction:**

Box and whisker plots were developed to quantify the distribution of the rotated plume origin data, Figure 7. The solid boxes represent the interquartile range, which is the middle 50 % of the data. The lower edge of the box is the first quartile (25th percentile), the upper edge is the third quartile (75th percentile); a line within the box marks the median (50th percentile) and an X marks the mean. The whiskers extend to the minimum and maximum values that are within 1.5 times the interquartile range, defined to be the furthest data points that are not considered outliers. Outliers were removed from the dataset before calculating the distribution statistics.

Our expectation was that the plume origin would be better characterized in the cross-wind direction where the plume is narrower, concentration gradients are large, and the gradients exist on both sides of the plume origin. This is not always the case, for example, with the plume of scene 4. We expect more variability in the along-wind direction, where the concentration gradient really only exists in one direction. What we find is that for eight of the plumes that were studied, variation in the origin identification in the cross-wind direction is indeed smaller than in the along-wind direction, as we expect. For four of the plumes, 2, 3, 7, and 10, the error in the cross-wind direction is larger than in the along-wind direction, and for three, the errors in across-wind and along-wind are comparable.

For the EMIT data, the pixel size is roughly 60 m. The whiskers in the cross-wind direction are 50 m to 110 m, except in two cases (scene 2, plume 1 and scene 4, plume 1). For the along-wind direction, the extent of the box and whisker plots is 1 pixel to 2 pixels (60 m to 120 m) for 12 of the 15 plumes. In future work, we will explore if errors for instruments with different pixel sizes are also 1 pixel to 2 pixels or if the error is still 50 m to 100 m, regardless of instrument characteristics. In addition, some teams assign plume origins only to pixels included in their plume mask, while others look at maps of infrastructure and use that information as part of the plume origin assignment. In future efforts, we will work to connect methodology choices to error characteristics.



**Figure 7: Box and whisker plots of the plume origin data for 15 plumes measured across all 4 scenes. The data has been rotated into the along-wind and cross-wind directions, and the statistics are for differences from the mean in meters. See Table 3 for legend to connect bars to scenes and plume numbers.**

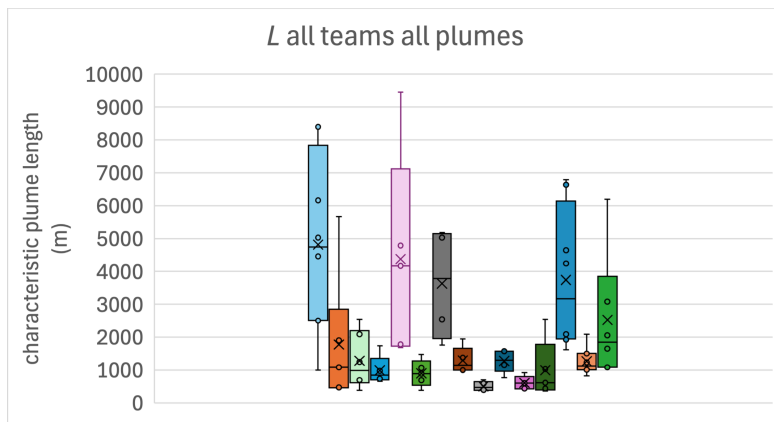
**Table 3: Plumes that are in the box and whisker plots (Figures 7,8, and 11)**

| Bar # | details         | Bar # | details          |
|-------|-----------------|-------|------------------|
| 1     | Scene 1 plume 1 | 9     | Scene 2 plume 8  |
| 2     | Scene 2 plume 1 | 10    | Scene 2 plume 10 |
| 3     | Scene 2 plume 2 | 11    | Scene 2 plume 11 |
| 4     | Scene 2 plume 3 | 12    | Scene 2 plume 12 |
| 5     | Scene 2 plume 4 | 13    | Scene 3 plume 1  |
| 6     | Scene 2 plume 5 | 14    | Scene 3 plume 2  |
| 7     | Scene 2 plume 6 | 15    | Scene 4 plume 1  |
| 8     | Scene 2 plume 7 |       |                  |

### 4.3. Plume segmentation (delineation) results

As touched on in the plume origin section, and illustrated in Figure 4, the results of plume delineation vary widely across teams. An assessment of the delineation approaches was not considered in this intercomparison exercise. As discussed in NIST IR 8575, there are many methods in use, such as clumping algorithms, machine learning, visual analysis or hand-drawn methods, and percentile thresholding. As a result, it is not surprising that the segmentation results vary widely. And as discussed above, it is not clear how the segmentation results impact the quantification, although very short or very elongated plumes in some circumstances have emission rate estimates far from the median.

The plume delineation does relate to the characteristic length found by each team. There are two ways that teams define the characteristic length. Some use the square root of the area of the plume, and others use a radial length from the origin. As we can see in Figure 8, there were seven cases where the value of  $L$  varies less than 1000 m across teams. These were cases where the median  $L$  ranged from 500 m to 1200 m. In contrast, there were 5 cases where the values spanned more than 5000 m across the teams. These were cases where the median  $L$  ranges from 2500 m to 5000 m. Overall, when the characteristics length is larger the range of values is also larger.



**Figure 8: Box and whisker plot of the characteristic plume length. The order of the bars is listed in Table 3.**

## 4.4. Emission rate estimates and key variables results

### 4.4.1. Initial comparison across all cases

Teams submitted estimates of the emission rate,  $Q$ , ( $\text{kg}\cdot\text{h}^{-1}$ ) and the associated uncertainty for each plume. The emission rate estimates from different teams are directly comparable. Figure 9 (a) contains box and whisker plots of the emission rate estimates from the mix and match experiment, as well as other key variables such as emission rate uncertainty, the wind speed and effective wind speed, and the characteristic length of the plume.

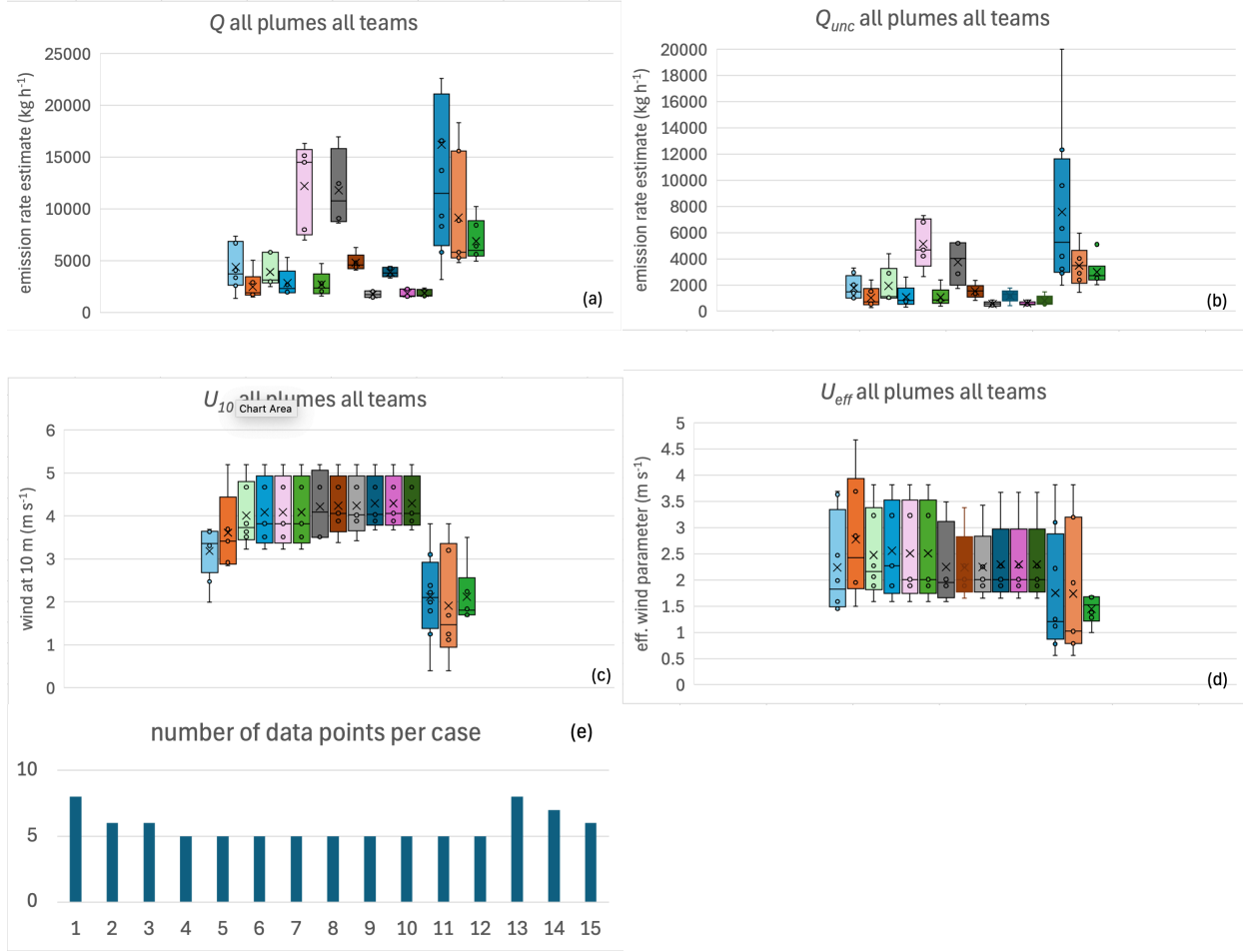
We can see that many of the plumes have emission rate estimates in the  $2000 \text{ kg}\cdot\text{h}^{-1}$  to  $5000 \text{ kg}\cdot\text{h}^{-1}$  range, and a few are much larger, in the  $10,000 \text{ kg}\cdot\text{h}^{-1}$  to  $15,000 \text{ kg}\cdot\text{h}^{-1}$  range. The absolute uncertainties tend to be smaller for the smaller emission rates and larger when the emission rates are larger.

The formulation of the uncertainty estimates varies across the teams. In general terms, most teams consider similar or the same sources of error, including error in the wind field, measurement error, effects of the spatial resolution of the measurement, error due to unaccounted for turbulence, and error due to plume segmentation. In most cases, the error terms are combined as a quadrature sum to derive the overall error in the emission rate. Very broadly, uncertainty values range from roughly 20 % to 50 %. All of the teams are reporting one standard deviation (1 sigma) in their uncertainty. The current experiment is not sufficient to evaluate the uncertainty estimates, but this can be a focus of later analysis that includes controlled release measurements.

Figure 9 (c) and (d) are the box and whisker plots for the 10 m wind ( $U_{10}$ ) and the effective wind parameter ( $U_{eff}$ ). The interquartile range for the 10 m winds for scene 2 (bars 2 through 12) are in the  $3.5 \text{ m s}^{-1}$  to  $5 \text{ m s}^{-1}$  range, and the values are nearly the same across all plumes, since they are all close together. For scene 1, wind speeds are lower ( $2.8 \text{ m s}^{-1}$  to  $3.8 \text{ m s}^{-1}$ ), and scenes 3 and 4 have the lowest winds for the scenes studied ( $1 \text{ m s}^{-1}$  to  $3 \text{ m s}^{-1}$ ). The scene 3 plume 2 winds speeds have the largest variability across teams of all the cases.

The effective wind parameter,  $U_{eff}$ , also discussed in more detail in section 4.4.3 (Formulations of  $U$ ) is more consistent across all of the cases than the 10 m wind. The interquartile values range from  $1.5 \text{ m s}^{-1}$  to  $4 \text{ m s}^{-1}$  for scenes 1 and 2, and from  $0.7 \text{ m s}^{-1}$  to  $3 \text{ m s}^{-1}$  for scenes 3 and 4. The cases almost all have medians that are smaller than the mean.

The number of data points that are included in each case are reported in panel (e). Not every team submitted results for every plume, so for some cases there are fewer than 8 data points.



**Figure 9: Box and whisker plots for the key quantities. Panel (a): Emission rate estimates,  $Q$ , Panel (b): Uncertainty on emission rate estimates (1 sigma), Panel (c): 10 m wind speed, Panel (d): effective wind parameter, Panel (e): number of data points included**

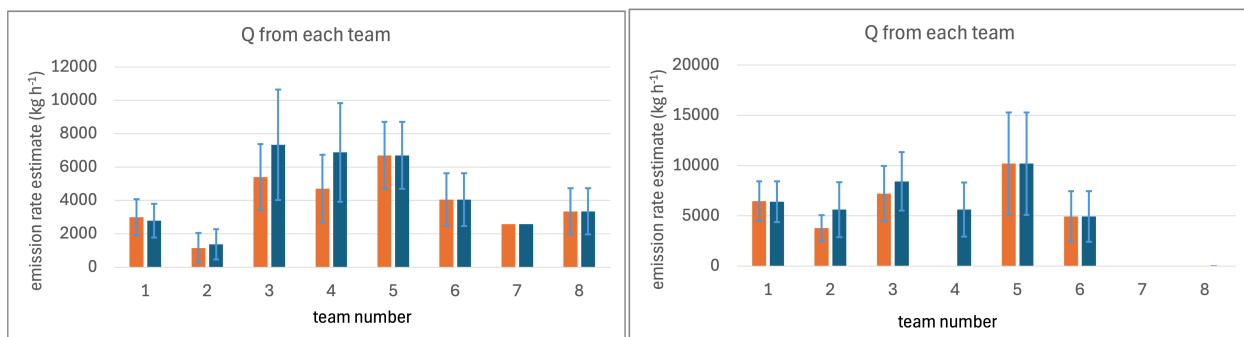
#### 4.4.2. Emissions rate estimates for two cases for two experiments

Figure 10 shows emission rate estimates from each team for the plumes of scenes 1 and 4. The orange bars are for results on the JPL L2, and the dark blue bars are for the mix and match experiment. Not all teams submitted results for all experiments and cases.

For scene 1, results for 6 of the 8 team's results agree to within their uncertainty bars for the JPL L2 case (orange bars), which range from 30 % to 80 % for plume 1. For the mix and match experiment (blue bars), the disagreement increases slightly. The change of L2 field resulted in an increase of emission rate estimates for 3 of the 4 teams. For scene 4 (right panel), there were 6 teams that submitted results. For this scene all teams' estimates agree within error bars for both experiments. Changing the L2 data source for scene 4 resulted in an increase in the emission rate estimate for two teams, while the rest changed little and emission rate estimates agree to within error bars. This sample size and experimental design are too small to make any



generalizations but are helpful in developing our ideas about how to formulate better intercomparison experiments.



**Figure 10: The emission rate estimates (kg·h<sup>-1</sup>) with uncertainty bars (1 sigma) for the single plumes of scene 1 (left panel) and scene 4 (right panel). The orange bars are for results on the JPL L2, and the dark blue bars are for the results of the mix and match experiment**

#### 4.4.3. Formulations of U

The estimated emission rate is directly proportional to the wind speed, so it is a key variable to assess. To assess how different terms contribute to the variability in emission rate estimates, we performed a test where the wind speed was held constant across all teams. As discussed in Section 3.3 (Integrated mass enhancement approach for emission rate estimates:), the relationship between  $U_{10}$  and  $U_{eff}$  varies across teams. We can understand the impact both by examining the relationships directly, and by calculating emission rates if  $U_{10}$  were constant across all teams.

Figure 11 shows five different formulations of the  $U_{eff}=f(U_{10})$  relationship used by the 8 participating teams. Teams 2, 7, and 8 use a direct substitution,  $U_{eff}=U_{10}$ , which reflects their definition of  $L$  as a radial plume length (distance from plume origin) and an assumption that the 10 m wind is representative of transport conditions over that distance. The remaining teams use an effective wind of the form  $U_{eff}=aU_{10}+b$  to match their definition of  $L$  as the square root of the plume area. The parameters  $a$  and  $b$  are derived using synthetic plume data from large eddy simulations (LES) driven with prescribed emission rates and  $U_{10}$  following Varon et al. (2018). The derivations are based on different LES datasets and make different assumptions about the methane retrieval uncertainty, pixel resolution, and wind conditions, which would all influence the observable plume extent—and therefore  $IME$  and  $L$ —for a given source rate  $Q$ . The parameter ranges are  $a \in [0.25, 0.49]$  and  $b \in [0.4, 0.62]$ .

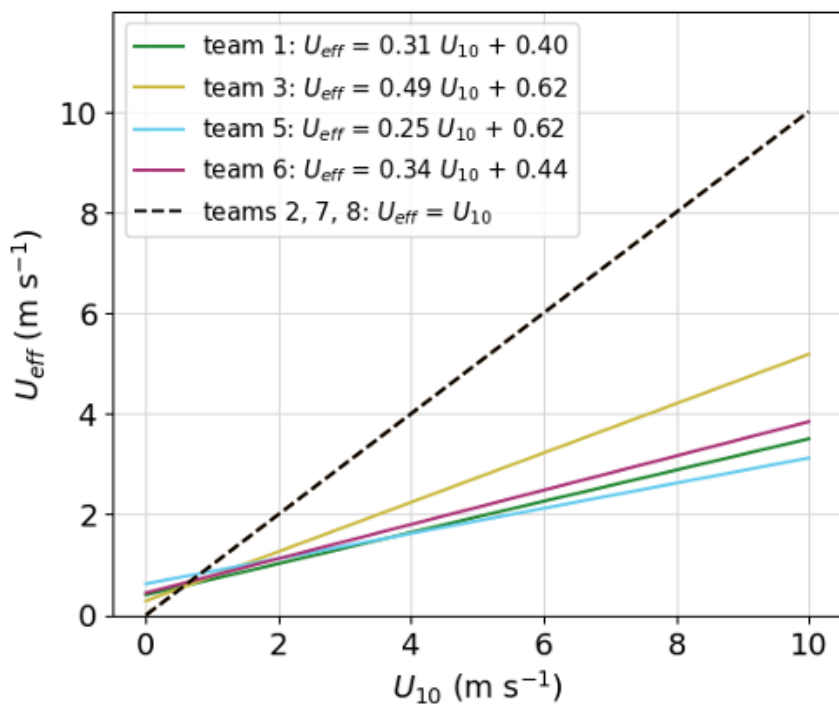
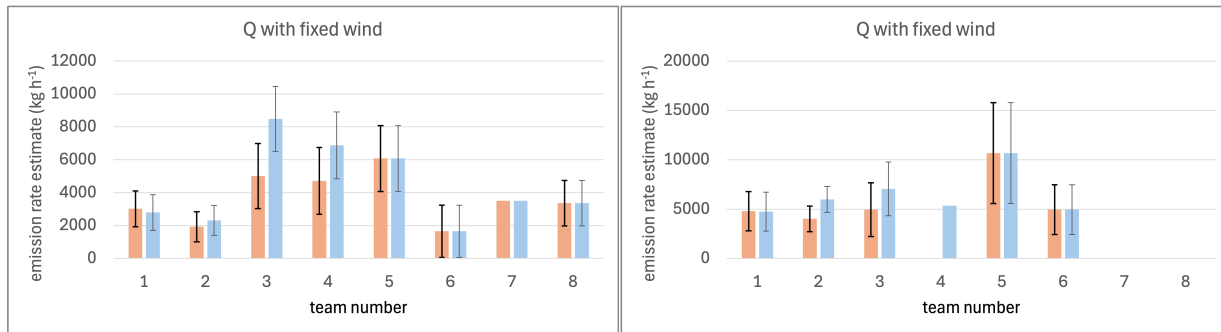


Figure 11: Different formulations of the IME-method effective wind speed  $U_{eff}$  as a function of 10-m wind speed  $U_{10}$  used by the participating teams. Solid lines define the plume length scale  $L$  for the IME method as the square root of the plume mask area. Dashed (

#### 4.4.4. Comparison when wind speed is held constant

In Figure 12, we show the estimated emission rates for two experiments where the wind speeds are held constant across all teams. These are shown for the results on the JPL L2 (orange bars) and the mix and match experiment (blue bars). The results for scene 1 (left panel) and scene 4 (right panel) are displayed. For scene 1, 5 of the 7 teams agree to within their 1 sigma error bars for the JPL L2 and fixed wind speed case. The teams that disagree have lower emission estimates. There is less agreement for the mix and match case, with 4 of the 7 teams agreeing to within their 1 sigma error bars, but note that this includes all the teams with lower estimates, in contrast to the teams that were in agreement for the JPL L2 case. For scene 4, the 4 of the 5 teams that provided estimates with error bars agree to within error bars for the JPL L2 case, with one team slightly out of agreement. For the mix and match case with fixed wind speeds, all of the 5 teams agree within the error bars. Again, with such a small set of cases, we cannot make firm conclusions about the contribution of each term to the variability across the teams.

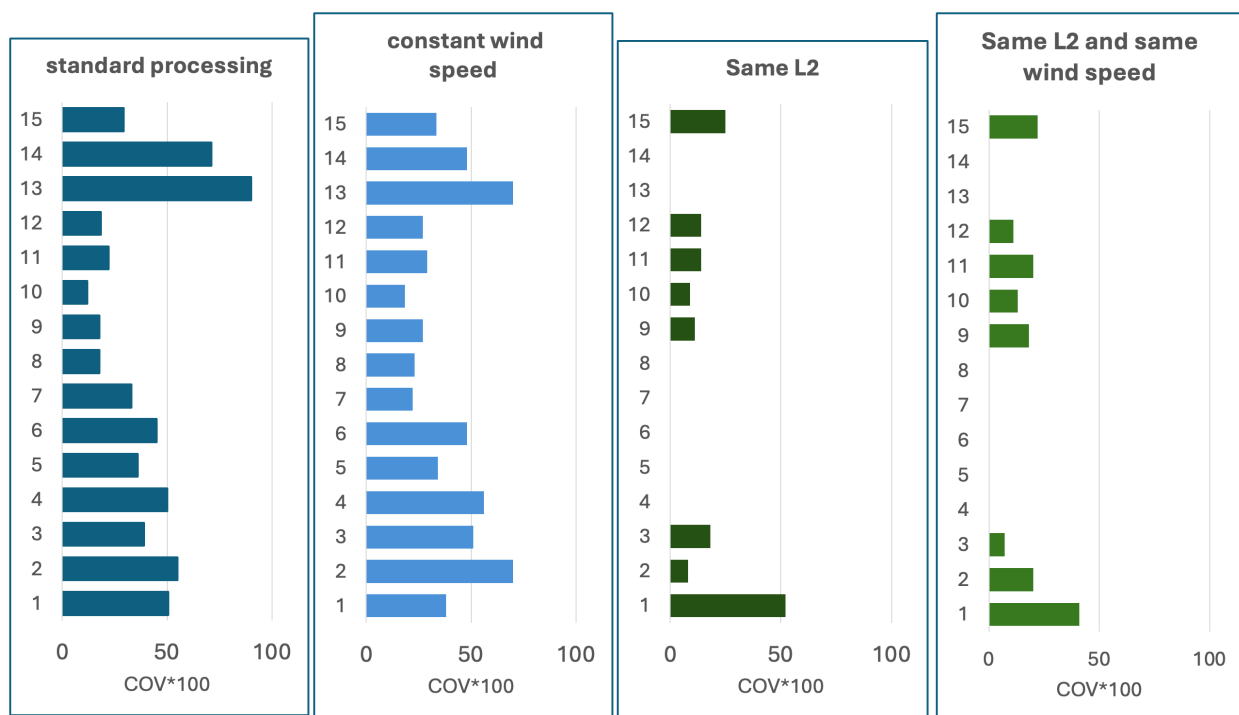


**Figure 12: The emission rate estimates (kg·h<sup>-1</sup>) with uncertainty bars for the plume of scene 1 (left panel) and scene 4 (right panel) when wind speeds are constant across all teams. The orange bars are for results on the JPL L2, and the blue bars are for the resu**

#### 4.5. Coefficient of Variation for emission estimates

To formalize analysis of the variation in emission rate estimates across groups, we use the COV. The COV is a standardized measure of dispersion and is calculated by dividing the standard deviation by the mean of the population of data. When the COV is higher, we can conclude that there is greater dispersion in the emission rate estimates relative to their average value. In our work we express it as a percentage. Note that in the statistics presented in Figure 12, there is a larger population (8 teams) for the standard processing and constant wind speed, whereas there are only 4 teams that contributed to the team developed L2 group of data.

Figure 13 shows the COV across 15 of the 16 plumes that were identified across the 4 scenes. There is one from scene 1, eleven in scene 2, two in scene 3, and one in scene 4. Table 3 has the details of the plumes included and the corresponding number on the charts.



**Figure 13: Coefficient of Variation (COV) for emissions estimate analysis experiments. Table 3 provides the details of the scenes and plumes for each bar.**

A variety of data combinations are included in Figure 13. The left panel, labeled standard processing, is the mix and match experiment, which includes analysis on team-developed L2 for 4 teams and JPL L2 for 4 teams. This is the type of data end users would expect to access, so it is useful to set expectations. We see that across all of the plume cases, the COV ranges from 20 % to 90 %, with the majority of the COV values in the 30 % to 50 % range. The COV data can be compared to the box and whisker plots of Figure 8, where we see that plumes 13 and 14 are clearly the cases with the widest distribution of data, indicated by the whiskers. Those cases are also among the larger emission rates and lower wind speeds, although that does not fully explain the COV of each case.

The second panel of Figure 13 shows COV for the case where the wind speed has been held constant for the mix and match L2 case. There is a mixture of impacts, where the COV increases for some cases, and COV decreases for other cases. The range of COV values is somewhat smaller with fixed winds, but fixing the wind does not result in a drastic reduction of COV overall.

The third and fourth panels specifically compare analyses performed on JPL L2 data. Not all plumes are included in this analysis. We see, by comparing the COV to the far left panel, that the COV in the third and fourth panels is similar to but slightly smaller relative to the first panel. So again, the variation in the L2 data does not explain all the variation across teams. The far right panel is an experiment where the L2 data and the wind speed are both held constant for a small set of plumes. Overall, COV are lower, but the cases included here are the lower COV

cases. For plumes 1, 2, 3, and 15 the COV is reduced with fixed winds and L2. With such a small set of cases, we cannot make generalized conclusions, but we can see that similar experiments where the emission rate is known would be useful.

## 5. Conclusions

An intercomparison exercise has been conducted with a collection of eight teams who detect and quantify methane emission from remote sensing data. This group of teams has been involved in a collaborative effort to document current practices, perform analysis across teams, and work towards developing documentary standards.

With this project, we have developed methods for comparing plume detections, plume origins and reframing the comparisons to a set of coordinates along the wind direction and across the wind direction. We also compared emission rate estimates as submitted and with variations, such as fixed wind speed across all teams and uniform L2 fields in the analysis. The report also includes statistics on the wind data used by the teams and the characteristic plume lengths,  $L$ . The plume origin analysis showed that the cross wind spread of plume origin results was 50 m to 110 m in all but one case, and across wind was 60 m to 120 m in 13 of the 15 cases. The EMIT footprint is 60 m, so the variations were roughly 1 to 2 pixels. We saw that the COV of the emission rate estimates was typically 30% to 50%, and fixing the wind speed and/or the L2 field did not drastically change this, except in 2 cases. The higher COV cases were often the cases with lower wind speeds, and likely unsteady winds, which can be challenging for the plume segmentation step of the analysis.

Plumes analyzed were not controlled releases, but methane emissions that appeared in the scenes. Without controlled releases, we cannot assess the bias of the measurements. Additionally, generalized conclusions from this experiment are also pre-empted by the small sample size. In addition, the plumes studied here had emission rates ranging from almost 1000  $\text{kg}\cdot\text{h}^{-1}$  to 15000  $\text{kg}\cdot\text{h}^{-1}$  with plume lengths from a few kilometers to 20 km or 30 km. The applicability of what we have learned is limited to plumes with similar scales. There are other aspects of the data analysis process that were not investigated, such as plume segmentation and its influence on  $L$  and the IME. We did not perform analysis of how the L2 fields themselves compared. We also saw that there was a case where the number of plumes identified in the scene varied across teams, and the reasons for this were not investigated in this effort. The experiment provided a snapshot of variances in methane plume estimates associated with different algorithms in use in spring of 2024, (and for many teams their methods have evolved since that time.)

## 6. Community Efforts and Next Steps

While it is interesting to develop a set of statistics from these intercomparison analyses, there are limits to what we can learn without knowledge of the true emission rate. There have been a number of large, controlled methane release experiments and we anticipate another intercomparison exercise that includes controlled release data from 2024 and 2025 so we can assess potential biases of the emission rate estimates between teams, and whether or not the true emission rates are within the estimated uncertainties. Controlled-release experiments also provide known values of the plume origins, so we can assess variability and accuracy of participants identified plume origins. However, there are limitations to controlled emission analyses. For example, the emission rates of controlled releases are generally less than 1500 kg·h<sup>-1</sup>, whereas the wild plumes we have worked with to date have larger estimated emission rates. There are also a range of surface reflectances and topographies not covered by controlled releases. In future work, we will seek to assess both controlled releases and wild plumes and understand what conclusions can be drawn that apply to each and both datasets.

Our effort complements other work in the community beyond the CEOS report efforts. There are new papers appearing in the scientific literature frequently on improvement of methods for plume detection and quantification, and intercomparison of results with different sensors. There has been a series of controlled release experiments led by Stanford University (now in collaboration with the University of Michigan) that have reported the statistics of remotely sensed emission rate estimates for known blind (teams do not know rates when analyzing the experiment) release rates for satellites (Sherwin et al., 2024a) and aircraft (El Abbadi et al., 2024). Additional measurements have been underway in 2024 and 2025 and will add to our knowledge of how a wide range of satellite measurements compare to the known controlled release rates. These studies focus on the aggregate statistics like the quantitative accuracy and the rate of true positives and false positives. They also have many teams analyzing multiple different instrument measurements.

A related activity is underway coordinated by Colorado State University at the Methane Emissions Technology Evaluation Center (METEC). The METEC performs controlled releases, but the initial focus was on low emission rates (less than 7 kg·h<sup>-1</sup>) and evaluation of ground based measurement systems. Recently, the METEC facility added a capability to do controlled releases in the range of 25 kg·h<sup>-1</sup> to 2000 kg·h<sup>-1</sup>, aimed at point source imager satellites and fixed wing aircraft measurements.

A second related activity is the MEDUSA project (Methane Emissions Detection Using Satellites Assessment) under the European Space Agency Climate Office. MEDUSA is a large, collaborative 3-year project that began in 2024 and is focused on the complex, heterogeneous system of satellites that are remotely sensing methane plumes. It brings together seven different research teams from Europe and Canada. The MEDUSA project will compare the various plume (emission) satellite products, both from the same instrument using different algorithms and/or produced by different research teams, as well as between different satellites. Where possible, they will validate the emission products through controlled releases. They are also developing a



framework for methane flux uncertainty calculations for the various types of instruments and performing a few case studies highlighting the combined use of these point source/hot spot methane observations. Some but not all of the teams that participated in this intercomparison are involved in the MEDUSA work, and there are teams working within MEDUSA that were not part of this effort.

In contrast to the other work, our effort is looking at details of the analysis process and we are focusing on the analysis by many teams of just one instrument measurement dataset. In addition, in future work we will analyze both wild plumes (plumes observed in the scene) as well as controlled release data, to begin connecting the well characterized experimental data to the data that is of broader use. Another aspect that we will consider is how the range of emission rates that can be achieved in controlled release experiments can be informative for the wider range of emission rates seen in the broader measurement sets.

## References

- [1] Carbon Mapper ATBD version 1.0.0  
[https://assets.carbonmapper.org/documents/L3\\_L4%20Algorithm%20Theoretical%20Basis%20Document\\_formatted\\_10-24-24.pdf](https://assets.carbonmapper.org/documents/L3_L4%20Algorithm%20Theoretical%20Basis%20Document_formatted_10-24-24.pdf)
- [2] Carvalho, D. (2019). An assessment of NASA's GMAO MERRA-2 reanalysis surface winds. *Journal of Climate*, 32(23), 8261-8281.
- [3] El Abbadi, S. H., et al. (2024), Technological Maturity of Aircraft-Based Methane Sensing for Greenhouse Gas Mitigation, *Environ. Sci. Technol.*, 58, 9591–9600.
- [4] Foote, M. D., Dennison, P. E., Thorpe, A. K., Thompson, D. R., Jongaramrungruang, S., Frankenberg, C., & 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), 6480-6492.
- [5] Frankenberg, C., Thorpe, A. K., Thompson, D. R., et al. (2016), Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region, *Proc. Natl. Acad. Sci.*, 113(35), 9734–9739, doi:10.1073/pnas.1605617113.
- [6] Guanter, L., Irakulis-Loitxate, I., Gorroño, J., Sánchez-García, E., Cusworth, D. H., Varon, D. J., ... & Colombo, R. (2021). Mapping methane point emissions with the PRISMA spaceborne imaging spectrometer. *Remote Sensing of Environment*, 265, 112671.
- [7] Guanter, L., Roger, J., Sharma, S., Valverde, A., Irakulis-Loitxate, I., Gorroño, J., ... & Zavala-Araiza, D. (2024). Multisatellite Data Depicts a Record-Breaking Methane Leak from a Well Blowout. *Environmental Science & Technology Letters*, 11(8), 825-830.
- [8] Kim, K. E., Lee, S. S., & Baik, H. S. (2016, October). Iterative matched filtering for detection of non-rare target materials in hyperspectral imagery. In *Image and Signal Processing for Remote Sensing XXII* (Vol. 10004, pp. 133-139). SPIE.
- [9] Pei, Z., Han, G., Mao, H., Chen, C., Shi, T., Yang, K., ... & Gong, W. (2023). Improving quantification of methane point source emissions from imaging spectroscopy. *Remote Sensing of Environment*, 295, 113652.
- [10] Schuit, B. J., Maasackers, J. D., Bijl, P., Mahapatra, G., Van den Berg, A. W., Pandey, S., ... & Aben, I. (2023). Automated detection and monitoring of methane super-emitters using satellite data. *Atmospheric Chemistry and Physics Discussions*, 2023, 1-47.
- [11] Sherwin, E. D., et al. (2024a), Single-blind test of nine methane-sensing satellite systems from three continents, *Atmos. Meas. Tech.*, 17, 765–782.
- [12] Sherwin, E. D., Rutherford, J. S., Zhang, Z., Chen, Y., Wetherley, E. B., Yakovlev, P. V., et al. (2024b). US oil and gas system emissions from nearly one million aerial site measurements. *Nature*, 627(8003), 328–334. <https://doi.org/10.1038/s41586-024-07117-5>
- [13] Varon, D. J., Jacob, D. J., McKeever, J., Jervis, D., Durak, B. O., Xia, Y., & Huang, Y. (2018). Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmospheric Measurement Techniques*, 11(10), 5673-5686.
- [14] Williams, J. P., Omara, M., Himmelberger, A., Zavala-Araiza, D., MacKay, K., Benmergui, J., ... & Gautam, R. (2025). Small emission sources in aggregate disproportionately account for a large majority of total methane emissions from the US oil and gas sector. *Atmospheric Chemistry and Physics*, 25(3), 1513-1532.

- [15] Worden, J., Green, P., Eldering, A., and Sherwin, E. (2025) Common Practices for Quantifying Methane Emissions from Plumes Detected by Remote Sensing. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency Report (IR) Publication ID. <https://doi.org/10.6028/NIST.IR.8575>
- [16] Zhang, X., Maasackers, J. D., Roger, J., Guanter, L., Sharma, S., Lama, S., ... & Aben, I. (2024). Global identification of solid waste methane super emitters using hyperspectral satellites. <https://eartharxiv.org/repository/view/7984/>

## **Appendix A. Acronyms & Abbreviations**

|         |                                                       |
|---------|-------------------------------------------------------|
| ATBD    | Algorithm Theoretical Basis Document                  |
| CEOS    | Committee on Earth Observation Satellites             |
| COV     | Coefficient of Variance                               |
| EMIT    | Earth Surface Mineral Dust Source Investigation       |
| ESA     | European Space Agency                                 |
| GIS     | Geographic Information System                         |
| IME     | Integrated Mass Enhancement                           |
| IMEO    | International Methane Emission Observatory (of UNEP)  |
| JPL     | Jet Propulsion Laboratory                             |
| L0      | Level 0                                               |
| L1      | Level 1                                               |
| L2      | Level 2                                               |
| LBNL    | Lawrence Berkeley National Laboratory                 |
| NASA    | National Aeronautics and Space Administration         |
| NIST    | National Institute of Standards and Technology        |
| NPL     | National Physical Laboratory, UK                      |
| SI      | Système International (International System of Units) |
| SRON    | Space Research Organization Netherlands               |
| TROPOMI | Tropospheric Monitoring Instrument                    |
| UPV     | Valencia Polytechnic University                       |
| UTM     | Universal Transverse Mercator                         |