# Machine Learning Models for Multi-sensor Detection of Methane Leaks in Hyperspectral Data

Vít Růžička[1,2], Gonzalo Mateo-García (Presenter)[1], Itziar Irakulis-Loitxate[1,3], Manuel Montesino San Martin[1], J. Emmanuel Johnson[1], Anna Allen[1], Carol Castaneda[1], Alma Raunak[1], Claudio Cifarelli[1], Luis Guanter[3]

1. International Methane Emissions Observatory, United Nations Environment Programme
2. University of Oxford, Oxford, United Kingdom
3. Universitat Politècnica de València (UPV)

UN environment programme

Introduction
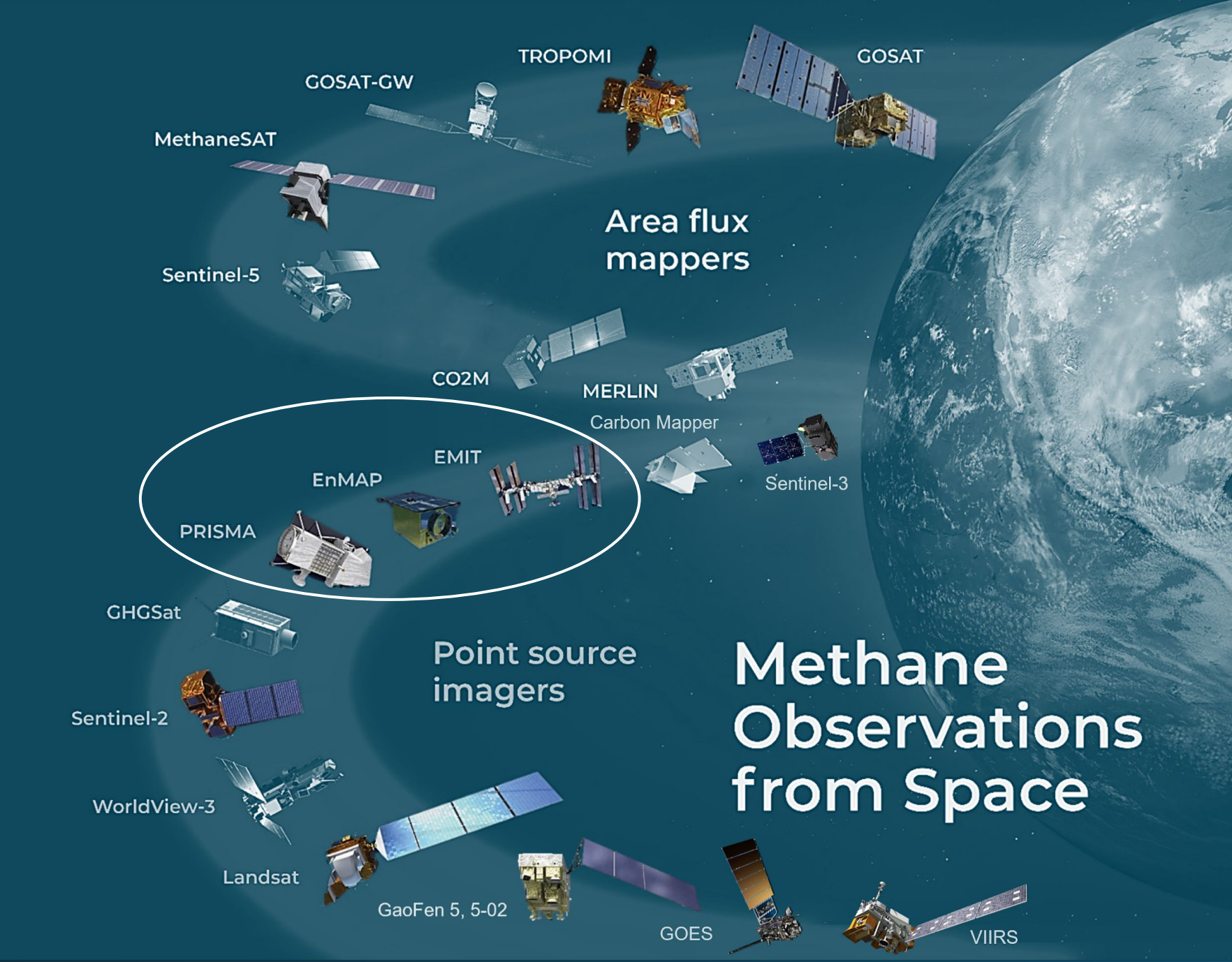
The **Methane Alert and Response System** (MARS) is a global satellite based initiative to detect and notify oil and gas methane emissions to governments and companies to swift mitigation action. Among the different satellites that MARS utilises, hyperspectral sensors such as **EMIT**, **PRISMA** and **EnMAP** offer an optimal trade-off between methane absorption sensitivity and spatial resolution that enables the **attribution of methane plumes to individual facilities**.

Scanning large archives of satellite imagery to find methane plumes is a challenging task because matched filter retrievals still produce a **significant amount of artifacts**. Since April 2025 MARS is aided by AI models to detect potential plumes in hyperspectral images of EMIT, PRISMA and EnMAP. These detections are displayed within a QA/QC tool where analysts log in and verify or reject the detections. If the detection is recent and can be attributed to a facility in the ground a formal notification is issued to government and operators and MARS case-managers engage with the parties to trigger mitigation action.

Datasets

The MARS team curates a large database of observed methane plumes across different satellites in the oil and gas, waste and coal sectors. (methanedata.unep.org). For this work we compiled the following datasets to train AI models:

MARS-EMIT

Extracted 256x256px labeled tiles and extended full granule test dataset.

RGB | WMF | Mask

+ Auxiliary information: wind, location embeddings.

Temporal split leaving 2024 data for validation and testing (y-axis shows the plume quantifications).

Split	Images (tiles 256x256, positive : bg tiles)	Sites	Original granules
Train	5746 (2796 : 2950)	1590	2250
Val	2316 (957 : 1359)	1244	954
Test	1709 (824 : 885)	1103	713
Full tile test	200 (100 : 100)	-	200

MARS-PRISMA

Between Oct. 2020 to May 2025

Split	Images (tiles 256x256, positive : bg tiles)	Sites	Full Tiles
Train	904 (381 : 523)	467	342
Test	226 (115 : 111)	192	115
Full tile test	115 (63 : 52)	-	115

MARS-EnMAP

Between Oct. 2023 to May 2025

Split	Images (tiles 256x256, positive : bg tiles)	Sites	Full Tiles
Train	391 (172 : 219)	293	156
Test	168 (81 : 87)	151	83
Full tile test	83 (53 : 30)	-	83

Methods & Results

We train U-Net models on the EMIT dataset and **fine-tune these models** on PRISMA and EnMAP.

EMIT

Model variant	Tiled dataset, segmentation: AUPRC	F1	Precision	Recall	Fulltile sources, individual events: Detected	Missed	False Alarms
baseline Mag1c (thr. 500, cross)	26.83	22.45	36.22	16.27	157	103	41991
baseline MF (thr. 500, cross)	31	27.7	39.61	21.29	176	84	110618
baseline WMF (thr. 500, cross)	28.35	23.4	38.8	16.75	187	73	79460
U-NET RGB+Mag1c	40.87 ±2.75	43.59 ±3.17	53.49 ±1.90	37.00 ±4.36	123 ±17	137 ±17	5801 ±2159
U-NET RGB+MF	56.06 ±2.49	53.92 ±2.14	63.04 ±3.97	47.51 ±4.31	174 ±16	86 ±16	6703 ±2145
U-NET RGB+WMF	68.26 ±3.06	63.07 ±2.46	72.99 ±1.17	55.64 ±3.81	182 ±18	78 ±18	3597 ±1833
U-NET RGB+WMF, Wind	68.97 ±3.16	63.68 ±2.77	69.28 ±2.75	59.01 ±3.69	190 ±17	70 ±17	5636 ±2610
U-NET RGB+WMF, Wind, Location	69.43 ±2.61	64.40 ±2.08	71.88 ±2.35	58.65 ±4.66	161 ±20	99 ±20	2425 ±890
Ensemble, 5x U-NET RGB+WMF	73.58	65.13	79.87	54.98	179	81	1527
Ensemble, 5x U-NET RGB+WMF, Wind	72.64	65.49	73.81	58.85	192	68	2120

+45% | -97%

PRISMA

Model variant	Tiled dataset, segmentation: AUPRC	F1	Precision	Recall	Fulltile sources, individual events: Detected	Missed	False Alarms
Baseline (WMF 500 thr, "ones" kernel)	N/A	26.05	23.14	29.8	120	21	57278
Train from scratch, 50 ep.	23.78 ±3.16	30.85 ±1.53	30.95 ±5.06	32.13 ±4.16	80 ±16	61 ±16	4178 ±2657
Zero-shot (EMIT U-NET RGB+WMF model)	31.59 ±1.99	40.63 ±2.95	31.08 ±2.82	58.93 ±3.88	111 ±12	30 ±12	5234 ±2595
Fine-tune (PRISMA), 9 ep.	34.19 ±3.79	43.50 ±2.18	39.69 ±5.30	49.08 ±2.83	97 ±4	44 ±4	2830 ±893
Fine-tune (EMIT+PRISMA), 1 ep.	32.65 ±2.04	41.76 ±1.98	32.53 ±3.07	59.16 ±4.40	110 ±8	31 ±8	4286 ±1510
Ensembled Zero-shot	35.6	44.04	37.02	54.56	113	28	2881
Ensembled Fine-tune (PRISMA), 9 ep.	35.18	44.29	37.53	54.02	100	41	2088
Ensembled Fine-tune (EMIT+PRISMA), 1 ep.	35.74	44.38	36.17	57.41	105	36	2690

+70% | +80%

EnMAP

Model variant	Tiled dataset, segmentation: AUPRC	F1	Precision	Recall	Fulltile sources, individual events: Detected	Missed	False Alarms
Baseline (WMF 500 thr, "ones" kernel)	N/A	35.37	37.82	33.21	113	21	20272
Train from scratch, 50 ep.	43.96 ±5.36	44.27 ±4.38	63.79 ±9.12	34.22 ±3.69	58 ±4	76 ±4	1349 ±584
Zero-shot (EMIT U-NET RGB+WMF model)	63.45 ±1.73	57.45 ±2.08	50.91 ±3.27	66.24 ±3.19	100 ±7	34 ±7	2105 ±1379
Fine-tune (EnMAP), 8 ep.	63.61 ±4.15	58.69 ±3.63	69.44 ±2.62	51.23 ±5.98	88 ±6	46 ±6	955 ±249
Fine-tune (EMIT+EnMAP), 2 ep.	64.65 ±1.62	59.50 ±2.04	54.46 ±3.71	65.85 ±2.41	101 ±5	33 ±5	1855 ±876
Ensembled Zero-shot	68.71	61.65	58.5	65.15	101	33	1303
Ensembled Fine-tune (EMIT+EnMAP), 2 ep.	69.53	63.66	61.68	65.77	101	33	1058

+80%

Test Predictions

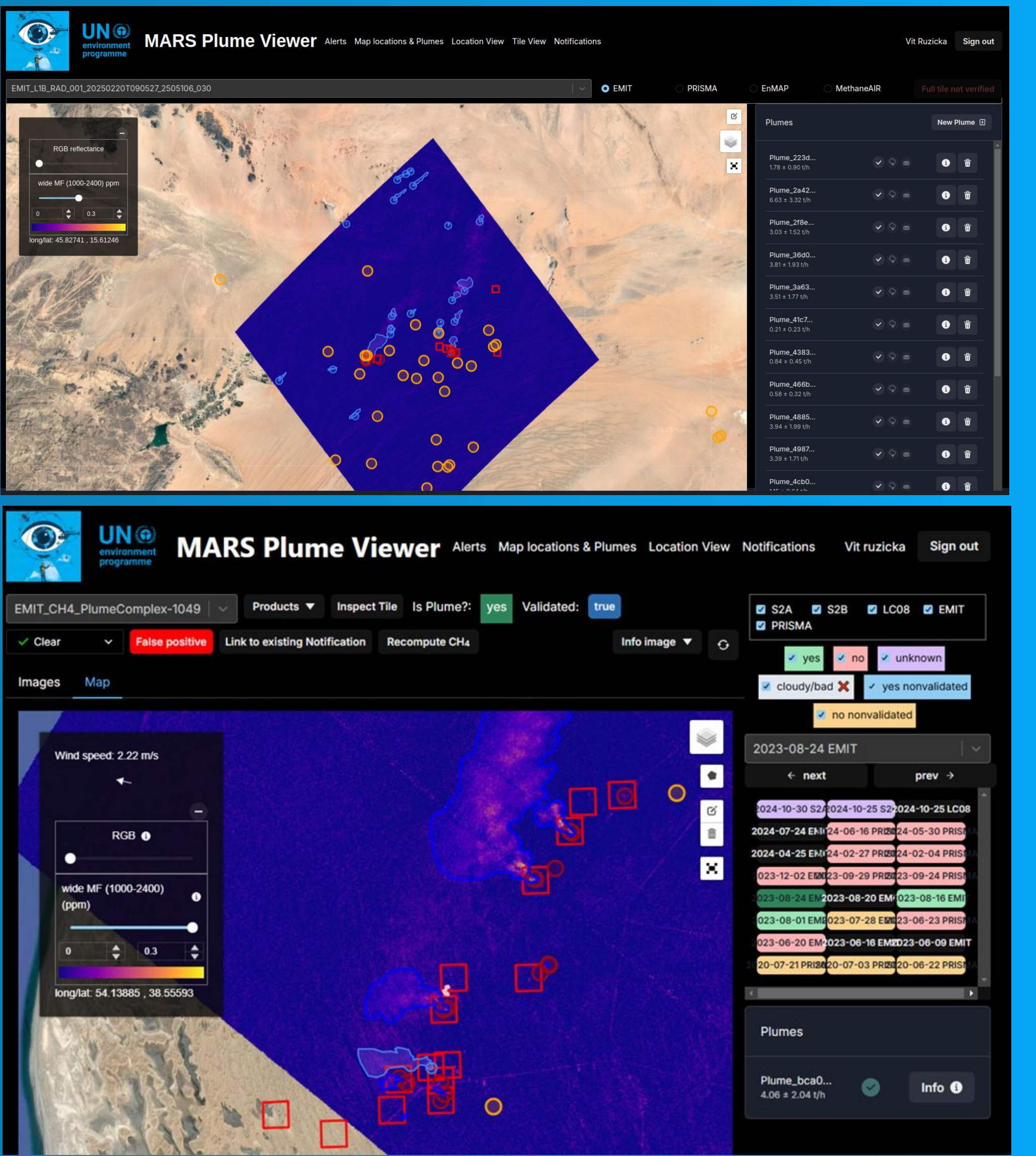RGB | WMF | Baseline | UNet | Ensemble

TN | FN | FP | TP

Operational predictions (full 76x74km)

Operational Deployment (* Early results since May 2025)

	EMIT	PRISMA	EnMAP
Processed tiles	2686	222	556
Verified Plumes detected (all sectors)	99	11	13
Different countries	19	4	4
Plumes notified	45	0	5

Key Takeaways

- **Generalisation across three** operational **sensors**
- **Deployment** in a real-world pipeline, **used daily**
- **Public release** of the **largest world wide available dataset of methane leak events**
- **Full granule predictions** (full tile tests)
- **Significant reduction of false alerts** with **model ensembles**

Růžička, V., Mateo-García, G., Irakulis-Loitxate, I., Montesino San Martin, M., Johnson, J.E., Allen, A., Cifarelli, C., Guanter, L.: Machine Learning Models for Multi-sensor Detection of Methane Leaks in Hyperspectral Data, Living Planet Symposium 2025, 23-27 June 2025, Vienna, Austria

June 25, 2025 - A.04.03 - POSTER