FROM SYNTHETIC TO REAL: A SCALABLE DATASET AND ML FRAMEWORK FOR GLOBAL METHANE MONITORING

16TH APRIL 2026 // 5TH ECMWF-ESA MACHINE LEARNING WORKSHOP

Enno Tiemann - OHB Digital Connect, Technical University Munich
Alexander Kläser – OHB Digital Connect
Adam Stewart – Technical University Munich
Philipp Hochstaffl – German Aerospace Center
Rochelle Schneider – Φ-lab, European Space Agency
Xiao Xiang Zhu – Technical University Munich / Munich Center For Machine Learning

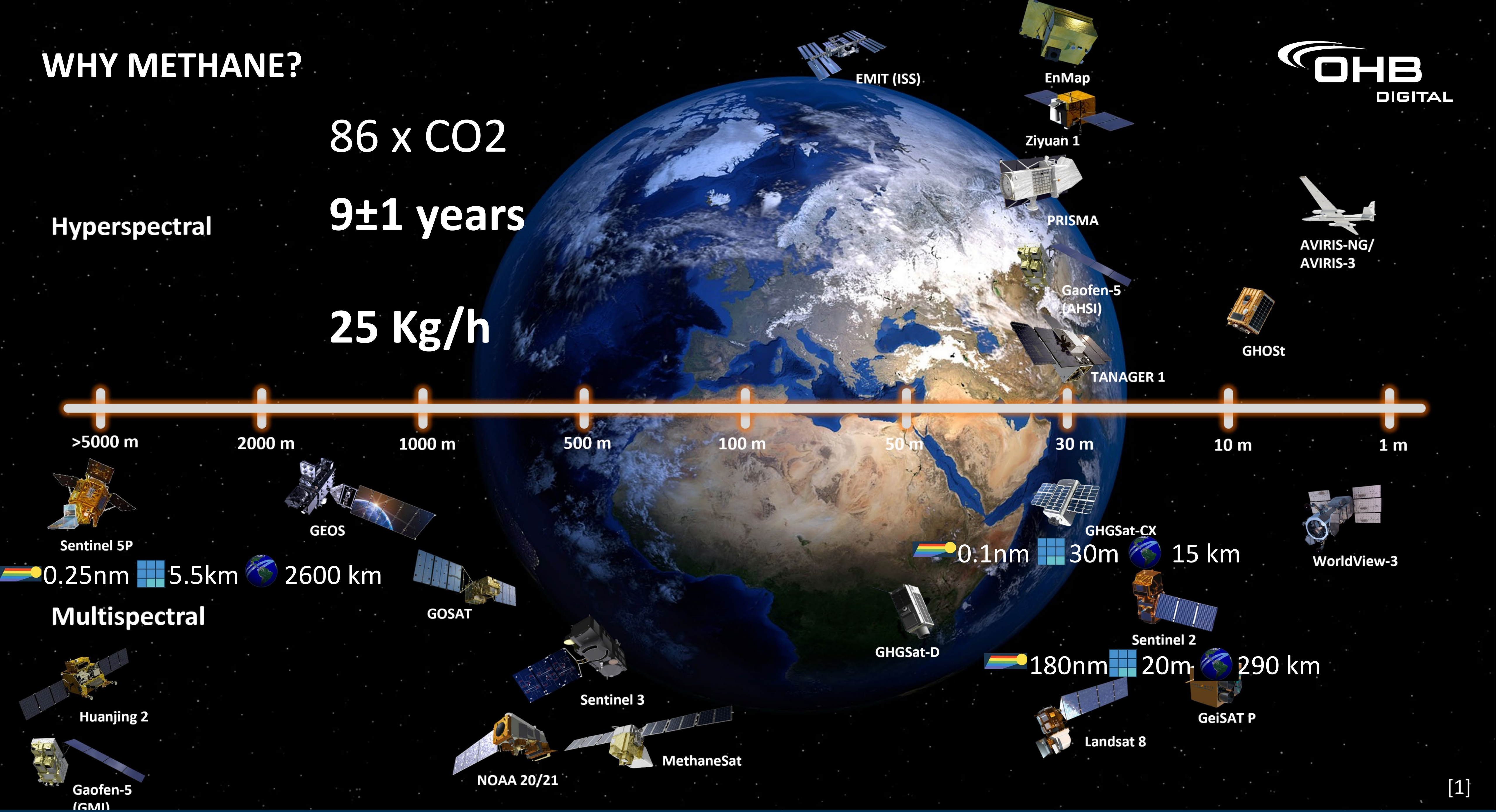
selected via the Open Space Innovation Platform (https://ideas.esa.int) as a Co-Sponsored Research Agreement and carried out under the Discovery programme of, and funded by, the European Space Agency

AGENDA

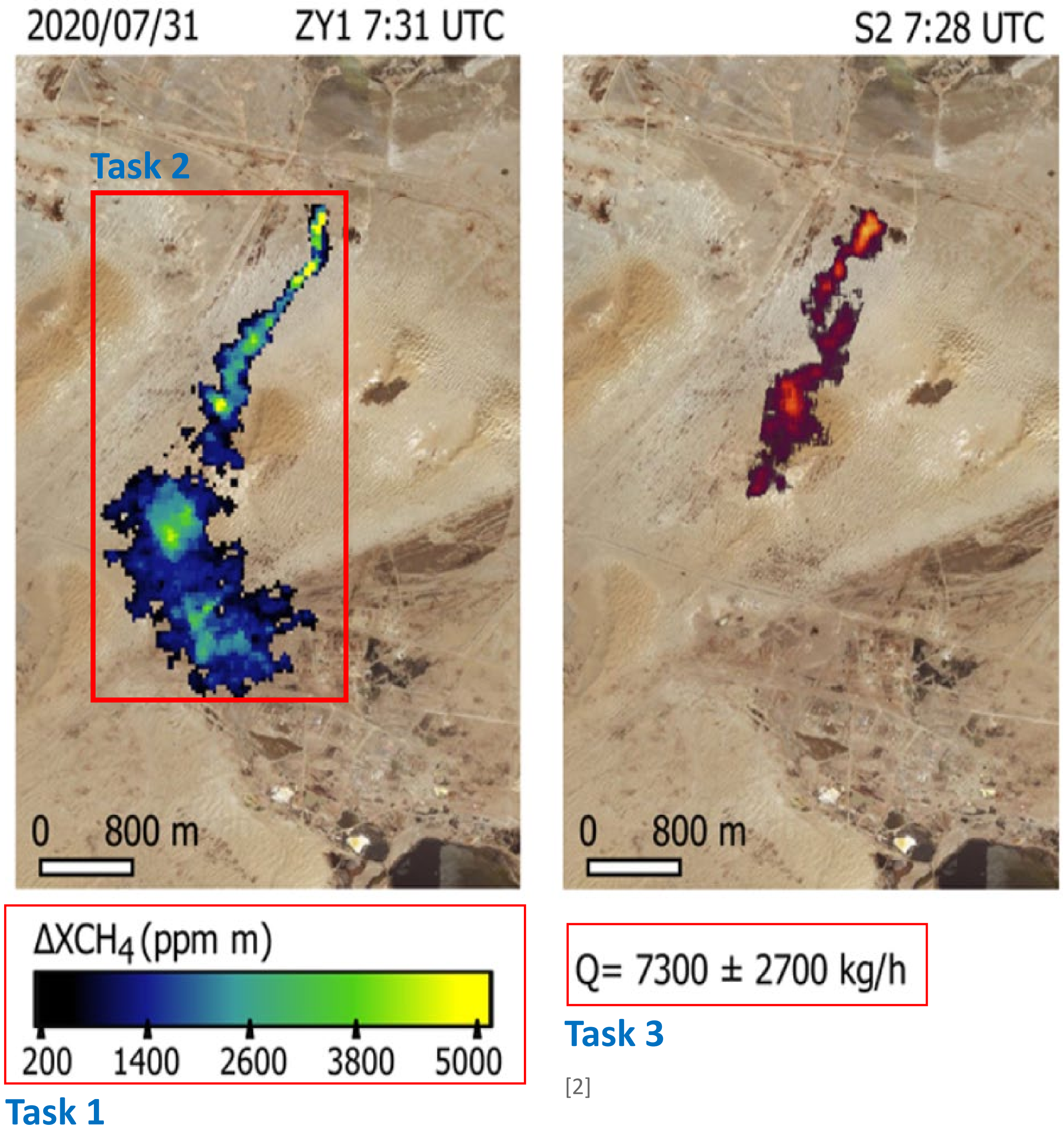
01

MOTIVATION

WHY METHANE?



[1]



[2]

METHODS & CURRENT LIMITATIONS

IS ML WORTH IT?

- SOTA: **Band ratio** (Multi-Band-Multi-Pass) → **Threshold** → **IME** (Integrated Mass Enhancement) at **1.6 µm** & **2.3 µm**
  - Each step needs specific tuning per scene (wind, reference, limits, refinements)
  - Detection limit: ~1 t/h
  - Accuracy: +/- 70% in some cases (Sentinel-2)

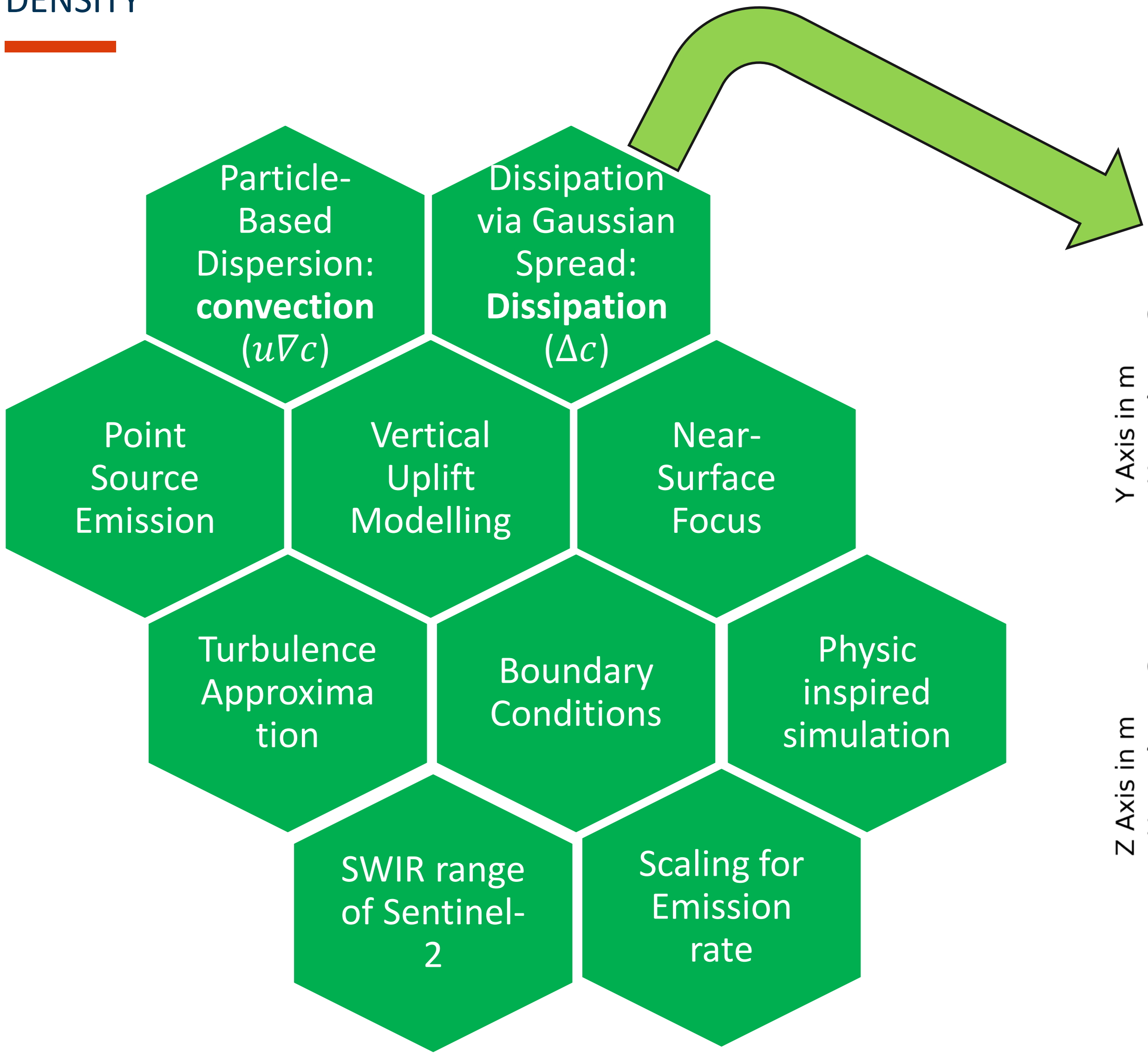
- **Motivation:** ML for improved automated methane monitoring
- Challenges: Data & ground truth -> e.g. via **synthetic data**
- LES with turbulence models are heavy and time intensive -> **Need for lightweight simulation** that are good enough
- We are aiming for a fast, automatic, generalized and robust approach -> Hard to do so let's start small ☺

02

SYNTHETIC DATA GENERATION

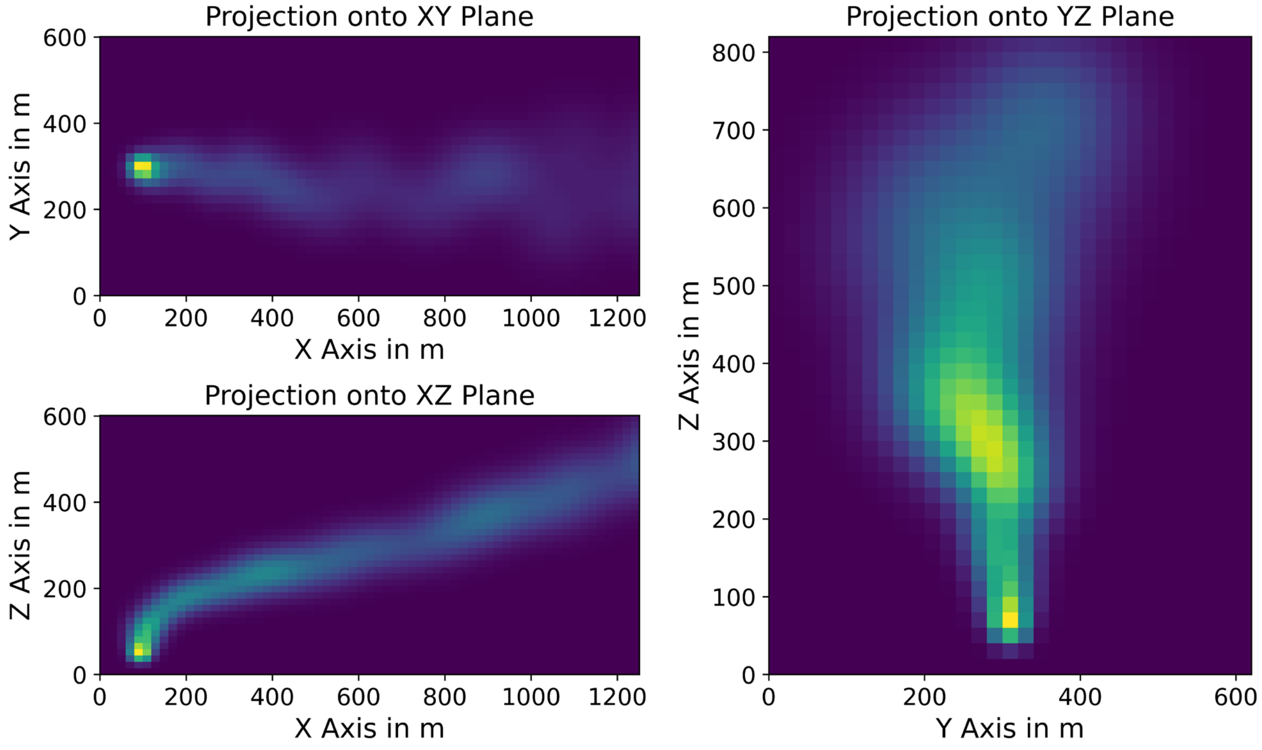
PROPOSED LIGHTWEIGHT SIMULATION

PARTICLE SIMULATION BASED ON WIND FIELDS & GAUSSIAN DENSITY



Calculation of concentration at specific point

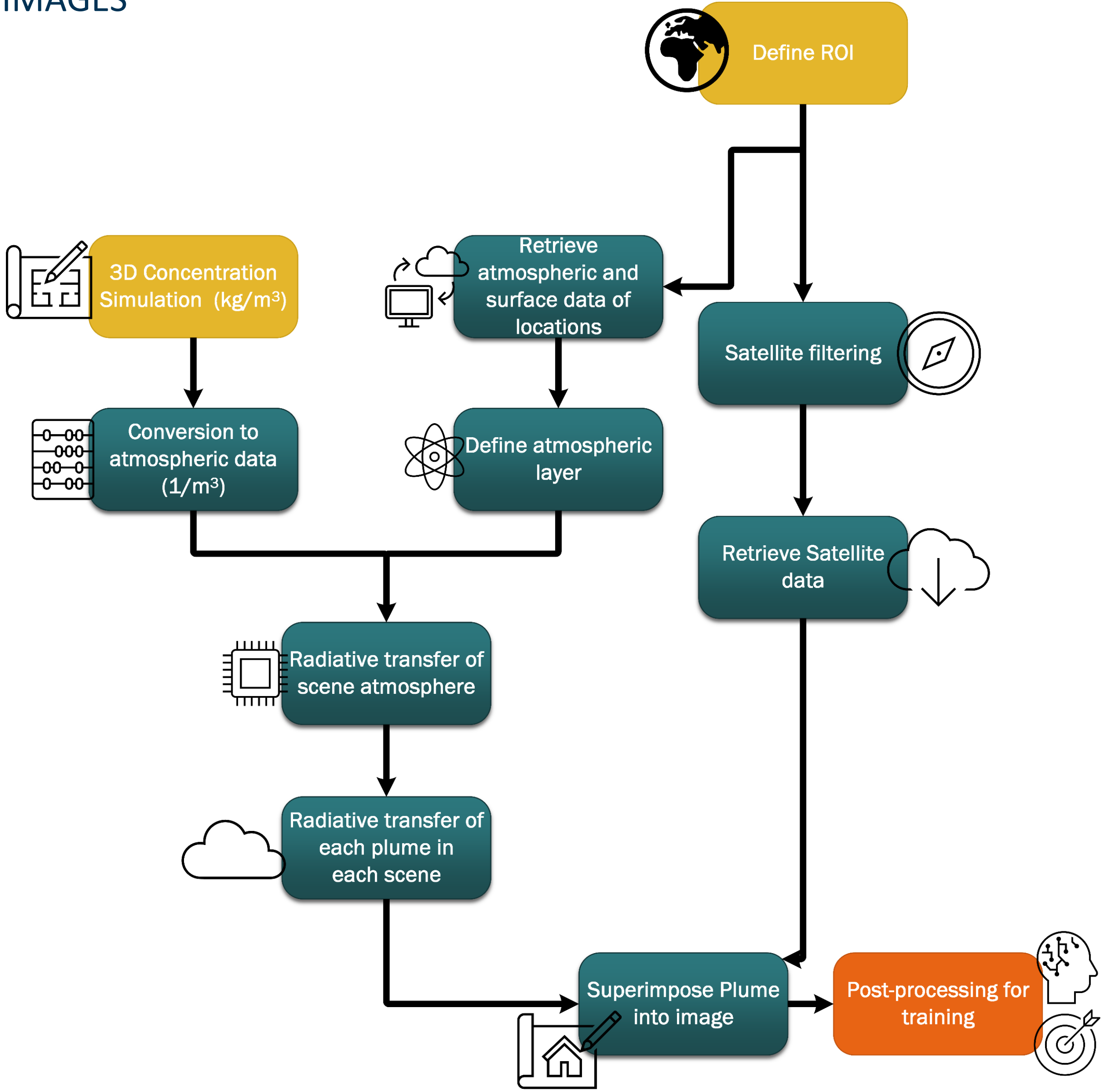
$$C(x) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma|^{\frac{1}{2}}} * \exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))$$



RADIATIVE TRANSFER AND WORKFLOW

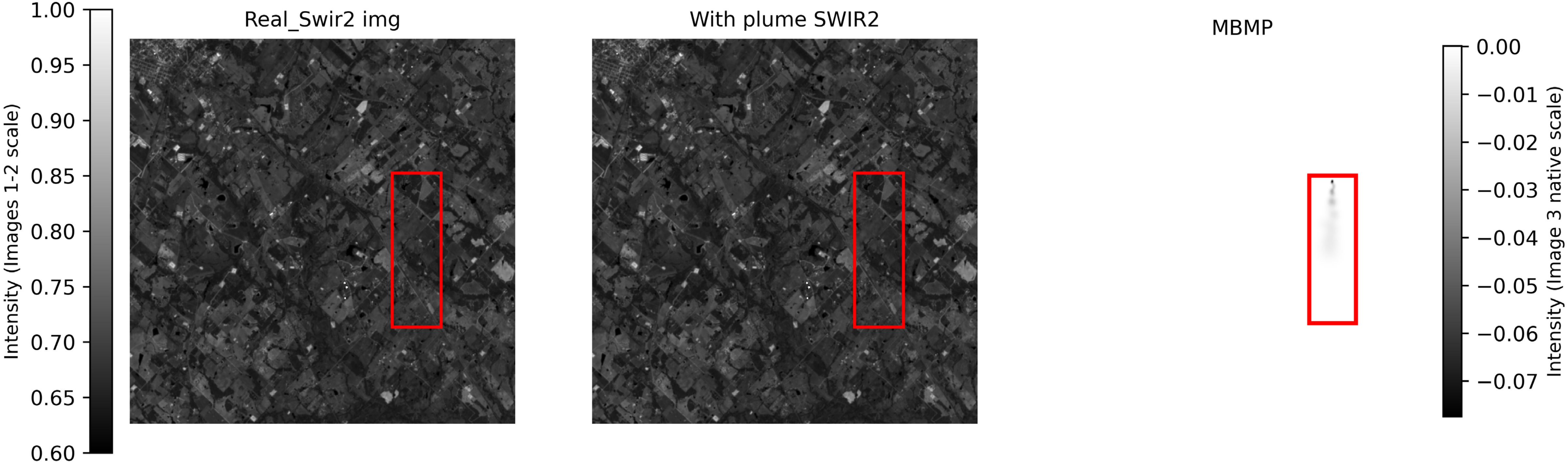
CONCENTRATIONS CH_4 → TOA ABSORPTION → SYNTHETIC PLUME IMAGES

- **Idea**: Only methane absorptions – else ideal conditions
  - Realistic noise through real remote sensing images
- Use Py4cATs (by DLR [3]) – 1D Radiative Transfer Model
- 1D slow for images → 2D adaptation
  - Fixed fine spectral grid interpolation
- Convolution with satellite specific spectral response function
- Integration of ideal ToA methane reflectance into ToA reflectance of real sensor observations
- Theoretical 3-4 % absorption in SWIR2 at 1000 kg/h envisaged



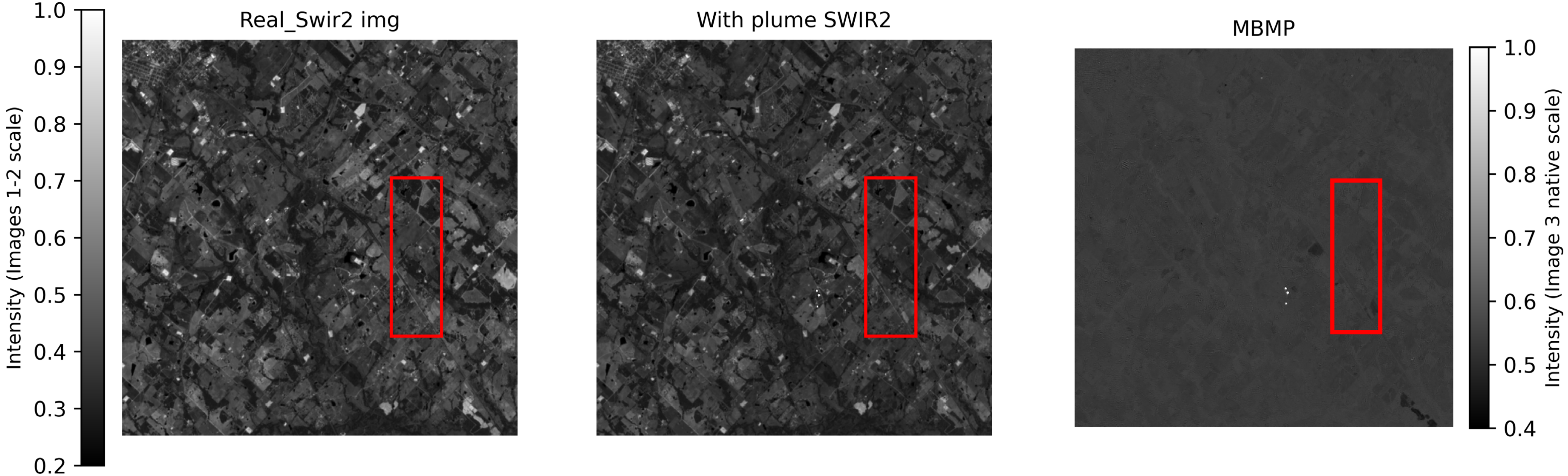
SUPERIMPOSED VS BASE IMAGE

1000 KG/H



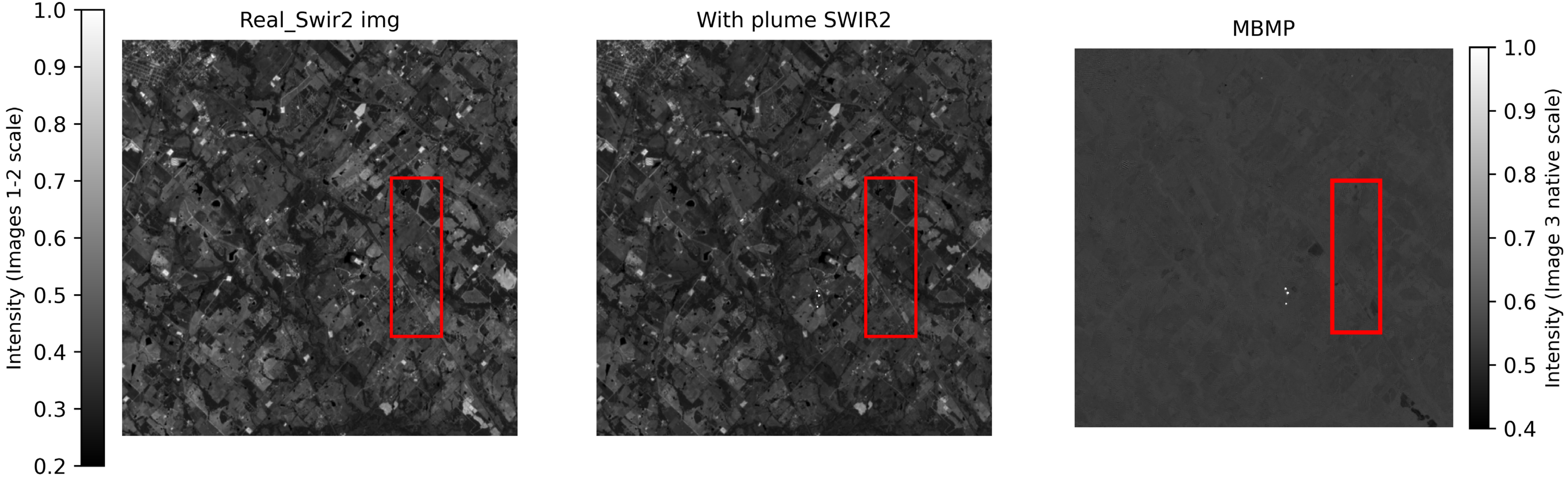
SUPERIMPOSED VS DIFFERENT TIMESTAMP

1000 KG/H



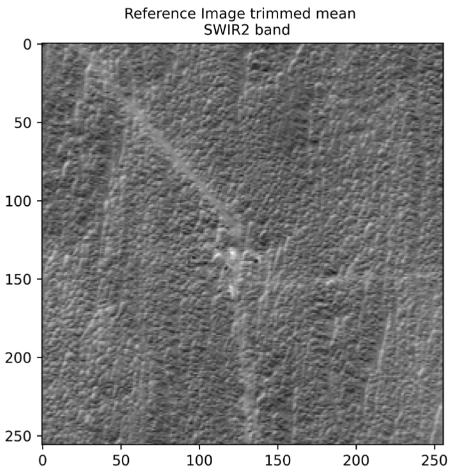
SUPERIMPOSED VS DIFFERENT TIMESTAMP

5000 KG/H



03

REAL-WORLD DETECTION DATASET

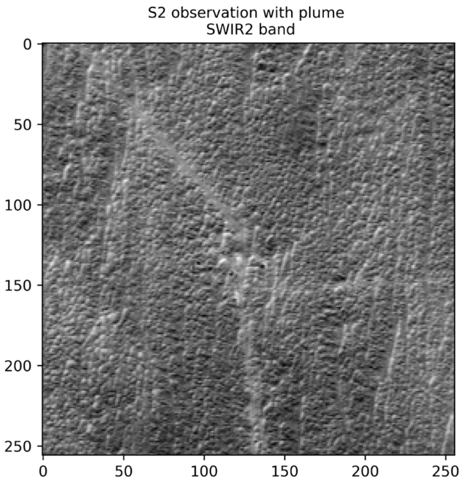


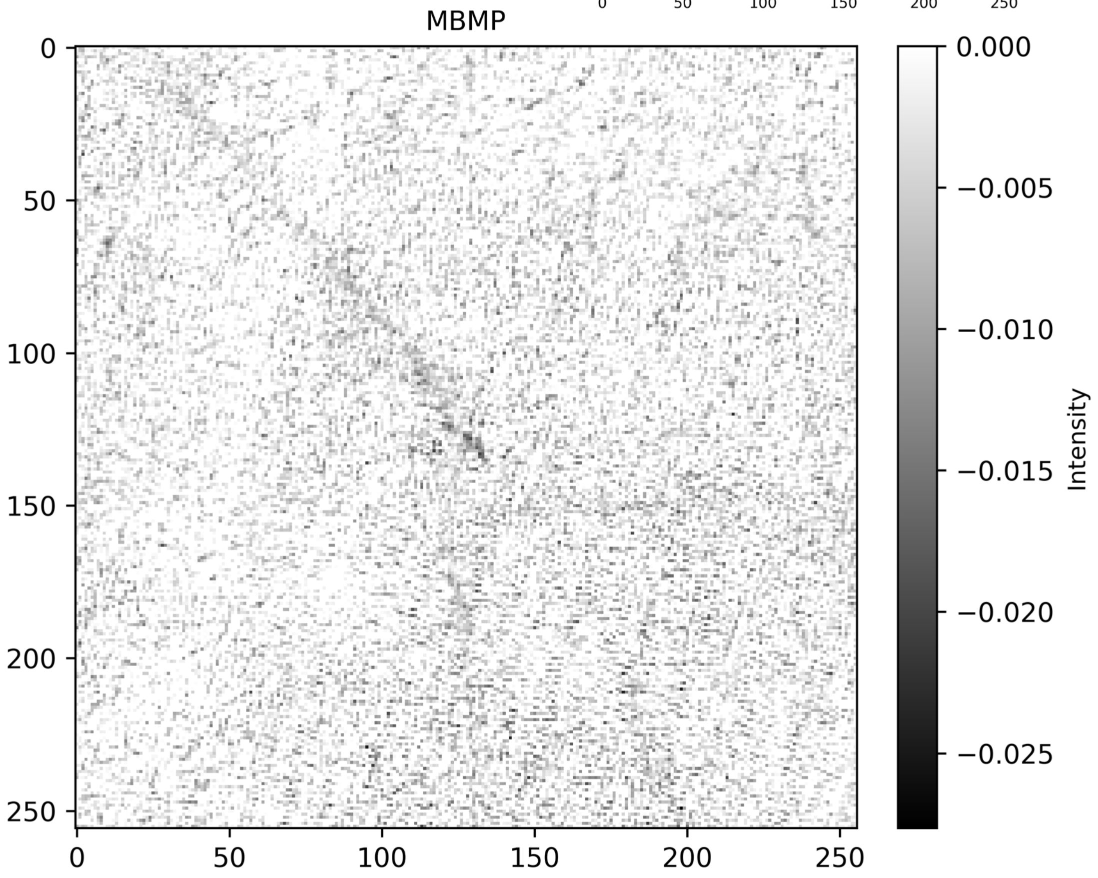
12.02.2023

~2882 kg/h

1174 kg/h std

5.8 m/s





REAL METHANE DETECTIONS

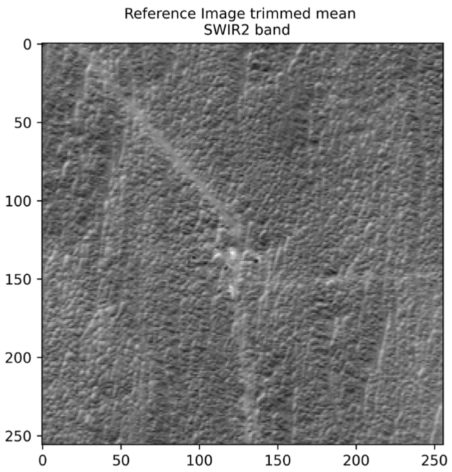
OF SENTINEL 2

- Source: Public methane database - International Methane Emissions Observatory [4]
- Contains: Polygons of plumes, detected emission rate and std, location, product name, wind information - of multiple satellites
- Plume or no plume?

REAL METHANE DETECTIONS

OF SENTINEL 2

- Source: Public methane database - International Methane Emissions Observatory [4]
- Contains: Polygons of plumes, detected emission rate and std, location, product name, wind information - of multiple satellites
- Plume or no plume? -> Exact detection approach & parameters are not specified

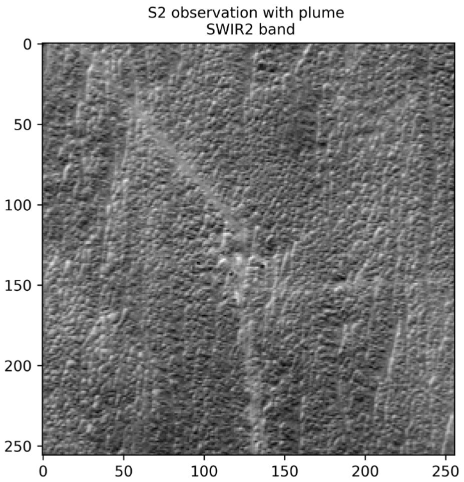


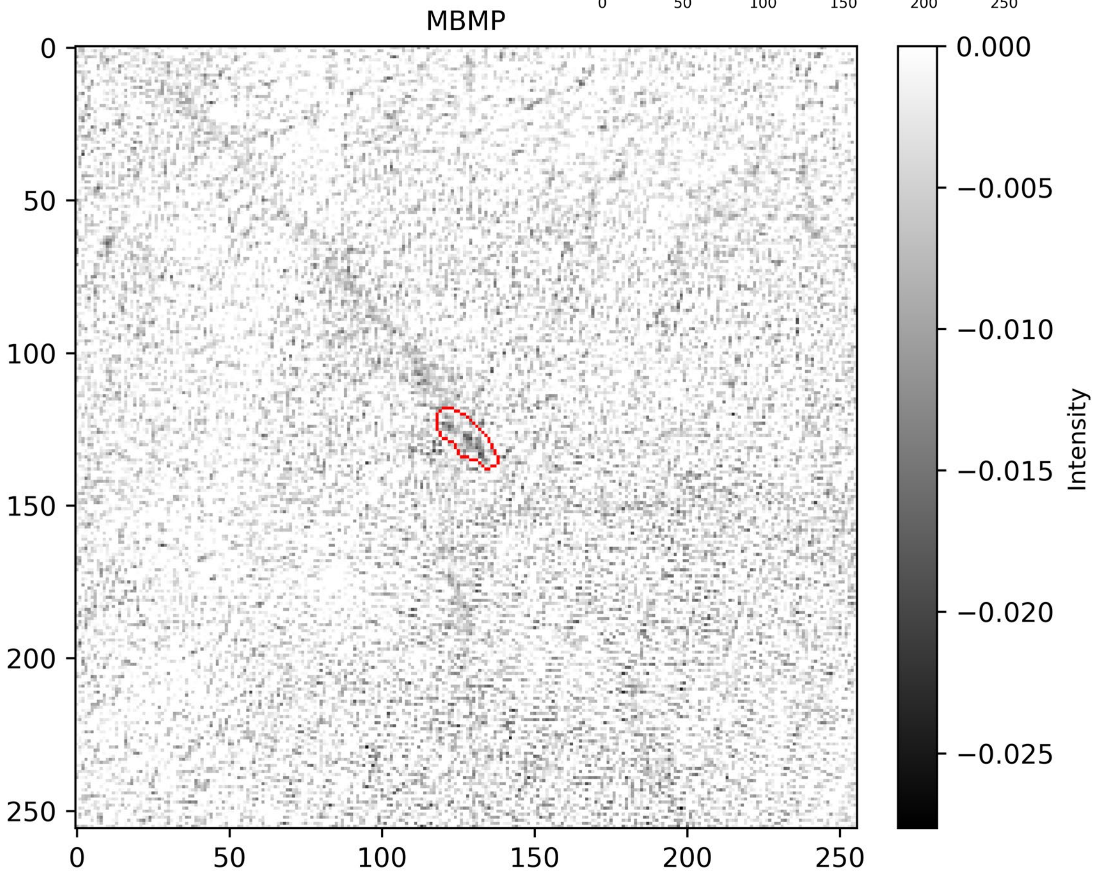
12.02.2023

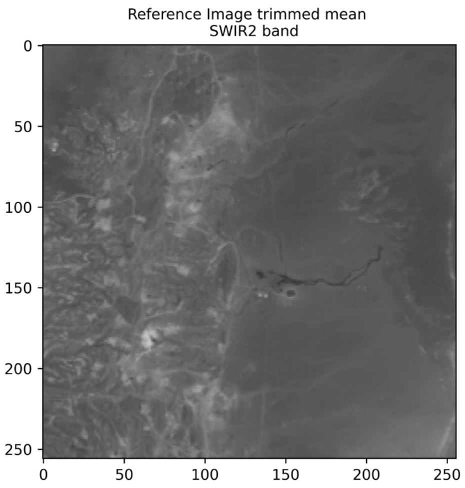
~2882 kg/h

1174 kg/h std

5.8 m/s





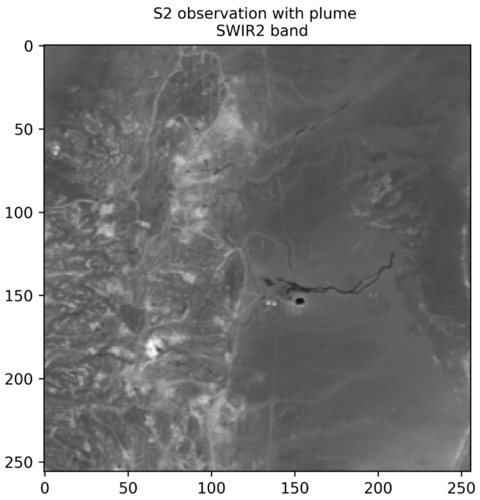


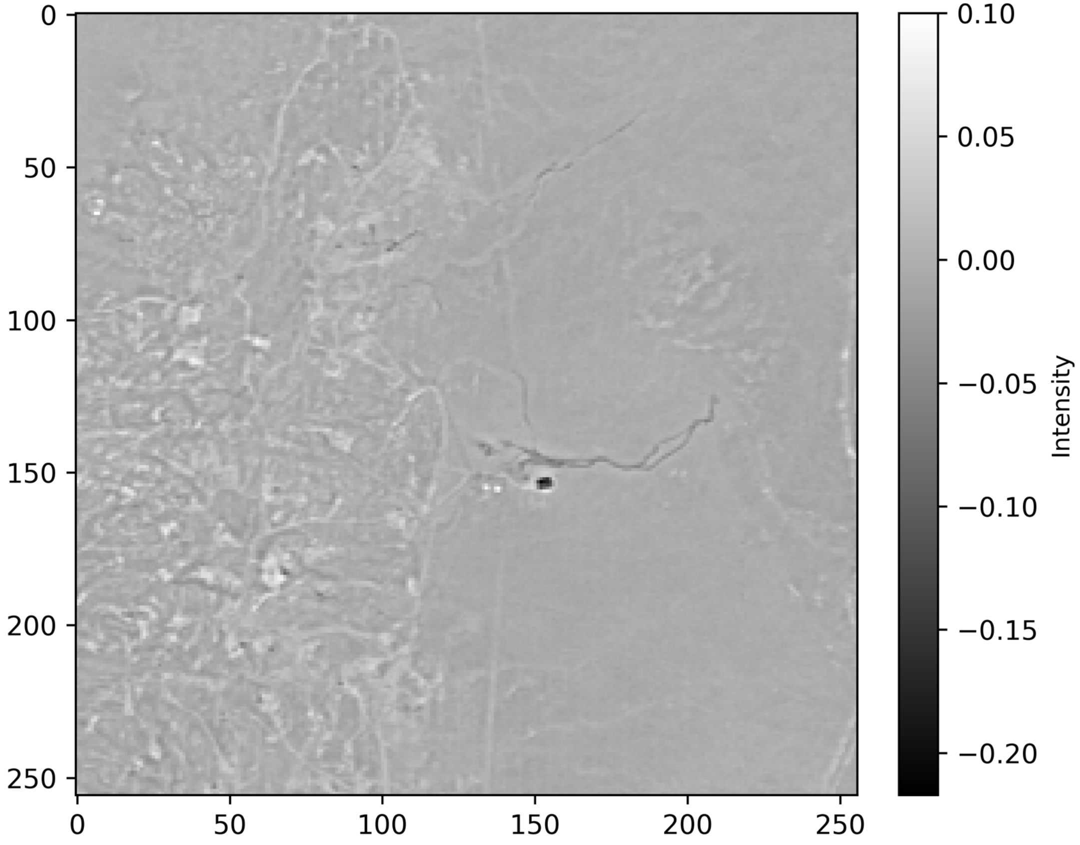
17.05.2021

~6336 kg/h

2575 kg/h std

5.7 m/s

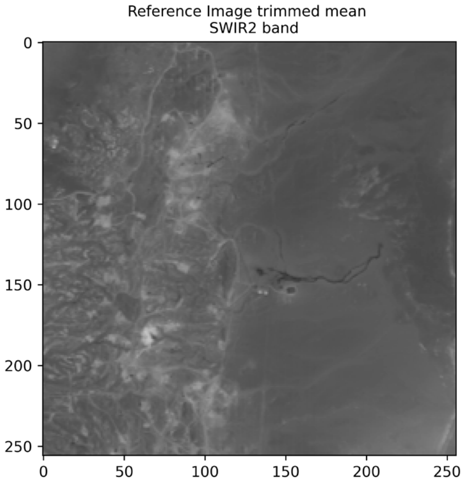




REAL METHANE DETECTIONS

OF SENTINEL 2

- Source: Public methane database - International Methane Emissions Observatory [4]
- Contains: Polygons of plumes, detected emission rate and std, location, product name, wind information - of multiple satellites
- Plume or no plume? -> Exact detection approach & parameters are not specified

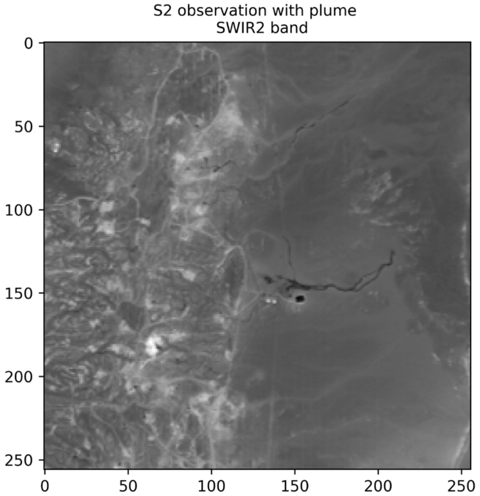


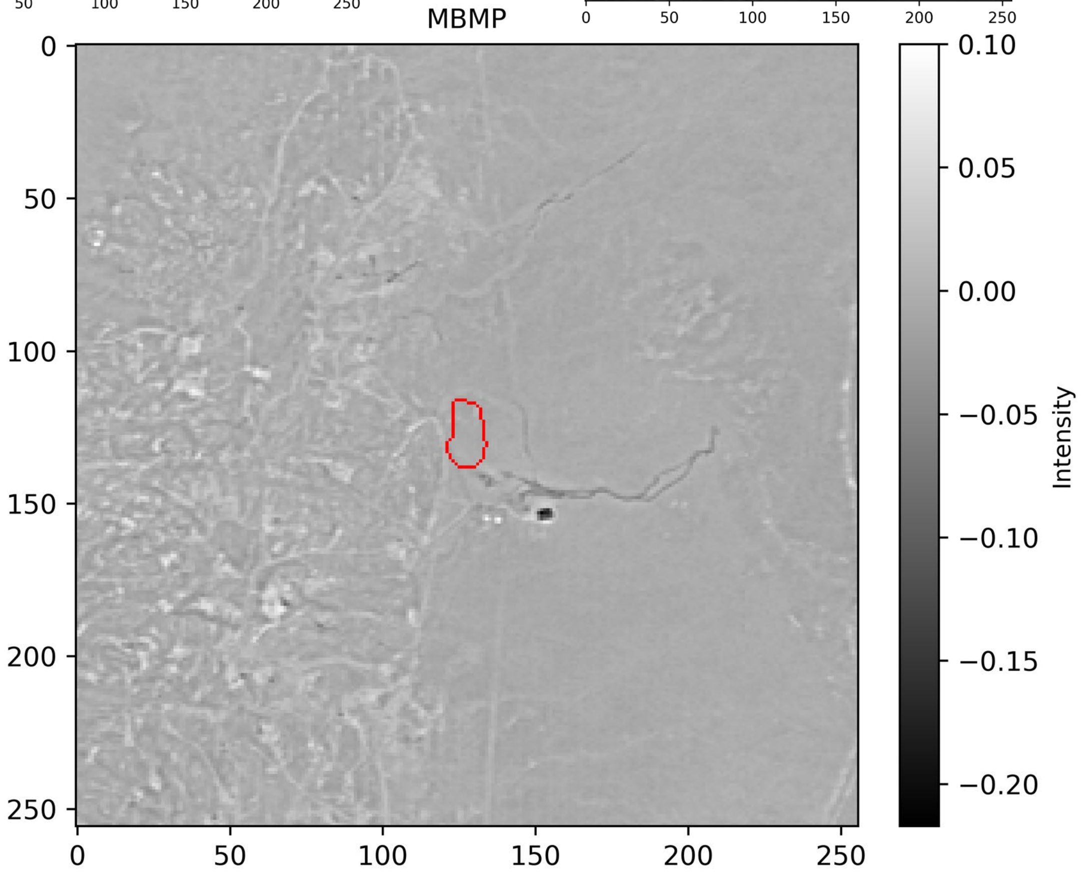
17.05.2021

~6336 kg/h

2575 kg/h std

5.7 m/s





REAL METHANE DETECTIONS

OF SENTINEL 2

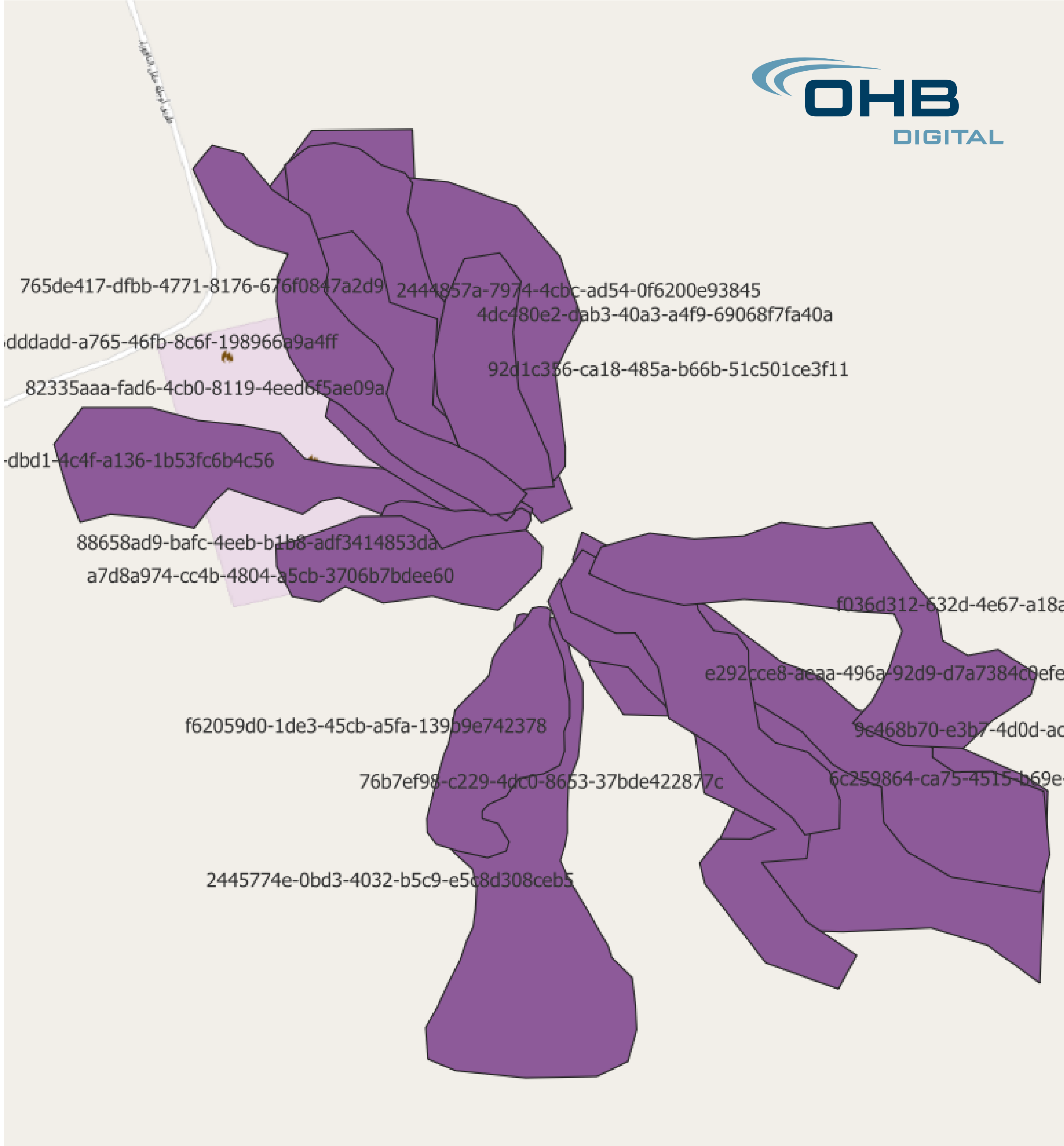
- Source: Public methane database - International Methane Emissions Observatory [4]
- Contains: Polygons of plumes, detected emission rate and std, location, product name, wind information - of multiple satellites
- Plume or no plume? -> Exact detection approach & parameters are not specified
- Many plumes on same location
  – Overall, 173 location cluster, 3515 detection over land

- **Validation** via MBMP – needs multiple observations, compares bands with intensity alignment
  – Usage of trimmed mean and median of +/-7 (mainly) cloud free observations as methane free reference observation
- IMEO data is well suited real-world Validation and Test set
  – Filtering on land scenes
  – Building ML-ready Dataset

DATASET LOCATIONS & INFORMATION

BASED ON DETECTION FROM IMEO

- Training set (70%) remaining
  - Either real detections – ~**2500 plume samples**
  - or synthetic plumes – randomized plume, random location of **33k overall scenes**
- Validation set **527 plume samples**
  - 53 plumes from 2 geographic clusters
  - 474 plumes randomly sampled
- Test set **528 plume samples**
  - 53 plumes from 3 geographic clusters
  - 475 plumes randomly sampled
- Background flat, sandy

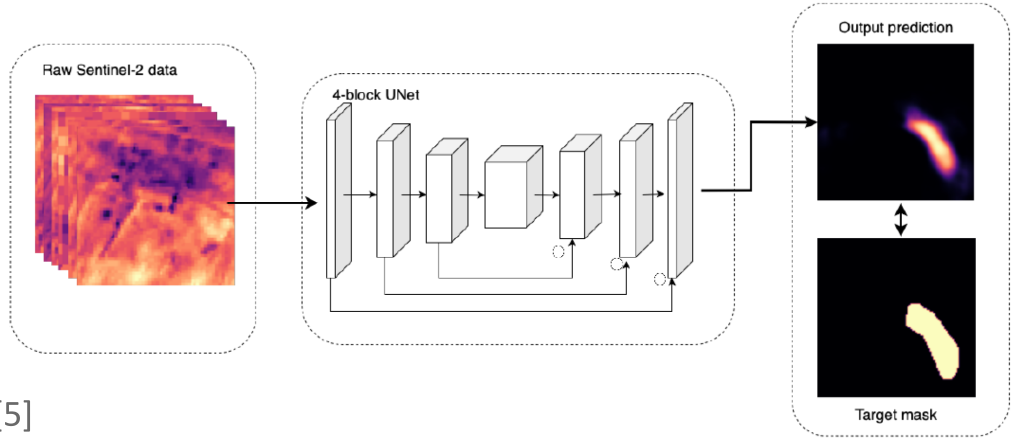


04

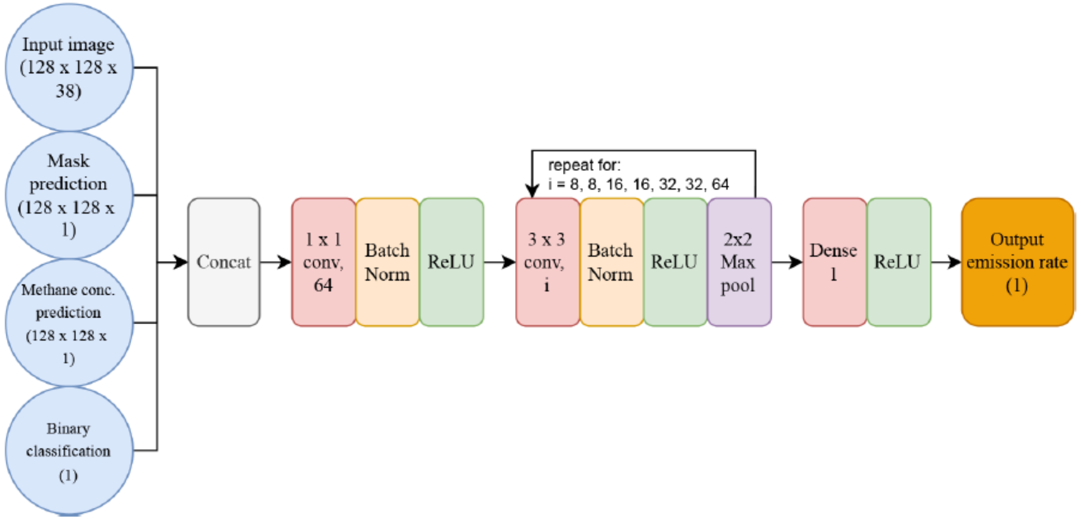
ML APPROACH AND RESULTS

MACHINE LEARNING APPROACHES

PLUME SEGMENTATION AND EMISSION RATE ESTIMATION



[5]

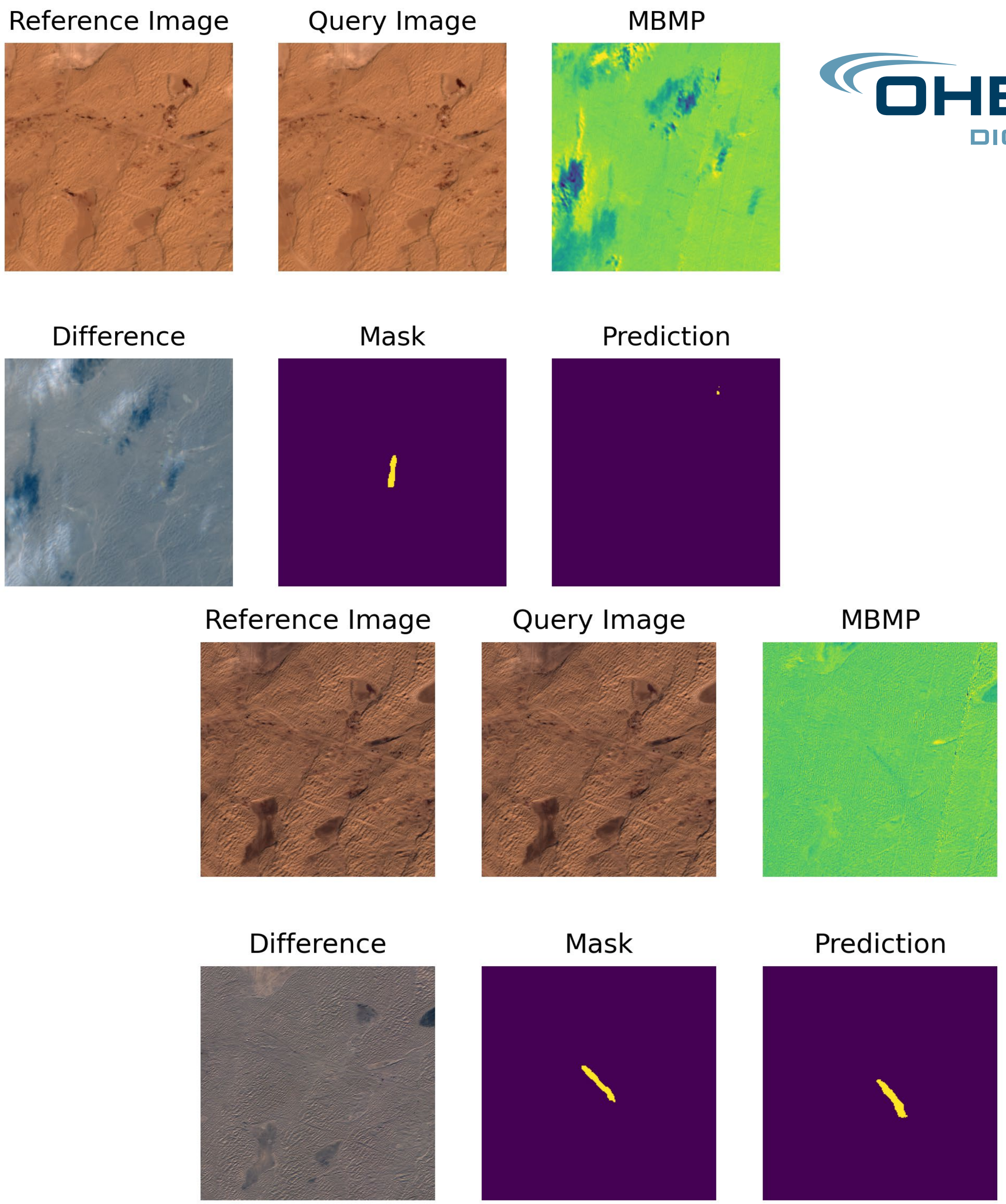


[6]

- About 15 publications apply ML to methane monitoring
  - Mostly CNN-based architectures using U-net
- Benefits:
  - Better detection limit (esp. for low SNR)
  - More robust (lower uncertainty)
  - Automatic processing (no parameter tuning needed)
- Challenges:
  - Few training datasets
  - Intensive work for manually annotation of plumes
  - (Nearly) no ground truth for emission rates
  - Missing common ground for model comparison

ML APPROACH & DATA AUGMENTATION

- Treating synthetical plume integration into images as data augmentation
- Suitable with Kornia for augmentation on GPU
- Flexible with simulated plumes – simulation + radiative transfer needed
- Our approach: **Change Detection on two images**
  - Flexible & data driven approach
  - Allows integration of auxiliary data
- 1st step: Segmentation of plume (later emission rate)
- Comparison of models:
  - Siamese U-net with Ensemble Channel Attention Module (ECAM) (SNU-net)
  - Fully-convolutional Siamese Concatenation / Difference (FC-Siam-conc / FC-Siam-diff)



PRELIMINARY RESULTS

WORK IN PROGRESS

- No hyperparameter tuning, only binary segmentation
- Fcsiamconc, Fcsiamdiff pretrained backbone on Imagenet

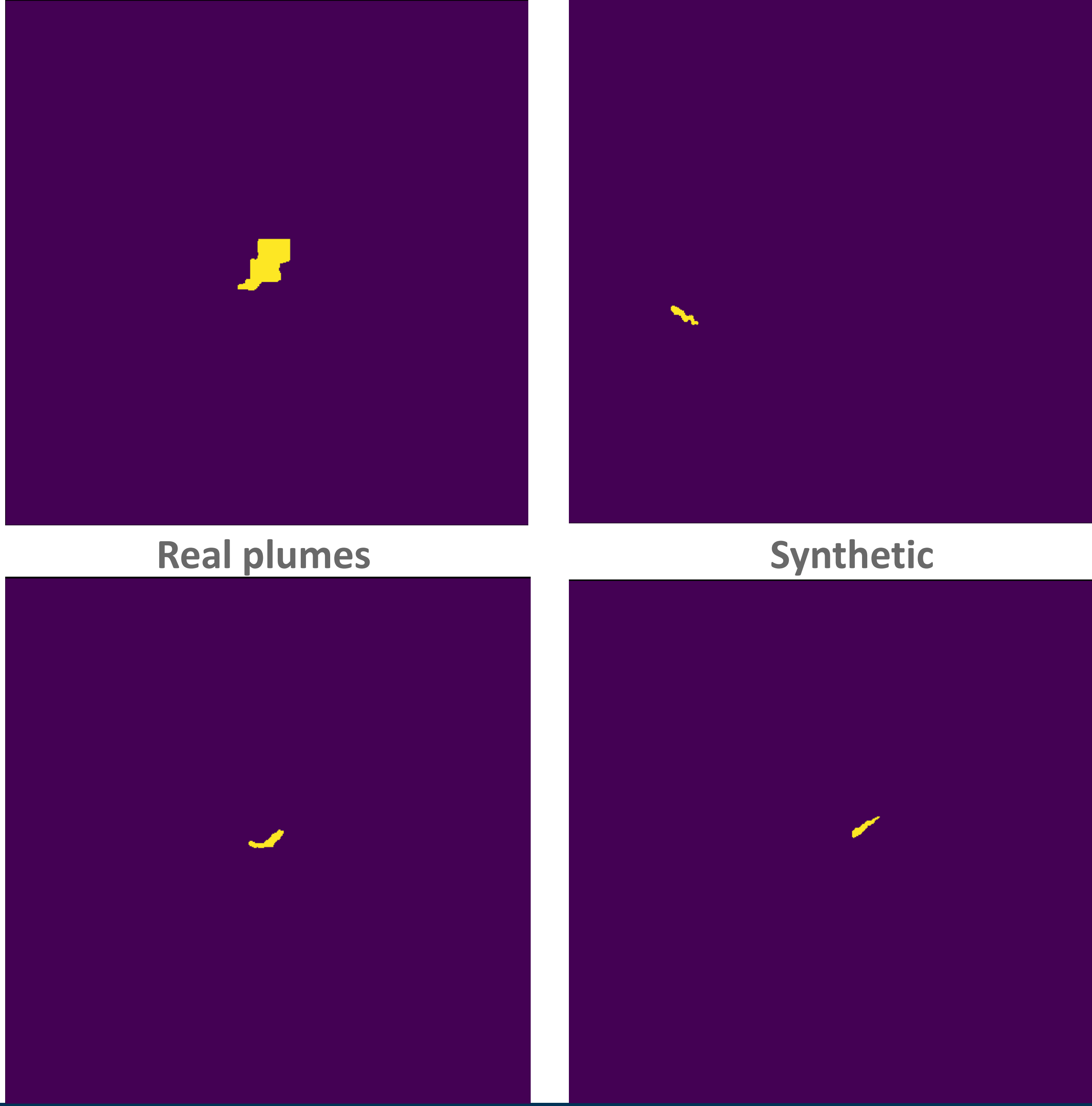
Method	Trained on	Precision	F1-Score	Recall
FcSiam-conc [7] (resnet50)	Real dataset	0.731	**0.555**	**0.447**
FcSiam-diff [7] (resnet18)	Real dataset	**0.805**	0.436	0.299
SNU-net 16 [8]	Real dataset	0.719	**0.555**	0.452
SNU-net 16	Synthetic dataset	0.280	0.151	0.175
SNU-net 16	Synthetic + real (2/27)	0.542	0.351	0.259

05

CONCLUSION & FUTURE WORK

CONCLUSION

- Dataset for real-world & synthetic methane plumes based on Sentinel-2 has been created
- Good first results for ML-based detection on real data
- First promising detection results with training on synthetic data
  - Integration of synthetic data into training process needs to be investigated (e.g., domain adaptation)
  - Adaptation of synthetic plumes to real-world conditions
  - Generate more diverse synthetical data with standard atmospheres
  - Models are based on U-net – denoising of plume?



NEXT STEPS

- Direct estimation of emission rate from observations
- Timeseries analysis for improved detection in a given scene
- Inclusion of other data such as CAMS & wind information

- Planned publication:
  - Detailed information on data generation
  - **Data** of absorption per pixel, IMEO derivative S2 Scenes, masks, reference image
  - Baseline models
  - Update of V1 Dataset [9]

REFERENCES

- [1]: Tiemann, Enno, et al. "Machine Learning for Methane Detection and Quantification From Space: A survey." IEEE Geoscience and Remote Sensing Magazine (2025).
- [2] Irakulis-Loitxate, Itziar, et al. "Satellites detect abatable super-emissions in one of the world's largest methane hotspot regions." Environmental science & technology 56.4 (2022): 2143-2152.
- [3]: Schreier, Franz, et al. "Py4cats—PYthon for computational ATmospheric spectroscopy." Atmosphere 10.5 (2019): 262.
- [4] UNEP IMEO Methane Sources and plumes https://methanedata.unep.org/ BY-NC-SA 4.0
- [5]: A. Vaughan, G. Mateo-García, L. Gómez-Chova, V. Růžička, L. Guanter, and I. Irakulis-Loitxate, "Ch4net: A deep learning model for monitoring methane super-emitters with Sentinel-2 imagery," Atmos. Meas. Tech., vol. 17, no. 9, pp. 2583–2593, 2024, doi: 10.5194/amt-17-2583-2024.
- [6] Joyce, Peter, et al. "Using a deep neural network to detect methane point sources and quantify emissions from PRISMA hyperspectral satellite images." Atmospheric Measurement Techniques 16.10 (2023): 2627-2640.
- [7] Daudt, Rodrigo Caye, Bertr Le Saux, and Alexandre Boulch. "Fully convolutional siamese networks for change detection." 2018 25th IEEE international conference on image processing (ICIP). IEEE, 2018.
- [8] Fang, Sheng, et al. "SNUNet-CD: A densely connected Siamese network for change detection of VHR images." IEEE Geoscience and Remote Sensing Letters 19 (2021): 1-5.
- [9]: Tiemann, E. GeoCH4PlumeNet - Global Dataset for Methane Point Source Detection and Quantification. Zenodo, 15 Aug. 2025, https://doi.org/10.5281/zenodo.16813369.

THANK YOU!

Enno Tiemann
PhD candidate / Systems Engineer
OHB Digital Connect GmbH

Mail: enno.tiemann@ohb.de
Web: https://ohb-dc.de/

selected via the Open Space Innovation Platform (https://ideas.esa.int) as a Co-Sponsored Research Agreement and carried out under the Discovery programme of, and funded by, the European Space Agency