

---

# Towards Methane Detection On Board Satellites

---

**Maggie Chen\***

University of Oxford

meiqi.chen@physics.ox.ac.uk

**Hala Lambdouar\***

University of Oxford

lamdouar@robots.ox.ac.uk

**Luca Marini\***

Delft University of Technology

1.marini@tudelft.nl

**Laura Martínez-Ferrer**

Universitat de València, Spain

laura.martinez-ferrer@uv.es

**Chris Bridges**

University of Surrey

c.p.bridges@surrey.ac.uk

**Giacomo Acciarini**

European Space Agency (ESA)

giacomo.acciarini@esa.int

## Abstract

Methane is a potent greenhouse gas and a major driver of climate change, making its timely detection critical for effective mitigation. Machine learning (ML) deployed on board satellites can enable rapid detection while reducing downlink costs, supporting faster response systems. Conventional methane detection methods often rely on image processing techniques, such as *orthorectification* to correct geometric distortions and *matched filters* to enhance plume signals. We introduce a novel approach that bypasses these preprocessing steps by using *unorthorectified* data (UnorthoDOS). We find that ML models trained on this dataset achieve performance comparable to those trained on orthorectified data. Moreover, we also train models on an orthorectified dataset, showing that they can outperform the matched filter baseline (mag1c). We release model checkpoints and two ML-ready datasets comprising orthorectified and unorthorectified hyperspectral images from the Earth Surface Mineral Dust Source Investigation (EMIT) sensor at <https://huggingface.co/datasets/SpaceML/UnorthoDOS>, along with code at <https://github.com/spaceml-org/plume-hunter>.

## 1 Introduction

Methane has a global warming potential approximately 84 times greater than that of carbon dioxide over a 20-year period [1]. A significant share of point-source methane emissions originates from “super-emitters” in the oil and gas sector, which are detectable from space. With the rising interest in wide-coverage, high-resolution Earth observation (EO), hyperspectral remote sensing imagery provides a unique opportunity to deploy ML models on board spacecraft for rapid greenhouse gas detection. Reliable methane detection at this scale offers a pathway to rapid response and effective mitigation efforts. Methane plume detection in hyperspectral satellite imagery typically involves two key processing steps: (1) *orthorectification*, which corrects geometric distortions caused by sensor viewing angle, terrain variations, and Earth curvature, and (2) the generation of methane enhancement

---

\*These authors contributed equally and are listed in alphabetical order.

products, often using *matched filters* [2, 3, 4] to strengthen weak plume signals by comparing image data against predefined spectral signatures. Both steps are computationally intensive, making them challenging to execute in resource-constrained environments on board spacecraft. Matched filters are also prone to high false-positive rates. Studies such as [5] improved segmentation accuracy by combining matched filter outputs with Red Green Blue (RGB) images in deep learning frameworks. More recently, end-to-end approaches that process hyperspectral bands directly with lightweight neural networks [6] further advanced plume detection. These methods remain reliant on orthorectified datasets. However, orthorectification is a costly process to perform in real time on board spacecraft [7]. To address this limitation, we present a novel unorthorectified dataset that more closely reflects the data obtained under realistic onboard conditions and demonstrate that methane plumes can also be reliably detected from less-processed satellite imagery.

## 2 Methodology

**Methane Hyperspectral data** Our datasets are constructed from hyperspectral observations acquired by the EMIT imaging spectrometer, flown on the International Space Station (ISS) and distributed via the NASA Earthdata portal<sup>2</sup>. Specifically, the L1B hyperspectral imagery [8] and L2B methane plume complexes [9], which serve as ground-truth annotations are used. Both products are provided at a spatial resolution of 60 m. In total, we curated 1,574 plume complexes acquired between August 10, 2022, and October 26, 2024. For the purpose of methane detection, a subset of 86 bands with wavelengths ranging between 1573-1699 nm and 2004-2478 nm is selected from the original 285 hyperspectral bands in the L1B images, following the procedure described in [6]. This subset captures the methane absorption spectrum, along with non-absorption and RGB bands, enabling ML models to better discriminate methane signals from background features.

**UnorthoDOS dataset generation pipeline** The schematic in Figure 1 illustrates the method used for generating our unorthorectified data. Orthorectification is defined as a mapping  $\tau$  that transforms pixel coordinates  $(x, y)$  from an angled (*off-nadir*) image  $I_{\text{unortho}}$  to orthogonal (*nadir*) coordinates  $(x_o, y_o)$  in the target image  $I_{\text{ortho}}$ , compensating for sensor geometry and terrain effects (see Figure 1). The transformation  $\tau$  is extracted from a geometric lookup table in the EMIT satellite data.

$$\begin{aligned} \tau : \quad I_{\text{unortho}} &\longrightarrow I_{\text{ortho}} \\ (x, y) &\longmapsto \tau(x, y) = (x_o, y_o) \end{aligned}$$

To generate the unorthorectified dataset, the inverse mapping  $\tau^{-1}$  is approximated by reconstructing  $I_{\text{unortho}}$  from  $I_{\text{ortho}}$  by using a pixel coordinate grid in the source (unorthorectified) plane, transformed using  $\tau$ . The transformed grid is then used to sample pixels from the orthorectified plume annotations back in the unorthorectified plane. For notational simplicity, we denote this operation as  $\tau^{-1}$ , although  $\tau$  is non-bijective and the images ( $I_{\text{unortho}}$  and  $I_{\text{ortho}}$ ) may differ in size. In particular, the inverse mapping  $\tau^{-1}$  is non-surjective, meaning that certain pixels in  $I_{\text{unortho}}$  lack a corresponding value in  $I_{\text{ortho}}$ ; these are filled using nearest-neighbor interpolation. Formally, our procedure for generating approximate *off-nadir* annotations can be expressed as:

$$\hat{I}_{\text{unortho}} = \Phi \circ \tau^{-1}(I_{\text{ortho}}),$$

where  $\Phi$  denotes the nearest-neighbor interpolation operator. Finally, for the purpose of ML training, both the hyperspectral images and their unorthorectified annotations are divided into smaller tiles. We refer to the resulting unorthorectified dataset as the **Unorthorectified Dataset for On Board Satellite methane detection (UnorthoDOS)** [10].

**Orthorectified benchmark** For comparison, an orthorectified dataset was generated using the orthorectified counterparts of the same L1B images used in the unorthorectified dataset, while retaining the original methane annotations from the L2B products. The orthorectified images are then added and divided into tiles of  $128 \times 128$  pixels, keeping only those containing over 80% non-padded pixels.

**Experimental setup** We present a simplified *tip and cue* paradigm [11], where a tip satellite coarsely localizes methane plumes, and a cue satellite refines the detection. The tip conducts image classification to determine whether a methane plume is present, while the cue uses semantic segmentation to map its extent (see Figure 2). The datasets are split to reserve a holdout test set of

---

<sup>2</sup><https://search.earthdata.nasa.gov/>

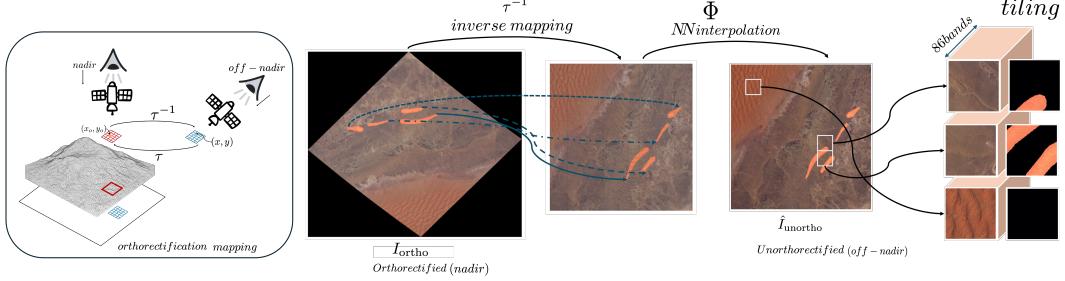


Figure 1: Schematic representation of the unorthorectified data generation process.

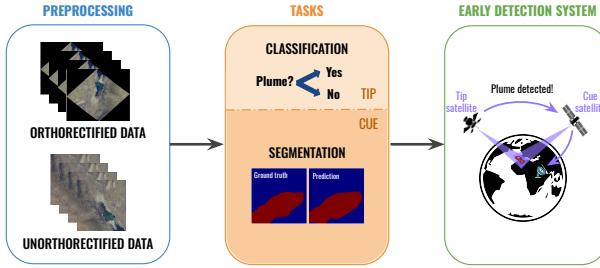


Figure 2: **Overview of the detection pipeline.** EMIT images are preprocessed into orthorectified and unorthorectified datasets using spatial jittering as data augmentation. The ML models provide an onboard early detection system by performing classification for the TIP satellite and segmentation for the CUE satellite.

tiles from 5% of L1B images, with the rest divided into training (80%) and validation (15%) sets. Given the limited number of labelled methane plumes, training and validation sets are augmented via spatial jittering with varying offsets to enhance model generalization. Input tiles are normalized by subtracting the mean and dividing by the standard deviation of each band. Per-tile methane plume classification and semantic segmentation are performed using a UNet [12] architecture. Two models are trained independently on the orthorectified and unorthorectified datasets using a Dice loss, the Adam optimizer [13], a learning rate of  $1e-4$ , and a batch size of 32 for 100 epochs. Training is done on a single 40GB NVIDIA A100 GPU. For evaluation, we report precision, recall, and F1-score for both image classification and semantic segmentation tasks. Additionally, we compute accuracy for image classification, and intersection over union (IoU) for semantic segmentation with model selection based on the highest validation IoU. Using the semantic segmentation model, an image tile is classified as containing a plume if at least one pixel has a positive segmentation prediction.

### 3 Results

Table 1 presents the image classification and semantic segmentation performance of models trained using the orthorectified and unorthorectified datasets, evaluated on image tiles in the corresponding holdout test sets. The performance on strong methane plumes, that is, with a maximum methane concentration threshold of  $\geq 900$  parts per million meter (ppm m), is also evaluated for each model and the mag1c [4] baseline. For semantic segmentation, the UNet outperforms the baseline, achieving IoU improvements of 288.03% for weak plumes and 536.36% for strong plumes. Figure 3 shows that UNet more effectively delineates the spatial extent of methane plumes. In contrast, mag1c produces fragmented outputs and exhibits a higher false positive rate, due to the misclassification of terrestrial features as methane emissions. Despite UNet’s superior performance over mag1c, absolute performance in the semantic segmentation of weak plumes remains limited, with IoU between 16% and 19% in both datasets. Both the segmentation and classification performances improve substantially for strong plumes in tiles with maximum concentration  $\geq 900$  ppm m. For segmentation, the IoU increases by 12.33% and 9.26%, respectively, for the orthorectified and the unorthorectified datasets. Whereas for plume classification on strong plumes, the recall increases by 27.12% for the orthorectified dataset and 22.72% for the unorthorectified dataset. Between the orthorectified and

unorthorectified approaches, the performances on both tasks are comparable. This consistency is observable in Figure 3, where the UNet model yields accurate plume predictions for both image types. In the image classification task, the orthorectified dataset yielded slightly superior results, potentially due to differences in spatial alignment of the two datasets or image quality. Conversely, segmentation performance was slightly better for the model trained on the unorthorectified dataset, with the exception of IoU.

Table 1: Image classification & semantic segmentation results. Best results per task and dataset are in **bold**.

Task	Model	Threshold (ppm m)	Precision	Recall	F1-Score	Accuracy	IoU
<b>Orthorectified</b>							
Img C	mag1c[4]	N/A	52.55	94.12	67.45	54.58	-
		$\geq 900$	39.60	<b>95.24</b>	55.94	46.84	-
Sem S	UNet	N/A	<b>95.56</b>	56.21	70.78	76.80	-
		$\geq 900$	94.60	83.33	<b>88.61</b>	<b>92.41</b>	-
Img C	mag1c[4]	N/A	41.67	15.69	22.80	-	4.76
		$\geq 900$	44.91	15.69	23.25	-	4.84
Sem S	UNet	N/A	79.23	19.91	31.82	-	18.47
		$\geq 900$	<b>82.77</b>	<b>32.52</b>	<b>46.69</b>	-	<b>30.80</b>
<b>Unorthorectified</b>							
Img C	UNet	N/A	<b>89.19</b>	48.89	63.16	71.48	-
		$\geq 900$	87.88	<b>71.61</b>	<b>78.91</b>	<b>85.65</b>	-
Sem S	UNet	N/A	88.41	19.85	32.42	-	16.91
		$\geq 900$	<b>89.08</b>	<b>33.01</b>	<b>48.16</b>	-	<b>26.17</b>

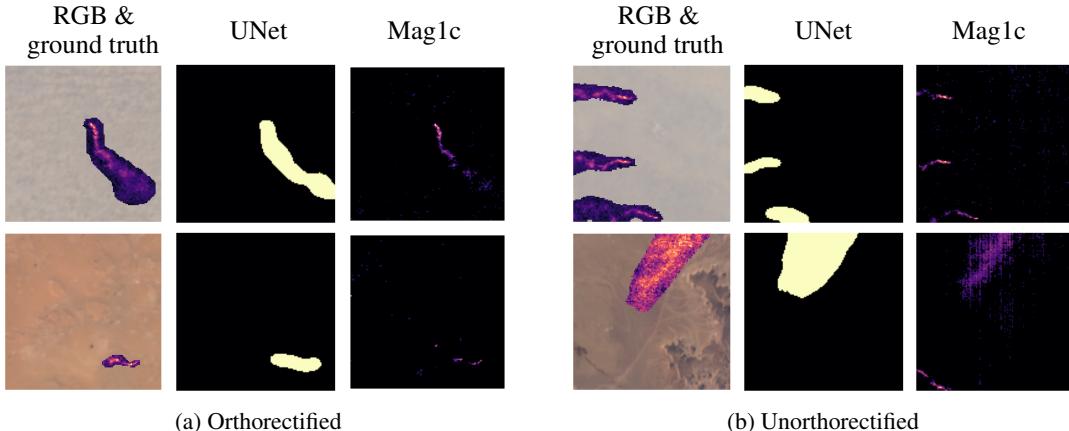


Figure 3: **Visualisation of semantic segmentation results on 2 example tiles from the orthorectified 3a and the unorthorectified 3b datasets each.** From *Left* to *Right* in each sub-figure: L1B tiles (only RGB bands shown for visualisation) overlaid with ground truth methane plume annotations; predicted semantic segmentation plume masks from UNet; segmentation predictions from mag1c [4]

## 4 Conclusions

In this work, we present UnorthoDOS, a novel dataset of unorthorectified satellite imagery designed to support ML-based methane plume detection for on board satellite deployment. We demonstrate that models trained on orthorectified and unorthorectified datasets achieve comparable performance, establishing that plume detection is feasible without orthorectification: a critical advantage for real-time detection in resource-constrained satellite scenarios. A key limitation of our study is the reduced sensitivity to weak methane plumes, which could be substantially mitigated by access to larger annotated hyperspectral image datasets, soon to be enabled by the growing number of satellite missions carrying dedicated sensors.

## Acknowledgments

This work has been enabled by FDL Earth Systems Lab (<https://eslab.ai/>) a public / private partnership between the European Space Agency (ESA), Trillium Technologies and the University of Oxford in partnership with Google Cloud, NVIDIA Corporation, Pasteur Labs and SCAN.

## References

- [1] G. Myhre, D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang. Anthropogenic and Natural Radiative Forcing. In *Climate Change 2013 - The Physical Science Basis*, pages 659–740. Cambridge University Press, 2014.
- [2] Dimitris Manolakis, Eric Truslow, Michael Pieper, Thomas Cooley, and Michael Brueggeman. Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms. *IEEE Signal Processing Magazine*, 31(1):24–33, 2013.
- [3] DR Thompson, I Leifer, H Bovensmann, M Eastwood, M Fladeland, C Frankenberg, K Gerilowski, RO Green, S Kratwurst, T Krings, et al. Real-time remote detection and measurement for airborne imaging spectroscopy: a case study with methane. *Atmospheric Measurement Techniques*, 8(10):4383–4397, 2015.
- [4] Markus D Foote, Philip E Dennison, Andrew K Thorpe, David R Thompson, Siraput Jongaramrungruang, Christian Frankenberg, and Sarang C Joshi. 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, 2020.
- [5] Vít Růžička, Gonzalo Mateo-Garcia, Luis Gómez-Chova, Anna Vaughan, Luis Guanter, and Andrew Markham. Semantic segmentation of methane plumes with hyperspectral machine learning models. *Scientific Reports*, 13(1):19999, November 2023.
- [6] Vít Růžička and Andrew Markham. Hyperspectralvlvits: General hyperspectral models for on-board remote sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2025.
- [7] Gabriele Meoni, Roberto Del Prete, Federico Serva, Alix De Beusscher, Olivier Colin, and Nicolas Longépé. Unlocking the use of raw multispectral earth observation imagery for onboard artificial intelligence. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17:12521–12537, 2024.
- [8] Robert Green. EMIT L1B At-Sensor Calibrated Radiance and Geolocation Data 60 m V001, 2022.
- [9] Robert Green, Andrew Thorpe, Philip Brodrick, Dana Chadwick, Amanda Lopez, Clayton Elder, Claire Villanueva-Weeks, Jay Fahlen, Red Willow Coleman, Daniel Jensen, Holly Bender, Quentin Vinckier, Chuchu Xiang, Winston Olson-Duvall, Sarah Lundeen, and David Thompson. EMIT L2B Estimated Methane Plume Complexes 60 m V001, 2023.
- [10] Maggie Chen, Hala Lambdouar, Luca Marini, Laura Martínez-Ferrer, Chris Bridges, and Giacomo Acciarini. UnorthodOS - Unorthorectified Dataset for On board Satellite methane detection, 2025.
- [11] CEOS - Committee o nEarth Observation Satellites. Greenhouse gas (ghg) emissions monitoring: Point source vs. wide area mappers, Aug 2024. Accessed: 2025-08-18.
- [12] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [13] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.