Monitoring Methane Emissions at Hassi Messaoud in 2021 and 2023 Using Sentinel-2 and Sentinel-5P: Assessing Algeria's Progress Toward its 2030 GHG Reduction Target

Nicholas Kinsella, supervised by Dr Robert McNabb.

Correspondence: Kinsella-N3@ulster.ac.uk

School of Geography and Environmental Sciences, Ulster University, Cromore Road, Coleraine, Co. Londonderry, BT52 1SA

Submitted for assessment : 7th of May 2025

Abstract. Methane (CH_4) is a potent greenhouse gas (GHG) with an outsized impact on near-term warming, and the oil and gas sector is the largest man-made contributor to post-industrial revolution anthropogenic CH_4 emissions. Accurate measurement and localisation of CH_4 emissions is crucial for mitigating the near-term effects of climate change, and for improving the efficiency of petrochemical facilities. A novel, open source, software tool - CH4 Sentinel - was developed to process Sentinel-2 data and quantify CH_4 super emitting point sources using a machine learning approach. CH4 Sentinel was applied to the Hassi Messaoud oil field in Algeria for the years 2021 and 2023, to determine if progress was being made to Algeria's 2030 GHG reduction goal. CH4 Sentinel processed Short Wave Infrared (SWIR) data from Sentinel-2's Multi Spectral Instrument (MSI) using a Multi-Band Multi-Pass method (MBMP) to visualise the plumes. Then a novel processor-light machine learning approach with an XGBoost regressor was employed to quantify super-emitting plumes in tonnes per hour (t/h). The regressor model showed an R^2 value of 0.81 and a RMSE of 1.57 t/h on unseen evaluation data. To provide context for plume measurements Sentinel-5P Tropospheric Monitoring Instrument (TROPOMI) data was processed using the Integrated Methane Inversion tool (IMI) for the years 2020 to 2023 to estimate total yearly emissions for the field. The IMI analysis revealed that while overall emissions at Hassi Messaoud are significantly underestimated by the Global Fuel Exploitation Inventory (GFEI), there has been a decrease since 2021 of 53,079 t/y. The Sentinel-2 analysis indicates that CH_4 super emissions have significantly reduced at Hassi Messaoud between 2021 and 2023 by an estimated as 22,636 tonnes per annum, approximately 42.6 % of the IMI reduction for that period.

1 Introduction

CH_4, a greenhouse gas with a warming potential 84 times greater than carbon dioxide (CO_2) over a 20-year period, has seen its atmospheric concentration increase over 250% since the industrial revolution (Pandey et al., 2023). Despite CH_4's short atmospheric lifespan when compared to CO_2 (Parker et al., 2020; Zhao et al., 2023), it has contributed to at least a quarter of anthropogenic warming (Varon et al., 2021; Vigano et al., 2008). Governments and industry are under increasing pressure to reduce CH_4 emissions: the 27th United Nations Climate Change Conference highlighted controlling CH_4 as the cheapest and fastest way to slow global warming in the years ahead (European Commission, 2022), and the 2021 Global Methane Pledge (GMP) was created to reduce anthropogenic emission levels by 30% of 2020 levels by 2030, with 157 countries participating (GMP, 2024; International Energy Agency, 2022; Malley et al., 2023). The oil and gas sector is of particular focus for efforts as it is the largest contributor to man-made CH_4 emissions, accounting for 35.76% of total anthropogenic CH_4 emissions (Jackson et al., 2020; Pandey et al., 2023).

Algeria ranks as the world's fourth-largest gas exporter, with an annual production of 108,667 million m^3 of natural gas in 2022 with that figure projected to increase in the coming years (International Trade Administration, 2024). It holds the title of Africa's leading natural gas producer and third-largest oil producer (Abada and Bouharkat, 2018). Additionally, Algeria is the second-largest natural gas supplier to Europe after Norway and the European Union has exerted increasing pressure on Algeria to boost its gas supply to satisfy the bloc's energy demands (Tahchi, 2024).

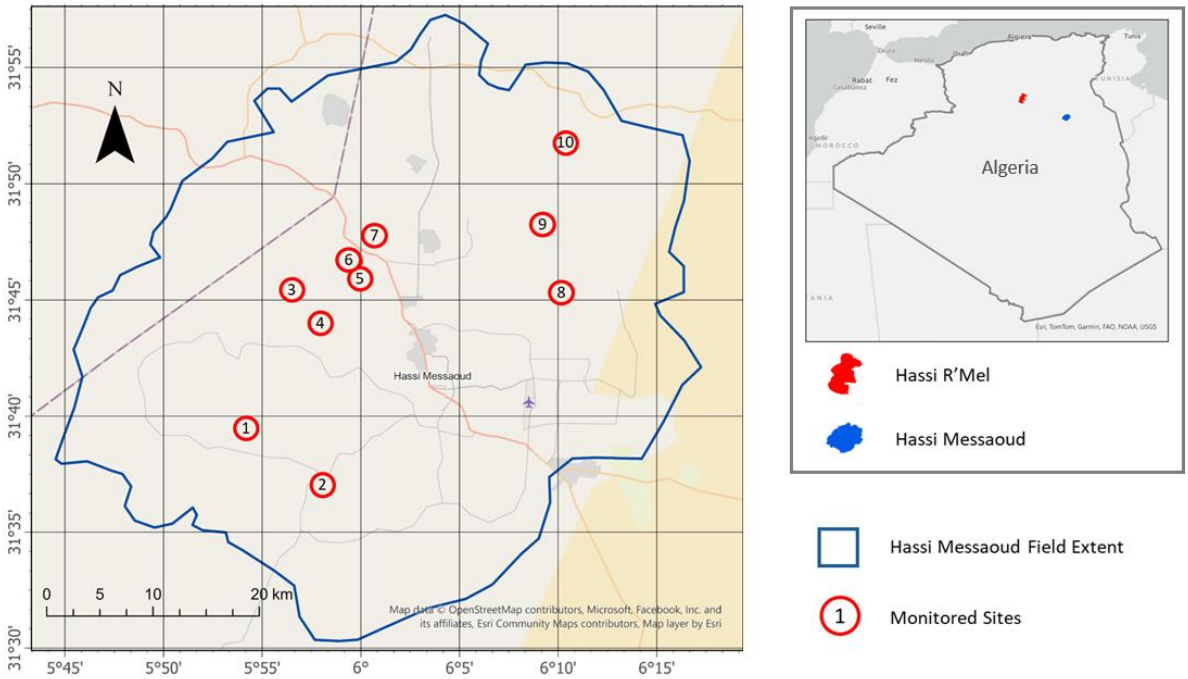
In 2015 Algeria made a pledge to reduce its GHG emissions by at least 7%, with a further 15% reduction subject to aid support. The reduction was scheduled to begin in 2021, culminating in 2030 (UNFCCC, 2015). Despite this, as of writing, Algeria has yet to join the GMP and currently does not have government led commitments to reduce CH_4 emissions specifically (International Energy Agency, 2024).

Algeria, like most countries, uses so called 'bottom-up' approaches to updating emission inventories (Government of Algeria, 2023; IPCC, 2006). These use sampling at a facility or component level, and component leakage models to estimate emissions. However research has indicated that bottom-up approaches are prone to underestimation of emissions (Dowd et al., 2023; Elkind et al., 2020; Karimi et al., 2021; Sherwin et al., 2024a; Vaughn et al., 2018; Wang et al., 2024a). Even developed countries can underestimate their emissions; Sherwin et al. (2024a) found actual oil and gas emissions to be around three times higher than US government inventory estimates. In Algeria's case, Naus et al. (2023) analysed the two largest petrochemical fields for the year 2020 and found that national inventories had overestimated emissions for a large gas field (Hassi R'Mel) by 53%, while an oil field (Hassi Messaoud) underestimated emissions by 79%.

The sensitivity of remote sensing instruments has dramatically improved in recent years, offering an opportunity to address many limitations of bottom-up approaches and produce more accurate CH_4 inventories (Elkind et al., 2020; Zhang et al., 2023). These 'top-down' methods use remote sensing via satellite, aircraft, or drone-based sensors. Weather permitting, they avoid the coverage issues associated with bottom-up point sampling, potentially capturing all emissions within an imaged area. However, since top-down approaches do not directly sample CH_4, it is crucial to ensure that detected emissions are genuine. Variations in surface albedo can create false emission artifacts (Varon et al., 2021), and gases such as water vapor and CO_2 can be mistaken for CH_4 by multispectral sensors, although this is less of an issue for hyperspectral instruments (Cusworth et al., 2020; Dowd et al., 2023; Maasakkers et al., 2022).

Environments conducive to accurate remote sensing of CH_4 typically exhibit minimal surface albedo variance and stable concentrations of water vapor and CO_2, making regions like the Algerian desert ideal for monitoring (Gorroño et al., 2023; Varon et al., 2021). Additionally, emissions from oil and gas operations are particularly suitable for top-down approaches due to their high concentration and intensity (Dogniaux et al., 2023; Varon et al., 2021). These parameters correspond well to Algeria's oil and gas sector making it an excellent candidate for a top-down analysis to evaluate the nation's progress towards its 2030 GHG reduction commitments.

Figure 1: Locations of sites monitored by this study in the Hassi Messaoud oil field with inset showing the locations of Hassi R'Mel and Hassi Messaoud fields in Algeria.



2 Hassi Messaoud oil field

Managed by state-owned company, Sonatrach, Algeria has well over a hundred identified oil and gas fields, two of which are considered giant (Abada et al., 2018, see Appendix 1). Situated in northern Algeria (Fig. 1), the Hassi R'Mel gas field covers an area of around 3,500 km^2, accounting for around 45% of Algeria's national gas production (Abada and Bouharkat, 2018; Naus et al., 2023; Rosenthal., 2023; Talamali., 2016). A further 350 km to the southeast lies the Hassi Messaoud oil field, which covers 1,600 km^2 and produces around 36% of Algeria's national oil production (Kamr Eddine Aissou., 2024; Naus et al., 2023). Despite being an oil field, Hassi Messaoud is by far the worse CH_4 emitter of the two, with TROPOMI detecting 154 super emissions since 2019 compared to just 8 in Hassi R'Mel over the same period (Kayrros, 2024). This is backed up by the study of Naus et al. (2023) which detected 1 super-emitter point source in Hassi R-Mel and 9 in Hassi Messaoud for the year 2020. The higher CH_4 emissions at Hassi Messaoud, despite being a smaller field than Hassi R'Mel, indicate the existence of opportunities to enhance operational efficiency, potentially benefiting both the environment and economic performance.

3 Objectives

Remote sensing surveys offer an objective way to assess Algeria's progress on their GHG reduction plan. Currently the only published remote sensing review of total CH_4 emissions from Hassi Messaoud by Naus et al. (2023) was conducted for 2020, before the 2021 start of Algeria's GHG reduction plan. Like Naus et al., the present study surveyed super-emitter events, defined as those detectable by Sentinel-2's MSI for the first year of planned reductions (2021) and the year 2023. To give these measurements context, overall emissions for Hassi Messaoud as measured by TROPOMI using the IMI, were measured yearly from 2020 to 2023.

In addition to this, the project aimed to produce open source software that could be used by Sonatrach and others to aid in their emission reduction goals in 2025 and beyond.

4 Platform Selection

Multispectral and hyperspectral sensors have been extensively used to monitor CH_4 point sources as well as overall atmospheric CH_4 levels (Dowd et al., 2023; Naus et al., 2023; Maasakkers et al., 2022; Sherwin et al., 2023; Varon et al., 2021). This study and its software used Sentinel-2 with Sentinel-5P. The following section will explain why these platforms were chosen with reference to other CH_4 detection capable satellites.

Remote sensing platforms detect CH_4 by capturing reflected light from the specific CH_4 absorption wavelengths that are found in the SWIR (Fig. 2). The platforms can be broadly categorised into area flux mappers (AFM) - designed to observe total emissions at the regional or global level, and point source imagers (PSI) - high spatial resolution sensors that can measure individual plumes (Jacob et al., 2022).

4.1 Area Flux Mappers

AFMs cannot detect most super-emitter plumes, although incorporating their data into the final analysis will provide total emissions for the year, providing valuable context for what proportion of detected super emitter plumes contribute to overall CH_4 emissions.

Sentinel-5P, launched in 2017, with TROPOMI for monitoring atmospheric pollutants, including CH_4, provides near daily data globally with a coarse resolution of 7.5 x 5.5 km. It retrieves CH_4 using a SWIR band of 2305-2385nm and a near infrared band of 757-774 nm (Copernicus, 2023; Ehret et al., 2022; Naus et al., 2023; Maasakkers et al., 2022; Siddiqui et al., 2024). It has also been shown to be able to detect plumes >5 t/h (Schuit et al., 2023).

With a higher spectral resolution, the two satellite Sentinel-3 constellation is an AFM that uses the multispectral Sea and Land Surface Temperature Radiometer (SLSTR) with a 500m spatial resolution and a return rate of 1-2 days at the equator. It has three SWIR bands, two of which intersect with CH_4 absorption lines (1580 - 1640 nm and 2230 – 2280 nm). SLSTR has a recorded CH_4 detection threshold of 8-20 t/h (Pandey et al., 2023).

Between these two, Sentinel-5P makes a good choice for measuring total emissions over large areas such as the Hassi Messaoud oil field, which is unsurprising as it was purpose built for this task (Reshi et al., 2024). For CH_4 plumes, it has a better detection threshold than Sentinel-3 and may separately be used for confirming plume measurements for very large emission events.

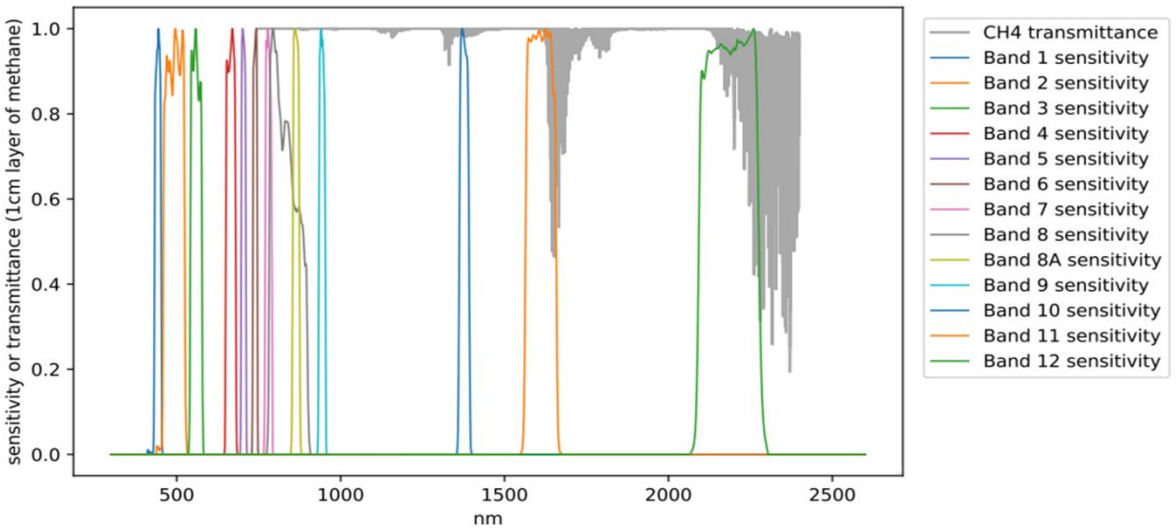


Figure 2: CH_4 transmittance overlayed with spectral sensitivity for Sentinel-2A bands (Ehret et al., 2021).

4.2 Hyperspectral sensors

In recent years several satellite based, high spatial resolution, CH_4 detection capable, hyperspectral sensors have come into operation. These include Carbon Mapper, GHGSat, PRISMA, EnMAP, EMIT and others (Jacob et al., 2022; Sherwin et al., 2024b). These systems have spatial resolutions of ~30 m or better (Guanter et al., 2021; Irakulis-Loitxate et al., 2022; Maasakkers et al., 2022; Roger et al., 2024), except for EMIT which is 60 m (Thorpe et al., 2023), giving them the high-resolution ideal for this project. They have a high number of bands, with as few as 174 in the case of PRISMA (Guanter et al., 2021) to as many as 421 in the case of Carbon Mapper (Zandbergen et al., 2022), which makes them effective at reducing emission artefacts and achieving lower detection thresholds. The highest detection threshold being EnMAP at ~1100 kg h^{-1} (Sherwin et al., 2024b) and the lowest being GHGSat at 42 kg h^{-1} (Dowd et al., 2023). However, where many of these systems become inappropriate for this project, is their need for tasking, as in the case of GHGSat, PRISMA and EnMAP. This means that a systematic survey of Hassi Messaoud is not likely, given scheduling demands (Sherwin et al., 2023). EMIT by contrast utilises a systematic approach, but has only been operational since 2022 and has a return rate over Hassi Messaoud of 68 days making it unsuitable (NASA Jet Propulsion Laboratory, 2024). Carbon Mapper provides a similar systematic survey approach and much more frequent revisit times of ~3.5 days, although the first satellites of the constellation only became operational in 2023, meaning that its data is not available for the 2021 analysis (Duren et al. 2020; Zandbergen et al., 2022).

4.3 Multispectral sensors

Multispectral sensors offer lower spectral resolution but have been able to detect CH_4 emissions in favourable conditions (Jacob et al., 2022). WorldView3 has a non-sun synchronous return rate of ~1 day, which would cause problems with emission artefacts, however it has 8 SWIR bands which reduce this issue, and a spatial resolution of 3.7-4.1 m, with an emission detection threshold of 200–300 kg h^{-1} (Irakulis-Loitxate et al., 2022; Sánchez-García et al., 2021). Unfortunately, like many of the hyperspectral satellite missions, WorldView3 needs to be tasked to a particular target and so is not suitable for this project (Sherwin et al., 2023).

Unlike most of the PSIs examined so far, Landsat 8/9 and Sentinel-2 operate by systematically surveying the Earth. This surveying method means that the missions make regular overpasses of Hassi Messaoud, providing the coverage needed for this project.

The Sentinel-2 satellite constellation, consisting of a pair of satellites, uses a push-broom spectrometer to provide systematic, sun-synchronous, global coverage every 2-5 days. Equipped with the 13-band Multispectral Instrument, including two SWIR bands at a 20m spatial resolution: at ~1560–1660 nm (SWIR-1) and at ~2090–2290 nm (SWIR-2). It can detect CH_4 emissions as low as 1.4 – 2.0 tons per hour in areas with low albedo variance, such as deserts like Hassi Messaoud (Gorroño et al,. 2023; Naus et al., 2023; Varon et al., 2021; Wang et al., 2024a). MSI has been noted to have more signal noise compared to other PSIs (Sherwin et al., 2024b). All MSI data is freely available

and accessible with the OpenEO API (Musial et al., 2024).

As with Sentinel-2, Landsat 8 and 9 are a systematic surveying, sun synchronous mission and they have a pair of SWIR bands of 1565–~1651 nm and 2106–~2294 nm (Masek et al., 2020) that are nearly identical spectrally to Sentinel-2's MSI. They also have a similar CH_4 detection threshold of 1.39 $t\,h^{-1}$ (Sherwin et al., 2024b). However, their lower spatial resolution of 30 m and 16 to 8 day return rate for 2021 and 2023 respectively, mean that that mission is less suitable for this study (Kabir et al., 2023; Masek et al., 2020). Landsat 8 and 9 data is available through the Earth Access API (Lopez et al., 2022).

MSI was used by Varon et al. (2021) to detect CH_4 emissions. The researchers were able to successfully retrieve the CH_4 plumes, as well as estimate their vertical column concentrations in $mol\,m^{-2}$. Another analysis by Dogniaux et al. (2023) of the Nord-Stream Pipeline attack in the Baltic also used Sentinel-2 and Landsat 8 imagery, measuring the gas leak to possibly be the largest recorded anthropogenic CH_4 emission in history.

Based on the available options Sentinel-2 emerges as the most suitable platform for CH_4 point detection, because of its global systematic coverage, regular revisit rates, high spatial resolution and pair of SWIR bands that coincide with CH_4's transmittance window. It has also been effectively utilised in multiple studies for detecting large CH_4 emissions (Dogniaux et al., 2023; Naus et al., 2023; Varon et al., 2021.).

5. Methods

The project's methodology is split into two workflows: one detecting and measuring super-emitting plumes with the PSI software using MSI data, and the other measuring overall emissions using TROPOMI data via a tool called Integrated Methane Inversion (IMI).

5.1 MSI SWIR Band Processing

The method for processing the MSI SWIR data was taken from Varon et al. (2021) who outlined three connected approaches for detecting CH_4 emissions using MSI data: The first, Single-Band Multi-Pass, utilised the Short-Wave Infrared band 12 (SWIR-2) on a day with a known emission, and then compared it to a day when there was no emission. Second was a Multi-Band Single-Pass method (MBSP) which took the Short-Wave Infrared band 11 (SWIR-1) and SWIR-2 readings for the same day, and extracted the CH_4 plume by least squares fitting the two band scenes, and then differencing the band datasets. The final method was a Multi-Band Multi-Pass method (MBMP), which effectively combines the two previous approaches by differencing SWIR-1 and SWIR-2 for a day with a known emission (without least squares fitting) and then again for a day with the emission absent. The MBMP method produced the best results with the least emission artefacts and so was chosen for this study.

The processing steps are summarised in Fig. 3. The Copernicus OpenEO API was used to download the Sentinel-2 L1C top-of-atmosphere product. This was then divided by 10,000 to obtain the reflectance values

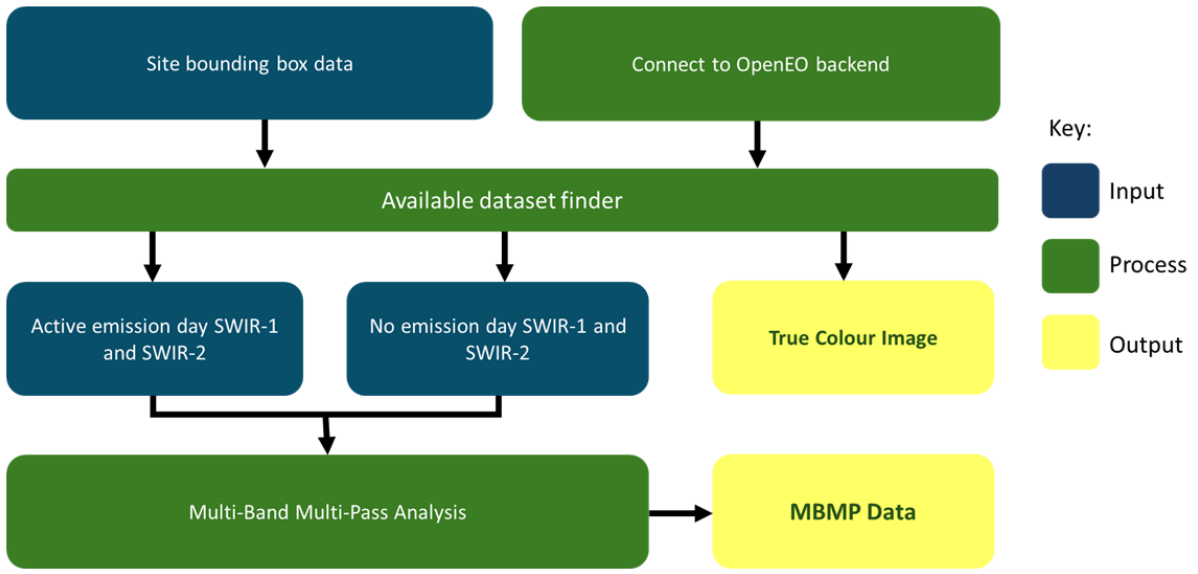


Figure 3: SWIR Band Processing Workflow

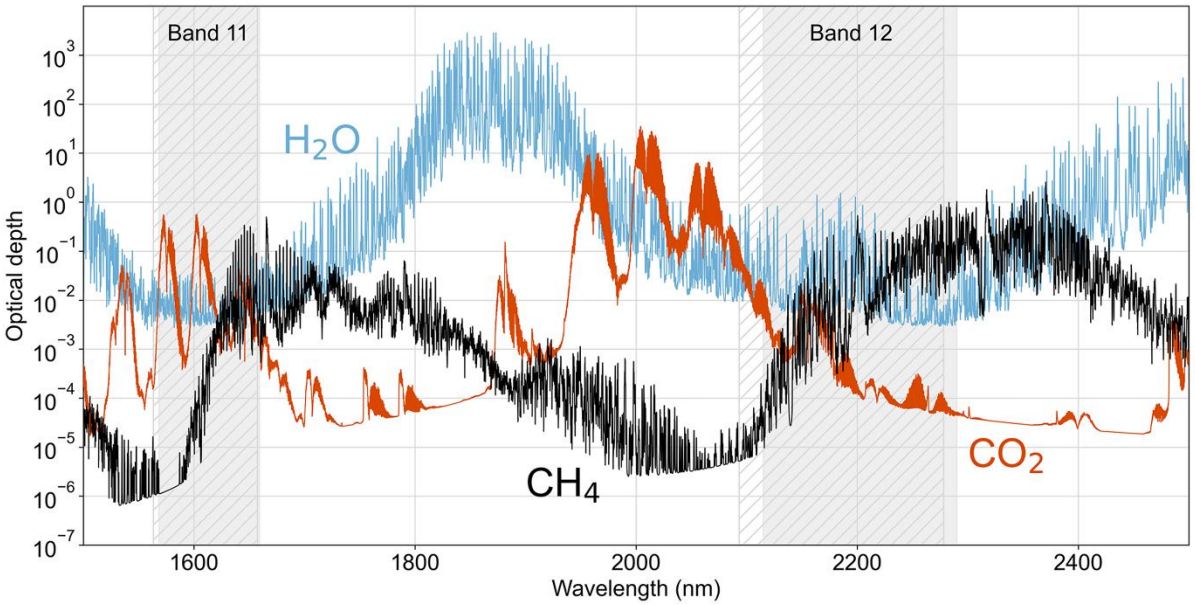


Figure 4: Absorption levels of H_2O, CO_2 and CH_4 as measured by optical depth in a slanted column of 40 degrees solar zenith and a satellite viewing angle of nadir (Varon et al., 2021).

for Short-Wave Infrared (SWIR) bands 11 (1565–1651 nm) and 12 (2106–2294 nm) (Copernicus, 2025; Masek et al., 2020). Visible spectrum bands (2, 3, and 4) were downloaded to aid in emission validation.

Scenes with ≤5% cloud cover were chosen for both the non-emission scene and active-emission scenes. No emission and active emission scenes were processed using the multi-band-single-pass equation as outlined in Varon et al. (2021):

$$MBSP = \frac{B12 - B11}{B11} \quad (1)$$

Where: $B12$ is the MSI SWIR-2 band reflectance and $B11$ is the MSI SWIR-1 band reflectance. Once the MBSP raster had been calculated, the following equation was then used to calculate the multi-band-multi-pass raster:

$$MBMP = ActiveMBSP - NoMBSP \quad (2)$$

Where $ActiveMBSP$ is the MBSP raster for the active emission scene and $NoMBSP$ is the MBSP raster for the no emission scene. The no emission scene was taken to be one satellite pass before the active emission scene, unless there were obscuring features interfering with the point source's visibility.

5.2 Plume Identification

This study used manual identification of valid plume features. To aid this, ten locations from peer-reviewed sources were identified as likely plume locations and used as the focus of the analysis (Table 1, Fig.1).

Table 1: Monitored sites at Hassi Messaoud.

Site	Location	Source
1	31.658, 5.905	Varon et al., 2021
2	31.617, 5.967	Naus et al. 2023
3	31.756, 5.942	Naus et al. 2023
4	31.734, 5.967	Naus et al. 2023
5	31.769, 6.001	Sanchez-Garcia et al., 2021
6	31.778, 5.995	Sanchez-Garcia et al., 2021
7	31.798, 6.011	Sanchez-Garcia et al., 2021
8	31.757, 6.168	Naus et al. 2023
9	31.807, 6.154	Gorroño et al., 2023
10	31.865, 6.174	Pandey et al., 2023

Different gas emission types exist at Hassi Messaoud and determining which is a genuine CH_4 emission was critical to the project. H_2O, CO_2 and CH_4's absorption wavelengths all coincide in MSI's SWIR bands, which could cause false positives, known as emission artefacts (Varon et al., 2021) (Fig.4). Care must therefore be taken in understanding what constitutes a genuine CH_4 emission. The following rationale was used to determine if a plume was a valid emission (Fig.5).

Not a CH_4 plume (Fig. 5a): If a plume is visible in the true-colour scene, it cannot be a CH_4 emission as the gas does not absorb light in the visual spectrum (Ehret et al., 2021; Varon et al., 2021).

Diffraction Spike (Fig. 5b): When a flare is lit on a relatively small emission, a four-pointed star like

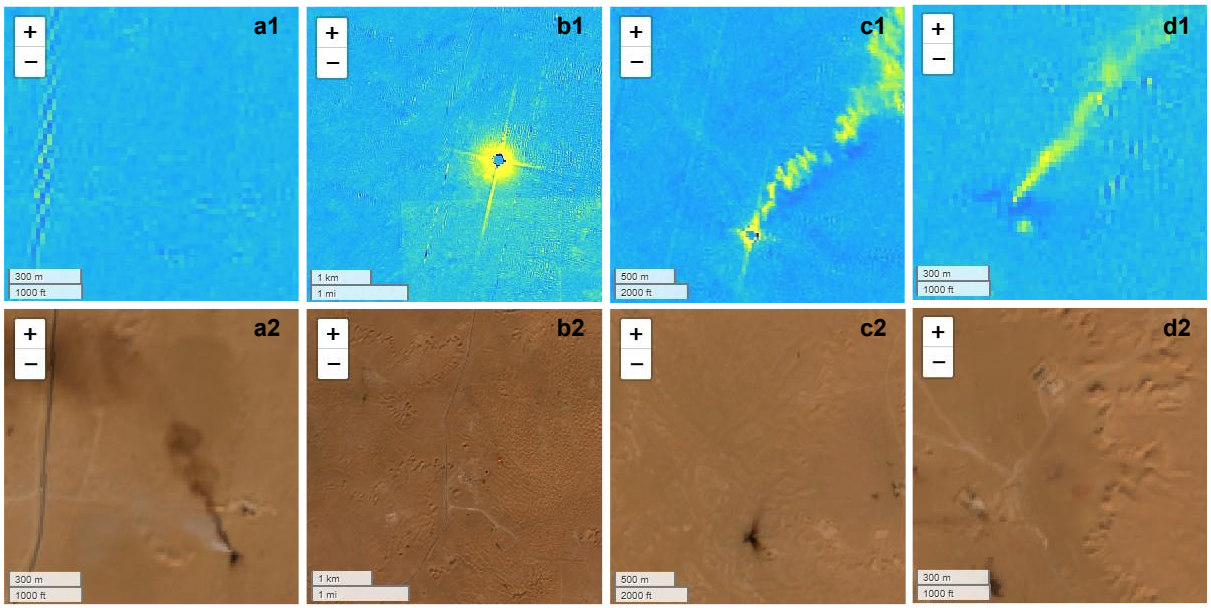


Figure 5: Emission examples from CH4 Sentinel. The top row shows the MBMP data and the bottom row shows the same location in true-colour. a = Not a CH_4 plume, b = Diffraction spike, c = Diffraction spike with plume and d = Emission with no diffraction spike.

diffraction spike will be visible at the plume's source, which is a lensing effect caused by the brightness of the flare (Massimetti et al., 2020). The resultant gasses produced will be CO_2 and H_2O (Musial et al., 2024) and although those gasses are likely present (Fig. 4), they are not at a sufficient threshold to be visible.

Diffraction Spike with plume (Fig. 5c): This scenario is like Fig. 5b but includes a plume because the amount of gas has reached MSI's detection threshold. As previously mentioned, the gases present in the plume will primarily be CO_2 and H_2O (Massimetti et al., 2020), with trace amounts of CH_4, as flaring is ~95% effective at converting CH_4 (Caulton et al., 2014).

Emission with no diffraction spike (Fig. 5d): When the CH_4 emission is emitted into the atmosphere with no flaring of the emission: This appears as a plume like signal in the MBMP image, but will be invisible in the true-colour image (Varon et al. 2021). There will be no star shaped diffraction spike. This represents a genuine CH_4 emission.

5.3 Emission Measurement

To quantify CH_4 emission rates (Q), MBMP reflectance data is ordinarily converted into scientific units, such as parts per billion (ppb) or moles per square meter (mol m^{-2}). Traditionally, this conversion relies on processor heavy, physics-based models (Jiang et al., 2024; Varon et al., 2018; Varon et al., 2021). Very recently, lightweight algorithmic approaches like Data Driven Emission Quantification (DDEQ) have emerged as an alternative (Kuhlmann et al., 2024). Both vectors of approach presented challenges for this project: DDEQ had data format compatibility issues, while the physics-based models demanded high processing power and rely on propriety software restricting their use.

The proprietary software in question is a Radiative Transfer Model (RTM) Python code owned by GHGSat (Varon et al., 2021). Although free for research purposes on request, it cannot be published online to download. Additionally, when used on an area the size of Hassi Messaoud with a typical desktop computer, it can take hours to run. To maintain an open and accessible project and minimise processing time, it was decided to avoid directly using the RTM in the project's software. Instead, the strategy involved using the Integrated Mass Enhancement (IME) method as outlined in Varon et al. (2018) and Varon et al. (2021), including the RTM to measure emission rates (Q_{IME}), using small bounding boxes to reduce processing times. These measurements then served as the dependent variables in training data for a regression model that can estimate the emission rate of user defined plumes (Q_M). The full workflow is summarised in Fig. 6 and explained in the following sections.

5.4 Q_{IME} and Training data.

The MBSP raster values were converted into concentration values using the same clear-sky RTM used by Varon et al. (2021). The RTM uses the High-

Resolution Transmission Molecular Absorption Database (HITRAN) along with Sun-satellite viewing geometry to simulate how much CH_4 is needed to produce an observed dimming in reflected sunlight, giving its results in mol/m^{-2} (Gordon et al., 2022). These values were converted to kilograms per square meter (kg/m²) by multiplying by CH_4's molar mass of 0.016 kg/mol (National Institute of Standards and Technology, 2023). These were then processed into a MBMP raster as earlier in equation 2.

CH_4 concentrations were thresholded using a binary mask to identify pixels with values ≥ the 95th percentile of the scene. These were designated as potential plume pixels. The manually tagged plume locations were then used to identify the actual plumes from non-plume high pixel value features (such as circular crop irrigation systems) and only accepting a tagged cluster of ≥50 pixels as a valid plume tag, to mitigate the issue of data noise inherent in the Sentinel-2 L1C data (Sherwin et al., 2024b). Convex hulls were then generated to encompass all adjacent plume pixels, segmenting them from the data for measurement.

The IME was measured by summing the CH_4 values in the plume area as expressed in Varon, et al. (2018):

$$IME = \sum_{j=1}^{N} \Delta\Omega_j A_j, \quad (3)$$

Where $_j$ is a single pixel, N is the total number of pixels in a plume, A_j is pixel area and $\Delta\Omega_j$ is the CH_4 concentration of that pixel.

The relationship between IME and emission rate (Q_{IME}) is defined as the mass of methane (IME) divided by the residence time of the gas (τ) in the detectable plume area, expressed as:

$$Q_{IME} = \frac{IME}{\tau} \quad (4)$$

τ can then be expressed as:

$$\tau = \frac{L}{W} \quad (5)$$

Where L is plume length and W is wind speed, meaning IME's relationship to Q_{IME} can be expressed as:

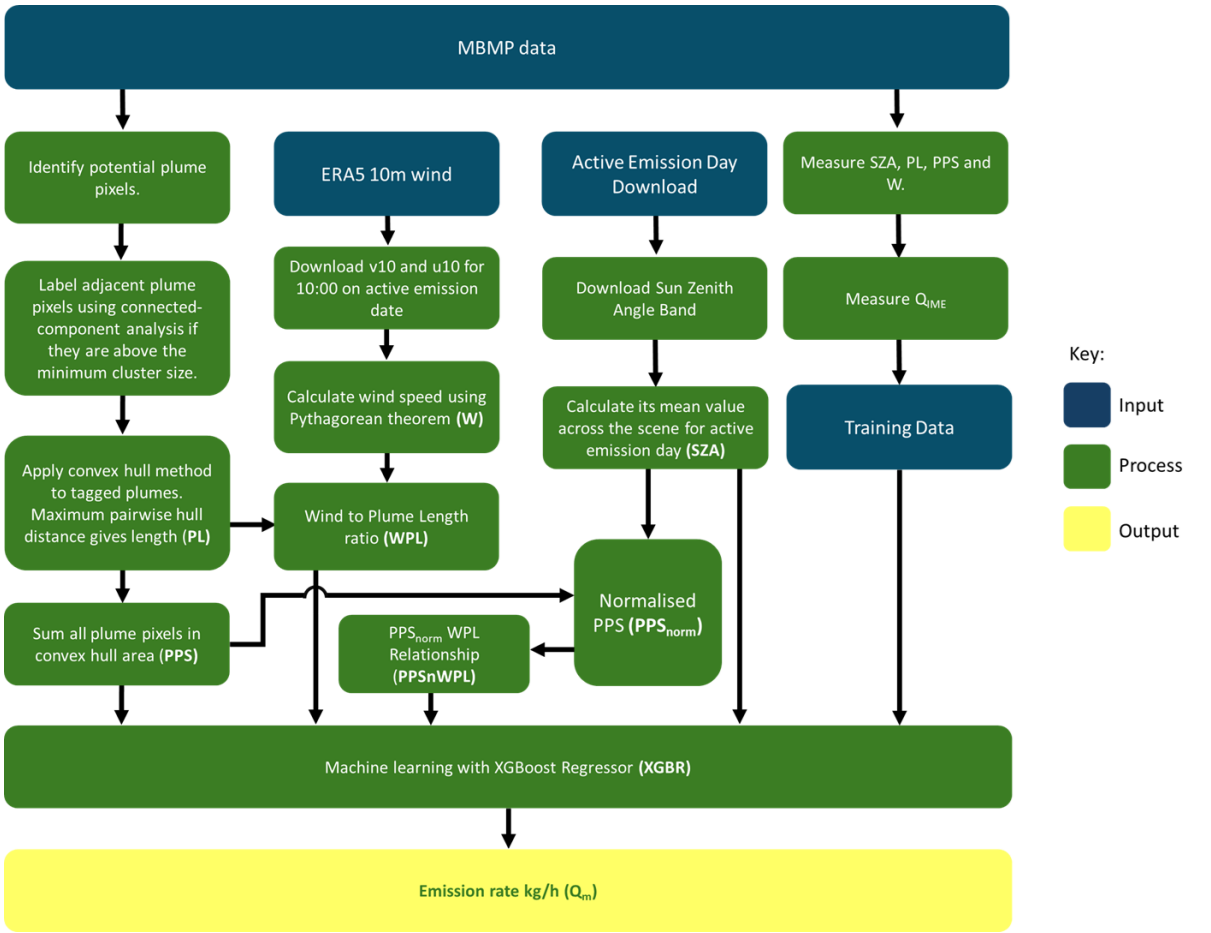


Figure 6: Workflow for super-emitting plume analysis.

$$Q_{IME} = \frac{W}{L} \times \text{IME} \quad (6)$$

The method in this analysis differs from Varon et al. (2018) as they used large eddy simulations to simulate the turbulent diffusion of the plume instead of assuming uniform movement with wind, as is the case in this study. This simpler method was also used in Irakulis-Loitxate et al. (2022).

Before collecting sample data, the Q_{IME} measurements were validated against ten Sentinel-2 plume measurements from peer reviewed sources in Hassi Messaoud that also used an IME method (Gorroño et al., 2023; Naus et al. 2023; Pandey et al., 2023; Varon et al., 2021) (Appendix 2). Thereafter, one hundred plumes were measured for Q_{IME} during 2019 and 2020, with seventy used to train the model and thirty randomly selected emissions to test model performance.

Q_{IME} served as the dependent variable for the predictive model and each of the 100 plumes were also measured for the following four variables:

1. Plume Pixel Sum (PPS): The sum of the MBMP pixel values in the plume's convex hull. Even without being converted using the RTM, a higher PPS suggests a larger methane mass, thereby increasing emission rate.

2. Wind Speed (W): Wind Speed affects how quickly a plume disperses and therefore reduces the observed PPS. Wind speed values are obtained from the Climate Data Store API ERA5 10 m dataset. The API gives wind speed data at an altitude of 10 m and was chosen owing to its accuracy in flat inland areas such as Hassi Messaoud (Gualtieri, 2022; Molina et al., 2021). ERA5 10 m gives wind speed as two components u10 (east-west) and v10 (north-south) with negative values representing wind moving west and south respectively and vice versa. The components form two sides of a right-angled triangle meaning the wind speed can be derived using the Pythagorean theorem, expressed as:

$$W = \sqrt{u10^2 + v10^2} \quad (7)$$

10:00 am on the active observation date was chosen as the time in all instances because the overpass time for Sentinel-2 is between 10:25 and 10:35 in the study area.

3. Plume Length (PL): The length of the plume is related to both PPS and W Calculated as the maximum pairwise distance between points on the convex hull.

4. Sun Zenith Angle (SZA): The angle between the sun and vertical, indicating how directly the sun is hitting the Earth's surface. Lower SZA values result in more illumination, affecting measured PPS. The value was calculated as the mean pixel value of SZA for the scene.

5.5 Feature Engineering

Some of the variables selected are not independent of one another. To capture relationships from the IME method, the following three features were derived.

Plume Pixel Sum normalised (PPS_{norm}): This is to account for scene brightness changes owing to the changing angle of the sun throughout the year. Solar radiance is highest when SZA is lowest at 0° and radiance decreases as SZA increases. Cronin (2014) suggested that reflectance can be partly normalised by using a cosine (cos) relationship of SZA. This is expressed as:

$$PPS_{norm} = \frac{PPS}{\cos(SZA)} \quad (8)$$

Wind to Plume Length ratio (WPL): Because the relationship between emission rate and IME is driven by residence time (Varon et al., 2018), a feature to capture this was created by dividing W by L as in equation 5 of the IME method. This is expressed as:

$$WPL = \frac{W}{PL} \quad (9)$$

PPS_{norm} WPL Relationship (PWPL): This was created to approximate the relationship described in equation 6 in the IME method. It is expressed as:

$$PWPL = PPS_{norm} \times \frac{W}{PL} \quad (10)$$

The final features used in the model were PPS, SZA, PPS_{norm}, WPL and PWPL.

5.6 Regression Model Choice

Initially a simple linear regression model was implemented to predict Q_M, however it produced nonsensical negative values for emission rates in some circumstances. Random Forest and Extreme Gradient Boost Regression (XGBR) models were explored as an alternative and these were then trained using the seventy training plumes. These were then tested against the thirty unseen evaluation emissions. XGBR was chosen as the regression algorithm after it emerged as the better predictor with a higher coefficient of determination score (R^2) and lower Root Mean Squared

Error (RMSE) in the evaluation and training data. XGBR uses decision trees much like a random forest, but instead of running all the trees in parallel, it runs them in series, with each tree attempting to improve on the previous tree's prediction (Chen & Guestrin, 2016).

5.7 XGBR Configuration

This section outlines the configuration choices made to optimise the performance of the XGBR model allowing it to handle outliers, skews and variable scaling effectively.

Objective: regression with squared loss: When training the model the objective is to have the closest possible match between Q_{IME} and Q_M. However, relatively small errors are of less concern than large ones. The squared error objective squares the difference between the Q_{IME} and Q_M so that larger errors in predicted emissions are more heavily penalised than smaller errors (Ciampiconi et al., 2023; XGBoosting, 2024a).

Log Transformation of Q_{IME}: The model training data is right-skewed and approximately log-normal - typical of point-source emission data (Frankenberg et al., 2016). In such cases, a small number of large values dominate the distribution and models trained using squared loss are especially sensitive to these outliers (Nielsen, 2016). Applying a log transformation reduces the influence of the large outliers by compressing extreme values, and producing a distribution closer to normal (Kuhn & Johnson, 2013), the effect of this is that the model will focus on minimising relative errors instead of absolute errors.

Standard Scaler: The independent variables in the model span different numerical ranges. Without scaling, features with larger magnitudes would dominate the training process. The standard scaler zeros all variable means to 0 and scales their standard deviation to 1, ensuring all variables contribute comparably to the model's predictions (Kuhn & Johnson, 2013; XGBoosting, 2024b)

5.8 XGBR Hyperparameters

Hyperparameters were optimised using the Optuna python library (Akiba et al., 2019). The seventy training plumes guided the optimisation and each hyperparameter iteration was scored by its R^2 on the evaluation set. Hyperparameters were chosen based on predictive strength and similarity of R^2 and RMSE scores between training and evaluation data (table 2).

Table 2: Hyperparameter configuration.

Parameter	Value
Estimators	156
Learning rate:	0.0226
Max depth:	9
Colsample bytre:	0.783
Subsample:	0.637
Gamma**:**	1.11e-05
reg_alpha:	0.221
reg_lambda:	0.289
min_child_weight:	6

5.9. Area Flux Mapper Analysis

Because smaller emissions are more common than larger ones (Frankenberg et al., 2016), it stands to reason that the CH_4 detection threshold of MSI is not picking up a significant number of emissions that fall below its threshold (Gorroño et al., 2023; Varon et al., 2021; Wang et al., 2024a) The study by Naus et al. (2023) indicated diffuse emissions contributed a similar amount to overall CH_4 emissions as the super emitters detectable by MSI. Therefore, to properly assess CH_4 emissions at Hassi Messaoud, these diffuse emissions must be quantified.

5.10 Area Flux Mapper Methodology

The Sentinel-5P TROPOMI CH_4 product provides atmospheric CH_4 concentrations in parts per billion (ppb), and its data has been shown to be in close agreement with ground-based measurements from the Total Carbon Column Observing Network (Liu et al., 2021; Lorente et al., 2021). Additional steps are required to estimate ground level CH_4 emissions from these observed whole atmosphere column concentrations, because CH_4 can travel significant distances from its source and may be present in high altitudes over otherwise low-emission areas, and vice versa, something not distinguished in the TROPOMI data product (Varon et al., 2022).

The IMI is designed to provide regional scale CH_4 emission data based on TROPOMI and a chemical transport model named GEOS-Chem. It was developed by Varon et al. (2022) and natively runs on Amazon Web Services cloud computing and storage (AWS). The IMI works by comparing the bottom-up inventories to top-down measurements by TROPOMI, thereby providing a more accurate measure of emissions than bottom-up data. In this study, the IMI was set-up on Ulster University's high performance

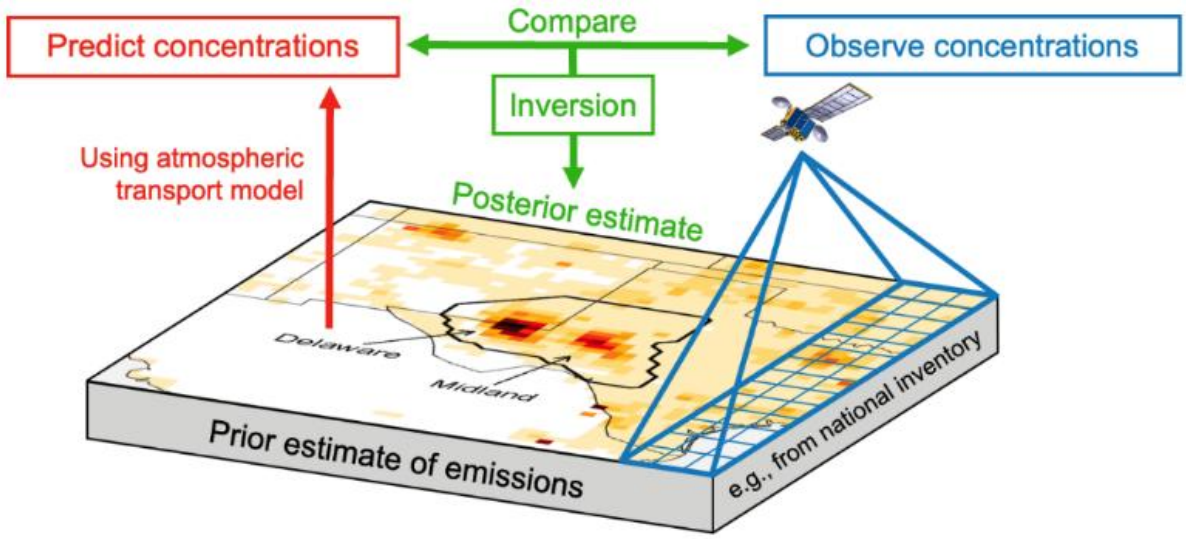


Figure 7: Processing steps and data sources for the IMI (Integrated Methane Inversion team, 2024)

computing cluster 'Atlas' by the project's supervisor Dr McNabb to avoid the significant costs associated with inversions using AWS.

Figure 7 shows the main processing steps. The inversion process starts with a "prior estimate of CH_4 emissions" based on several bottom-up inventories. For fuel exploitation emissions like at Hassi Messaoud, it uses GFEI v2.0 (Estrada et al., 2024). Like all gasses, CH_4 is affected by weather and other atmospheric influences, so the GEOS-Chem chemical transport model is applied to the prior estimate data, to create a model of predicted atmospheric concentrations, that should be seen by Sentenel-5P. The prediction is then compared to Sentenel-5P's actual observations. The 'inversion' involves adjusting the predicted concentrations so that they better match the observed concentrations. This results in a "posterior estimate" which should be a better reflection of reality.

Inversions were performed for whole calendar years 2020 to 2023 using a bounding box of (South-West) 30.72952° N, 5.027957° E and (North-East): 32.72952° N, 7.027957° E. 2020 was included to validate the total results of Naus et al. (2023) and 2022 to provide a continuity of four sequential years. A user manual detailing the IMI's configuration and use on AWS for non-experts was written and included in this project's Github repository linked at the end of this paper.

6. XGBR Model Performance

The IME tool used to collect the training data for the XGBR was compared to ten known emissions from academic literature. Eight of the emissions were predicted within the margin of uncertainty stated in their research paper. Taken together these eight had an R^2 of 0.81 and an RSME of 1.57 t/h. Of the two remaining, one plume with a study estimate of 2.1 t/h could not be measured due to being too faint. The other plume with a study estimate of 2.7 t/h returned a Q_{IME} value of 0.38 t/h which is far below the detection threshold of MSI (Gorroño et al,. 2023), indicating it is erroneous (fig.8).

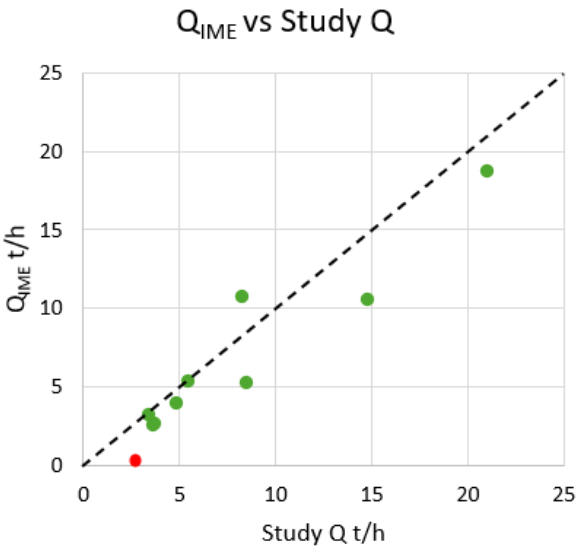


Figure 8: The dashed black line represents a perfect prediction. Erroneous outlier marked in red.

Once the 100 plume samples were collected, 70 randomly selected samples were used to train and 30 were used to evaluate the XGBR model's performance. Figure 9 presents how closely the Q_M values of XGBR matched the those measured by the IME.

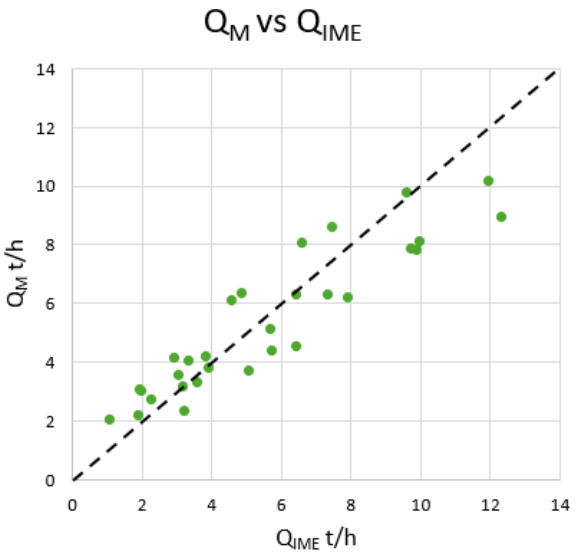


Figure 9: Comparison of Q_{IME} and Q_M. The dashed black line represents a perfect prediction.

The model has a training R^2 of 0.8547 with a RMSE of 1.30 t/h. The unseen evaluation dataset indicated an R^2 of 0.8113 and a RMSE of 1.34 t/h (Fig. 9)

A feature importance analysis indicates that PWPL contributes to nearly half of the overall prediction, followed by smaller contributions by PPS, PPS_{norm}, and WPL, with SZA contributing the least (Fig. 10).

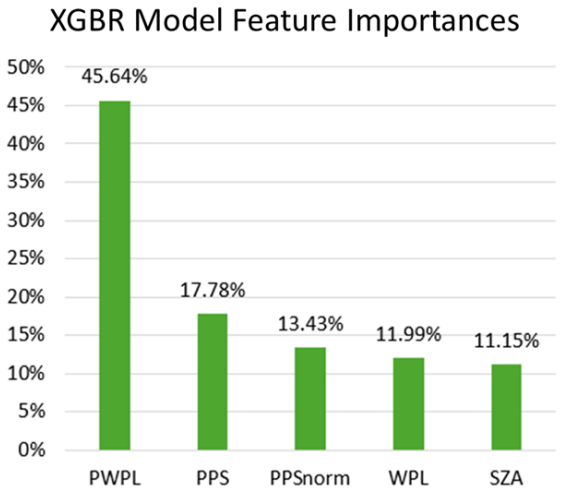


Figure 10: Percentage contribution of features to XGBoost Regressor.

7. Results

7.1 CH4 Sentinel Results for 2021

Of 145 Sentinel-2 overpasses, 97 had the less than 5% cloud cover suitable for the 2021 super-emitter analysis. Plume activity was detected at sites 4, 5, 6, 7 and 9 with sites 5 and 6 being the most active. The total estimated emissions for 2021 for each site are shown in Table 3 and visualised in Fig. 11.

Table 3: Result summary for 2021 with Annual Emission Estimates by Site.

Site	Detections	Year %	Estimated days active	Average (t/h)	Estimate for year (t/y)
4	1/97	1.03	4	9.63	924
5	40/97	41.24	151	4.26	15,440
6	44/97	42.68	166	3.95	15,743
7	9/97	9.28	34	3.67	2,995
9	6/97	6.18	23	4.02	2,222
				Total:	37,324

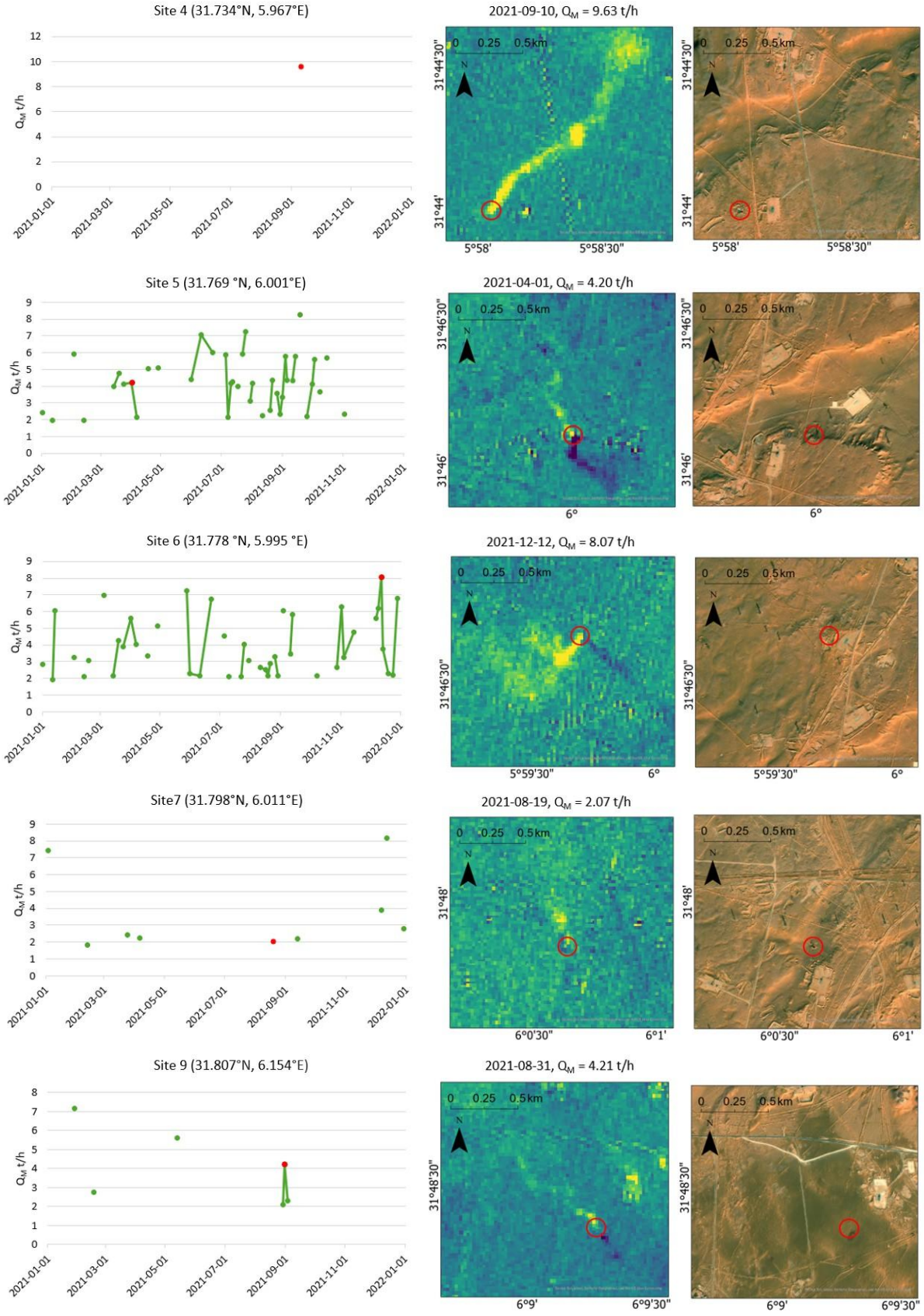


Figure 11: 2021 results with timeseries, MBMP and True-colour panels. Consecutive emissions are marked with connecting lines in the timeseries and the emission displayed in the panel is marked in red.

7.2 CH4 Sentinel Results for 2023

Of 147 Sentinel-2 overpasses, 112 had the less than 5% cloud cover suitable for the 2023 super-emitter analysis. Plume activity was detected exclusively at site 6. The total estimated emissions for 2023 for each site are shown in table 4 and visualised in Fig. 12.

Table 4: Result summary for 2023 with Annual Emission Estimates by Site.

Site	Detections	Year %	Estimated days active	Average (t/h)	Estimate for year (t/y)
6	47/112	41.96	153	4.00	14,688
				Total:	14,688

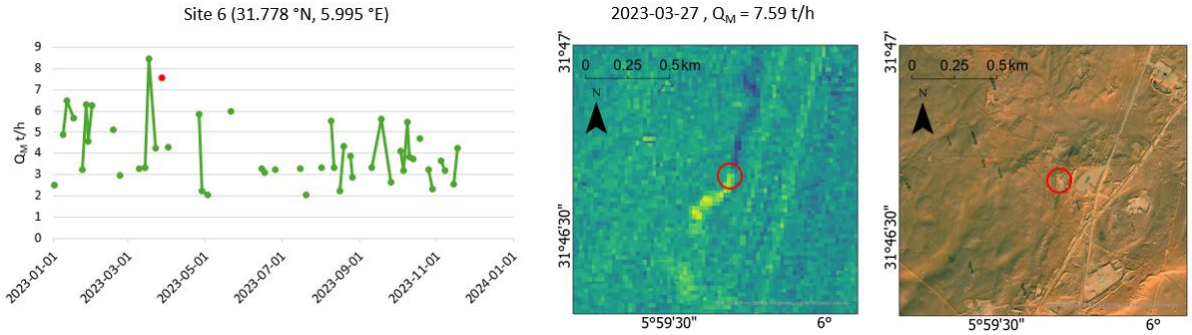


Figure 12: 2023 results with timeseries, MBMP and True-colour panels. Consecutive emissions are marked with connecting lines in the timeseries and the emission displayed in the panel is marked in red.

7.3 IMI Analysis 2020 – 2023

Posterior emission measurements based on the IMI indicate that the field has reduced overall emissions by 94,195 t/y since 2020 (31.82%) or 53,079 t/y since 2021 (20.83%). The GFEI based prior emissions are nearly static for all 4 years (Table 5, Fig. 13).

Table 5: Result summary for IMI analysis.

Year	**2020**	**2021**	**2022**	**2023**
Prior emissions	90,325	91,337	92,134	92,131
Posterior emissions	296,032	254,916	200,575	201,837

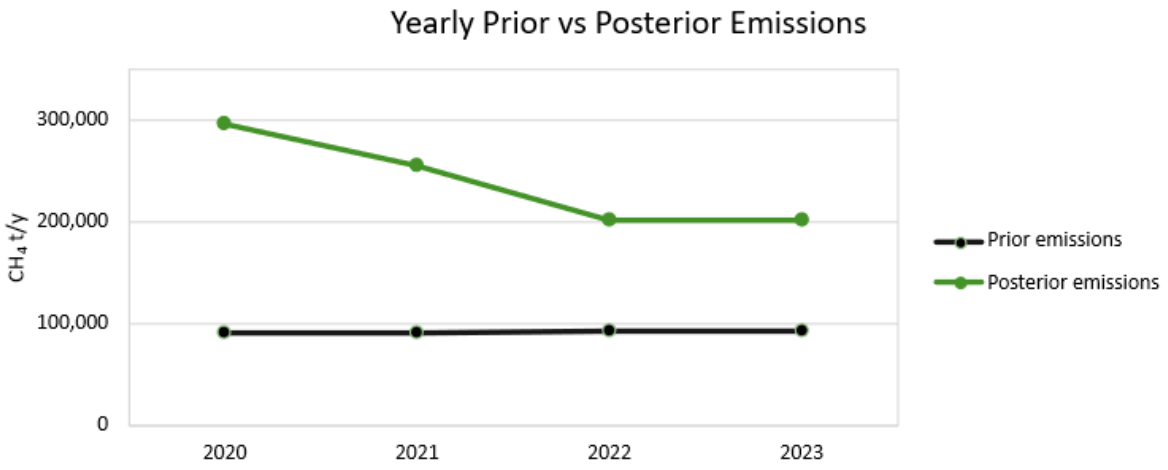


Figure 13: Prior emissions are the bottom-up estimates taken from GFEI data and Posterior emissions are the top down measurements taken from the IMI analysis.

8. Discussion

Emission estimates derived from MSI are reported to have uncertainties ranging from 30% to 70%, due to its low spectral resolution of 100–200 nm (Gorroño et al., 2023; Naus et al., 2023). Measurements have also been noted to have more uncertainty at lower wind speeds (Naus et al., 2023). This makes them unreliable when measuring a single day emission, however over longer periods such errors will average out as the number of measured samples grows, meaning that using them for a yearlong survey as in this report, is appropriate.

8.1 Q_{IME} Measurements

Had the IME method implementation been identical to those used in the academic literature, then the Q_{IME} values should have matched those in the research. That they did not is because the approaches adopted in each study differ. Varon et al. (2021) used a binary mask in a bounding box that was relatively tight around the study plume, whereas in this study, a convex hull around the plume was employed. Naus et al. (2023) averaged a 500 m × 500 m box downwind of the emission source of the previous ten satellite passes for the 'no emission' scene, as opposed to Varon et al., (2021) and the present study, which both used a single reference scene.

Another difference is that Varon et al. (2021) used large eddy simulations to simulate the turbulent diffusion of the plume instead of assuming uniform movement with wind, as was the case here.

The two emissions from Naus et al. (2023) that were unable to be measured correctly were the smallest of the peer reviewed emissions. It is possible that reducing the binary mask threshold from the 95th percentile to 90th would have helped but the intention was to stick to the Varon et al. (2021) method as closely as possible.

8.2 Q_M Measurements

When the variables PPS, W, PL, and SZA were used in the XGBR alone, the maximum evaluation R^2 score achieved after hyperparameter tuning was 0.38. The inclusion of the engineered features PPS_{norm}, WPL, and PWPL, alongside PPS and SZA, significantly improved the model's performance, indicating that these features captured additional relationships not fully exploited by the original variables. However, removing PPS and SZA and using only the engineered features resulted in a notable drop in R^2 to 0.51, suggesting that while the engineered features provided valuable information, they did not fully encapsulate the relationship between the measured variables and Q_M.

The method of calculating Q_{IME} could be improved in a few ways. One would be incorporating SZA into the estimation of PL. Given that increased solar radiance enhances plume visibility, as seen with PPS_{norm}, it is reasonable to expect that brighter conditions would correspond to longer detected plumes and vice versa. Additionally, the ERA5 10 m Windspeed was taken at 10:00am, however with the satellite overpass happening at ~10:30am, it would be better to use the average of the 10:00am and 11:00am values. Finally, as with the Varon et al. (2021), turbulent plume diffusion should be factored into the WPL variable to give a more accurate reflection of the residence time (equation 9).

Based on training and evaluation data, the XGBR model explains 81–85% of Q_{IME} emission variation and has an RMSE of 1.3-1.34 t/h. This level of uncertainty is not unusual for MSI methane column retrieval methods and peer reviewed sources in Hassi Messaoud give uncertainty ranges of 14-67% (Gorroño et al., 2023; Naus et al. 2023; Pandey et al., 2023; Varon et al., 2021).

8.3 Emissions in 2020, 2021 and 2023

This analysis indicates that significant improvements have been made in reducing CH_4 at Hassi Messaoud over the period analysed. However, GFEI-reported methane inventories significantly underestimate total emissions when compared to this analysis.

Naus et al.'s (2023) analysis of 2020 emissions estimated overall totals of 220,000 t/y from Hassi Messaoud using a Sentinel-2/TROPOMI inversion workflow. Of this, 129,536 tonnes were attributed to 337 plume detections across 10 locations. As Naus et al. used a different methodology to calculate overall emissions, this analysis ran the IMI for the year 2020 to ensure compatibility of results. It produced a total emission estimate of 296,032 t/y - approximately 1.35 times higher than the Naus et al. value. By comparison, the GFEI value is 90,325 t/y, which is a significant underestimate. Based on the IMI value and the Naus et al. super-emitter measurements, super emitters contributed approximately 44 % of total emissions in 2020. It should be noted that the two emissions that could not be detected by CH4 Sentinel were both from Naus et al., indicating that their super emitter analysis

methodology was able to capture smaller emissions than was possible with CH4 Sentinel, making their super emitting total not perfectly comparable with this study's results. Despite this, Naus et al. observed 9 super emitting point sources, nearly double what was observed by this study in 2021 and nine times more than in 2023 strongly indicating that the observed reductions in super emitter activity is genuine.

For 2021, this paper's analysis detected 100 super-emitting plume events from five point sources, with sites 5 and 6 dominating emissions, contributing 40 and 44 plume detections respectively. If it is assumed that the detected emissions lasted at least until the next cloud-free satellite pass then the estimated emissions from the detected super emitters in 2021 are 37,325 tonnes; equivalent to powering 140,354 households for a year or a town the size of Bradford, England (Office for National Statistics, 2021, U.S. Environmental Protection Agency, 2024). The IMI revealed 2021 emissions to be 254,916 t/y. This is again much higher than the GFEI of 91,337 t/y, but represents nearly a 13.9% reduction on the previous year. Based on the IMI value of this analysis, super emitters contributed approximately 14.6 % to total emissions, which is also a significant decrease from the 2020 Naus et al. (2023) value.

The 2023 analysis detected 47 emission events all coming from a single point source at site 6. Assuming again that the detected emissions lasted at least until the next cloud-free satellite pass, the total amount of CH_4 released was 14,688 tonnes of CH_4. This is enough energy for 55,232 households for a year; a town the size of Cambridge, England (Office for National Statistics, 2021, U.S. Environmental Protection Agency, 2024). The IMI revealed 2023 emissions to be nearly the same as that of 2022 with a total of 201,837 t/y, once again far above the GFEI value of 92,131. This year saw the lowest contribution of super emitters to total emissions of 7.3 %. This represents a significant decrease from 2021 and 2020 levels.

The analysis estimates the reduction in emissions from super emitters from 2021 to 2023 at 22,636 t/y and overall reductions in that time at 53,079 t/y, meaning that the reduction in super emitter activity accounted for 42.6 % of total reductions over that period. If the overall analysis is extended to 2020, the overall reduction increases to 94,195 t/y indicating that reductions began before the official timeline of 2021 (UNFCCC, 2015). Whether this reflects an improvement in management or a decrease in activity is an important question. Specific oil production rates for Hassi Messaoud are not publicly available, but overall production quantities of Algeria are. Figure 14 shows that 2023 production rates of crude oil are an average of 56,000 barrels higher than in 2021, despite the observed decrease in super emitting plume activity by this study (OPEC, 2025). This suggests that the observed reduction in super emitting plumes likely represents genuine improvement and are not simply the product of declining production levels.

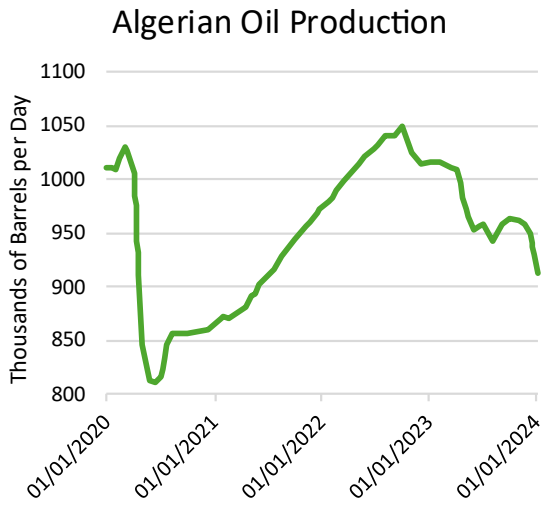


Figure 14: Crude Oil Production in Algeria 2020–2024 (OPEC, 2025)

Liquid natural gas (LNG) is 90-95% CH_4 (Al-Enazi et al., 2021) meaning LMG commodity prices are a good reflection of the value lost in these emissions. At average European Natural Gas Import prices for 2021 and 2023, the gas lost in 2021 from the detected super emitters was worth $27.92 million USD and the quantity of gas lost in 2023 was worth $11.58 million USD (BP, 2021; World Bank, 2025). As super emitters have been shown to be readily detectable using the methods in this research, remote sensing presents an excellent opportunity to reduce these losses.

9 Conclusions

The CH4 Sentinel software demonstrated its ability to localise and quantify methane emissions in a statistically robust manner in the largely spectrally homogenous Algerian desert. Its performance would undoubtably decline when used in spectrally heterogeneous settings (Varon et al., 2021). Notwithstanding that limitation, it remains a valuable, open-source, lightweight tool that could be used in similar environments to improve the economic performance of petrochemical fields while minimising the environmental impact – particularly if cost-

effective capture systems are adopted, such as those incentivised by the European Union under its 'You collect, we buy' scheme, which aims to encourage Sonatrach, Algeria's state petrochemical company, to capture CH_4 (European Commission, 2023).

The reason for prolonged unflared emissions is unclear, as flaring should be standard practice in the disposal of uncaptured CH_4 (Caulton et al., 2014, IPCC, 2006). Either they were known but not managed correctly, or they went undetected by field operators for at least days or weeks at a time. Both scenarios highlight the utility of remote sensing-based emission monitoring for both operators and regulators.

As more top-down analyses for CH_4 are undertaken globally, it is being increasingly realised that bottom-up inventories are frequently underestimating emissions by large margins (Naus et al., 2023; Nesser et al., 2024; Riddick et al., 2024; Varon et al., 2022; Wang et al., 2024b). Likewise, the IMI analyses in this study revealed large underestimations of GFEI values for all years examined. Moreover, despite an overall reduction of a third of emissions in this analysis, the GFEI values remained nearly constant. This suggests that the bottom-up methods of emission estimation used by GFEI are failing to capture either the magnitude of emissions at Hassi Messaoud or even how they change over time.

If the reductions in emissions seen at Hassi Messaoud are mirrored at the other oil and gas fields in Algeria then it bodes well for meeting the 2030 GHG reduction target, at least regarding CH_4. The field reduced its total emissions by nearly a third, which is above the 21% stipulated in the emission reduction plan (UNFCCC, 2015). For a more comprehensive analysis, additional research is needed to quantify the reduction efforts at the other fields as well as quantifying CO_2 emissions, something this study has not investigated.

Data and Code Availability. Complete code and detailed instructions for implementing CH4 Sentinel as well as running the IMI on AWS are available in the GitHub repository https://github.com/Nick-Kinsella/CH4-Sentinel.

Please note that the radiative transfer model used to generate the Q_{IME} training data is proprietary software owned by GHGSat Inc., it cannot be shared publicly without permission from GHGSat Inc. or Dr. Daniel Varon, although it is available free of charge for research purposes if requested from them.

Acknowledgements. Thanks are extended to this project's supervisor, Dr Robert McNabb, for spending many hours setting up the IMI on Ulster University's high-performance computer system, thereby saving the author considerable expense in running that analysis. Dr Daniel Varon of Harvard University was also instrumental in supporting this research with numerous email exchanges and great patience in explaining his research to someone outside his domain. The success of this project would not have been possible without each of their significant contributions.

References

Abada, Z. and Bouharkat, M., 2018. Study of management strategy of energy resources in Algeria. *Energy Reports*, 4, pp.1-7. Available at: https://www.researchgate.net/publication/328663133_Study_of_management_strategy_of_energy_resources_in_Algeria [Accessed 18 Sep. 2024].

Akiba, T., Sano, S., Yanase, T., Ohta, T. and Koyama, M., 2019, July. Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 2623-2631).

BP, 2021, Approximate conversion factors – Statistical Review of World Energy – updated July 2021. Available at: https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2022-approximate-conversion-factors.pdf [Accessed: 14 March 2025].

Caulton, D.R., Shepson, P.B., Cambaliza, M.O., McCabe, D., Baum, E. and Stirm, B.H., 2014. Methane destruction efficiency of natural gas flares associated with shale formation wells. Environmental science & technology, 48(16), pp.9548-9554.

Chen, T. and Guestrin, C., 2016, August. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

Ciampiconi, L., Elwood, A., Leonardi, M., Mohamed, A. and Rozza, A., 2023. A survey and taxonomy of loss functions in machine learning. arXiv preprint arXiv:2301.05579, pp 11

Copernicus, 2025. Sentinel-2 L1C Documentation. [online] Available at: https://documentation.dataspace.copernicus.eu/APIs/SentinelHub/Data/S2L1C.html [Accessed 28 January 2025].

Copernicus, 2023. S5P Mission Performance Centre Methane [L2__CH4___] Readme. Available at: https://sentinels.copernicus.eu/documents/247904/3541451/Sentinel-5P-Methane-Product-Readme-File [Accessed 16 Jul. 2024].

Cronin, T.W., 2014. On the choice of average solar zenith angle. Journal of the Atmospheric Sciences, 71(8), pp.2994-3003.

Cusworth, D.H., Duren, R.M., Thorpe, A.K., Tseng, E., Thompson, D., Guha, A., Newman, S., Foster, K.T. and Miller, C.E., 2020. Using remote sensing to detect, validate, and quantify methane emissions from California solid waste operations. *Environmental Research Letters*, 15(5), p.054012.

Dogniaux, M., Maasakkers, J.D., Varon, D.J. and Aben, I., 2024. Report on Landsat 8 and Sentinel-2B observations of the Nord Stream 2 pipeline methane leak. *Atmospheric Measurement Techniques*, *17*(9), pp.2777-2787.

Dowd, E., Manning, A.J., Orth-Lashley, B., Girard, M., France, J., Fisher, R.E., Lowry, D., Lanoisellé, M., Pitt, J.R., Stanley, K.M. and O'Doherty, S., 2023. First validation of high-resolution satellite-derived methane emissions from an active gas leak in the UK. *EGUsphere*, 2023, pp.1-22.

Duren, R.M., Guido, J., Herner, J., Rao, S., Green, R.O., de Belloy, M., Schingler, R., Ardila, D.R., Thorpe, A.K., Cusworth, D. and Falk, M., 2020, December. Carbon Mapper: global tracking of methane and CO 2 point-sources. In *AGU Fall Meeting Abstracts* (Vol. 2020, pp. A247-01).

Elkind, J., Blanton, E., Denier Van Der Gon, H., Kleinberg, R.L. and Leemhuis, A., 2020. Nowhere to hide: The implications of satellite-based methane detection for policy, industry and finance. *Columbia Center on Global Energy Policy*. Available at: https://energypolicy.columbia.edu [Accessed 15 Nov. 2021].

Ehret, T., De Truchis, A., Mazzolini, M., Morel, J.M., d'Aspremont, A., Lauvaux, T., Duren, R., Cusworth, D. and Facciolo, G., 2022. Global tracking and quantification of oil and gas methane emissions from recurrent sentinel-2 imagery. *Environmental science & technology*, *56*(14), pp.10517-10529.

Estrada, L.A., Varon, D.J., Sulprizio, M., Nesser, H., Chen, Z., Balasus, N., Hancock, S.E., He, M., East, J.D., Mooring, T.A. and Oort Alonso, A., 2024. Integrated Methane Inversion (IMI) 2.0: an improved research and stakeholder tool for monitoring total methane emissions with high resolution worldwide using TROPOMI satellite observations. EGUsphere, 2024, pp.1-31.

European Commission, 2022. Global methane pledge ministerial at COP27. Available at: https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_22_6995 [Accessed 16 Jul. 2024].

European Commission, 2023. EU announces €175m financial support to reduce methane emissions at COP28. *IP/23/6057*. Available at: https://ec.europa.eu/commission/presscorner/detail/en/IP_23_6057 [Accessed 18 Sep. 2024].

Frankenberg, C., Thorpe, A.K., Thompson, D.R., Hulley, G., Kort, E.A., Vance, N., Borchardt, J., Krings, T., Gerilowski, K., Sweeney, C. and Conley, S., 2016. Airborne methane remote measurements reveal heavy-tail flux distribution in Four Corners region. Proceedings of the national academy of sciences, 113(35), pp.9734-9739.

Gordon, I.E., Rothman, L.S., Hargreaves, R.J., Hashemi, R., Karlovets, E.V., Skinner, F.M., *et al.*, 2022. The HITRAN2020 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 277, p.107949. https://doi.org/10.1016/j.jqsrt.2021.107949

Gorroño, J., Varon, D.J., Irakulis-Loitxate, I. and Guanter, L., 2023. Understanding the potential of Sentinel-2 for monitoring methane point emissions. *Atmospheric Measurement Techniques*, 16(1), pp.89-107.

Government of Algeria, 2023. National Inventory Report of Algeria: 2023 Submission under the United Nations Framework Convention on Climate Change (UNFCCC). October 2023. Available at: https://unfccc.int/sites/default/files/resource/NIR_Algeria_Final%20VF%2022102023%20rev%206.pdf [Accessed 18 Mar. 2025]

Guanter, L., Irakulis-Loitxate, I., Gorroño, J., Sánchez-García, E., Cusworth, D.H., Varon, D.J., Cogliati, S. and Colombo, R., 2021. Mapping methane point emissions with the PRISMA spaceborne imaging spectrometer. *Remote Sensing of Environment*, *265*, p.112671.

Integrated Methane Inversion team, 2024. Integrated Methane Initiative: Diagram explaining the methodology. [online] Available at: https://imi.seas.harvard.edu/ [Accessed 9 December 2024].

International Energy Agency, 2022. Global Methane Tracker 2022. *IEA, Paris*. Available at: https://www.iea.org/reports/global-methane-tracker-2022 [Accessed 16 Jul. 2024].

International Energy Agency, 2024. Global Methane Tracker 2024: Tracking pledges, targets, and action. Available at: https://www.iea.org/reports/global-methane-tracker-2024/tracking-pledges-targets-and-action [Accessed 17 Sep. 2024].

International Trade Administration, 2024. Algeria - Oil and Gas (Hydrocarbons). *U.S. Department of Commerce*. Available at: https://www.trade.gov/country-commercial-

guides/algeria-oil-and-gas-hydrocarbons [Accessed 17 Sep. 2024].

Intergovernmental Panel on Climate Change (IPCC), 2006. IPCC guidelines for national greenhouse gas inventories: Volume 2 – Energy, Chapter 4: Fugitive emissions. Hayama, Japan: IGES. Available at: https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/2_Volume2/V2_4_Ch4_Fugitive_Emissions.pdf [Accessed 18 Mar. 2025].

Irakulis-Loitxate, I., Gorroño, J., Zavala-Araiza, D. and Guanter, L., 2022. Satellites detect a methane ultra-emission event from an offshore platform in the Gulf of Mexico. *Environmental Science & Technology Letters*, 9(6), pp.520-525.

Jacob, D.J., Varon, D.J., Cusworth, D.H., Dennison, P.E., Frankenberg, C., Gautam, R., Guanter, L., Kelley, J., McKeever, J., Ott, L.E. and Poulter, B., 2022. Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmospheric Chemistry and Physics*, *22*(14), pp.9617-9646.

Jackson, R.B., Saunois, M., Bousquet, P., Canadell, J.G., Poulter, B., Stavert, A.R., Bergamaschi, P., Niwa, Y., Segers, A., and Tsuruta, A., 2020. Increasing anthropogenic methane emissions arise equally from agricultural and fossil fuel sources. *Environmental Research Letters*, 15(7), p.071002. Available at: https://doi.org/10.1088/1748-9326/ab9ed2 [Accessed 4 Jul. 2024].

Jiang, Y., Zhang, L., Zhang, X. and Cao, X., 2024. Methane retrieval algorithms based on satellite: A review. *Atmosphere*, 15(4), p.449. Available at: https://doi.org/10.3390/atmos15040449 [Accessed 29 Sep. 2024].

Kayrros, 2024. Methane Watch. Available at: https://methanewatch.kayrros.com/map [Accessed 29 Sep. 2024].

Kabir, S., Pahlevan, N., O'Shea, R.E. and Barnes, B.B., 2023. Leveraging Landsat-8/-9 underfly observations to evaluate consistency in reflectance products over aquatic environments. *Remote Sensing of Environment*, *296*, p.113755.

Kamr Eddine Aissou, K., Amri, K., Bendada, R. and Saibi, H., 2024. Variations in the present-day stress field and its implications for hydrocarbon development in the Hassi Terfa field, Hassi Messaoud, Algeria. *Journal of Applied Geophysics*, 222, p.105326. Available at: https://doi.org/10.1016/j.jappgeo.2024.105326.

Karimi, N., Ng, K.T.W. and Richter, A., 2021. Prediction of fugitive landfill gas hotspots using a random forest algorithm and Sentinel-2 data. *Sustainable Cities and Society*, 73, p.103097.

Kuhlmann, G., Koene, E., Meier, S., Santaren, D., Broquet, G., Chevallier, F., Hakkarainen, J., Nurmela, J., Amorós, L., Tamminen, J. and Brunner, D., 2024. The ddeq Python library for point source quantification from remote sensing images (version 1.0). *Geoscientific Model Development*, *17*(12), pp.4773-4789.

Kuhn, M. and Johnson, K., 2013. Applied Predictive Modeling. New York: Springer, pp.32–33.

Liu, M., Van Der A, R., Van Weele, M., Eskes, H., Lu, X., Veefkind, P., De Laat, J., Kong, H., Wang, J., Sun, J. and Ding, J., 2021. A new divergence method to quantify methane emissions using observations of Sentinel-5P TROPOMI. *Geophysical Research Letters*, *48*(18), p.e2021GL094151.

Lopez, L., Barrett, A., Steiker, A.E., Scheick, J., Lowndes, J.S.S. and Robinson, E., 2022, December. A Python Package to Search and Access NASA Earth Science Data. In *AGU Fall Meeting Abstracts* (Vol. 2022, pp. IN22C-0318).

Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., Aan De Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D. and Pollard, D.F., 2021. Methane retrieved from TROPOMI: Improvement of the data product and validation of the first 2 years of measurements. *Atmospheric Measurement Techniques*, *14*(1), pp.665-684.

Naus, S., Maasakkers, J.D., Gautam, R., Omara, M., Stikker, R., Veenstra, A.K., Nathan, B., Irakulis-Loitxate, I., Guanter, L., Pandey, S. and Girard, M., 2023. Assessing the relative importance of satellite-detected methane superemitters in quantifying total emissions for oil and gas production areas in Algeria. *Environmental Science & Technology*, 57(48), pp.19545-19556.

Nielsen, D., 2016. *Tree boosting with xgboost-why does xgboost win" every" machine learning competition?* (Master's thesis, NTNU)

Nesser, H., Jacob, D.J., Maasakkers, J.D., Lorente, A., Chen, Z., Lu, X., Shen, L., Qu, Z., Sulprizio, M.P., Winter, M. and Ma, S., 2024. High-resolution US methane emissions inferred from an inversion of 2019 TROPOMI satellite data: contributions from individual states, urban areas, and landfills. Atmospheric Chemistry and Physics, 24(8), pp.5069-5091.

Maasakkers, J.D., Varon, D.J., Elfarsdóttir, A., McKeever, J., Jervis, D., Mahapatra, R., Pandey, S., Lorente, A., Borsdorff, T., Foorthuis, R., Schuit, B., Tol, P., van Kempen, T., van Hees, R. and Aben, I., 2022. Using satellites to uncover large methane emissions from landfills. *Science Advances*, 8(32), eabn9683. Available at: https://doi.org/10.1126/sciadv.abn9683 [Accessed 4 Jul. 2024].

Massimetti, F., Coppola, D., Laiolo, M., Valade, S., Cigolini, C. and Ripepe, M., 2020. Volcanic hot-

spot detection using SENTINEL-2: A comparison with MODIS–MIROVA thermal data series. Remote Sensing, 12(5), p.820.

Malley, C.S., Borgford-Parnell, N., Haeussling, S., Howard, I.C., Lefèvre, E.N. and Kuylenstierna, J.C., 2023. A roadmap to achieve the global methane pledge. *Environmental Research: Climate*, 2(1), p.011003.

Masek, J.G., Wulder, M.A., Markham, B., McCorkel, J., Crawford, C.J., Storey, J. and Jenstrom, D.T., 2020. Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, *248*, p.111968.

Molina, M.O., Gutiérrez, C. and Sánchez, E., 2021. Comparison of ERA5 surface wind speed climatologies over Europe with observations from the HadISD dataset. International Journal of Climatology, 41(10), pp.4864-4878.

Musial, J., Leszczenski, J., Bojanowski, J., Milcinski, G., Vrecko, A., Clarijs, D., Dries, J., and Marquard, U. (2024) 'Overview of the Copernicus Data Space Ecosystem APIs', *EGU General Assembly 2024*, Vienna, Austria, 14–19 April. Available at: https://doi.org/10.5194/egusphere-egu24-10109.

NASA Jet Propulsion Laboratory, 2024. EMIT Data Portal: Coverage and Forecasts. Available at: https://earth.jpl.nasa.gov/emit/data/data-portal/coverage-and-forecasts/ [Accessed: 6 October 2024].

National Institute of Standards and Technology (2023) 'Methane', NIST Chemistry WebBook. Available at: https://webbook.nist.gov/cgi/cbook.cgi?ID=74-82-8 [Accessed: 28 February 2025].

Office for National Statistics (ONS), 2021. Usual resident population counts and median age, by individual BUA, England (excluding London), 2021. Census 2021. Available at: https://www.ons.gov.uk/visualisations/dvc2257a/fig1/datadownload.xlsx [Accessed 18 March 2025].

Organization of the Petroleum Exporting Countries (OPEC) (2025) Algeria Crude Oil: Production. CEIC. Available at: https://www.ceicdata.com/en/indicator/algeria/crude-oil-production [Accessed: 14 March 2025].

Pandey, S., van Nistelrooij, M., Maasakkers, J.D., Sutar, P., Houweling, S., Varon, D.J., Tol, P., Gains, D., Worden, J. and Aben, I., 2023. Daily detection and quantification of methane leaks using Sentinel-3: a tiered satellite observation approach. *Atmospheric Measurement Techniques*, 16(3), pp.2383-2398.

Parker, R.J., Webb, A., Boesch, H., Somkuti, P., Barrio Guillo, R., Di Noia, A., Kalaitzi, N., Anand, J.S., Bergamaschi, P., Chevallier, F., Palmer, P.I., Feng, L., Deutscher, N.M., Feist, D.G., Griffith, D.W.T., Hase, F., Kivi, R., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D.F., Roehl, C., Sha, M.K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V.A., Warneke, T., Wennberg, P.O. and Wunch, D., 2020. A decade of GOSAT Proxy satellite CH4 observations. *Earth System Science Data*, 12(4), pp.3383-3412. Available at: https://essd.copernicus.org/articles/12/3383/2020/ [Accessed 16 Jul. 2024].

Reshi, A.R., Pichuka, S. and Tripathi, A., 2024. Applications of Sentinel-5P TROPOMI Satellite Sensor: A Review. *IEEE Sensors Journal*.

Riddick, S.N., Mbua, M., Santos, A., Hartzell, W. and Zimmerle, D.J., 2024. Potential Underestimate in Reported Bottom-up Methane Emissions from Oil and Gas Operations in the Delaware Basin. Atmosphere, 15(2), p.202.

Roger, J., Irakulis-Loitxate, I., Valverde, A., Gorroño, J., Chabrillat, S., Brell, M. and Guanter, L., 2024. High-resolution methane mapping with the EnMAP satellite imaging spectroscopy mission. *IEEE Transactions on Geoscience and Remote Sensing*.

Rosenthal, J., 2023. In Reverse: Natural Gas and Politics in the Maghreb and Europe. *FPRI: Foreign Policy Research Institute*. United States of America. Available at: https://coilink.org/20.500.12592/bf2q0b [Accessed 29 September 2024].

Sánchez-García, E., Gorroño, J., Irakulis-Loitxate, I., Varon, D.J. and Guanter, L., 2021. Mapping methane plumes at very high spatial resolution with the WorldView-3 satellite. *Atmospheric Measurement Techniques Discussions*, pp.1-26.

Schuit, B.J., Maasakkers, J.D., Bijl, P., Mahapatra, G., Van den Berg, A.W., Pandey, S., Lorente, A., Borsdorff, T., Houweling, S., Varon, D.J. and McKeever, J., 2023. Automated detection and monitoring of methane super-emitters using satellite data. *Atmospheric Chemistry and Physics Discussions*, *2023*, pp.1-47.

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

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

Sherwin, E.D., El Abbadi, S.H., Burdeau, P.M., Zhang, Z., Chen, Z., Rutherford, J.S., Chen, Y. and Brandt, A.R., 2024b. Single-blind test of nine methane-sensing satellite systems from three

continents. *Atmospheric Measurement Techniques*, *17*(2), pp.765-782.

Siddiqui, A., Halder, S., Kannemadugu, H.B.S. et al., 2024. Detecting methane emissions from space over India: Analysis using EMIT and Sentinel-5P TROPOMI datasets. *Journal of the Indian Society of Remote Sensing*.

Tahchi, B., 2024. Algerian gas to strengthen energy security of the European Union: Policy, capacity and strategy. Energy Reports, 11, pp.3600-3613.

Talamali, S., Chaouchi, R. and Benayad, S., 2016. Sedimentological evolution of the Lower Series formation in the southern area of the Hassi R'Mel field, Saharan Platform, Algeria. *Arabian Journal of Geosciences*, 9, p.481. DOI: 10.1007/s12517-016-2506-7.

Thorpe, A.K., Green, R.O., Thompson, D.R., Brodrick, P.G., Chapman, J.W., Elder, C.D., Irakulis-Loitxate, I., Cusworth, D.H., Ayasse, A.K., Duren, R.M. and Frankenberg, C., 2023. Attribution of individual methane and carbon dioxide emission sources using EMIT observations from space. *Science advances*, *9*(46), p.eadh2391.

United Nations Framework Convention on Climate Change (UNFCCC) (2015) *Algeria – INDC (English unofficial translation)*, 3 September. Available at: https://tinyurl.com/4uvvt34w [Accessed: 1 March 2025].

Unite States Environmental Protection Agency (EPA) (2024) *Greenhouse Gas Equivalencies Calculator*. Available at: https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator [Accessed: 14 March 2025].

Varon, D.J., Jacob, D.J., McKeever, J., Jervis, D., Durak, B.O., Xia, Y. and Huang, Y., 2018. Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmospheric Measurement Techniques*, 11(10), pp.5673-5686.

Varon, D.J., Jervis, D., McKeever, J., Spence, I., Gains, D. and Jacob, D.J., 2021. High-frequency monitoring of anomalous methane point sources with multispectral Sentinel-2 satellite observations. *Atmospheric Measurement Techniques*, 14, pp.2771-2785.

Varon, D.J., Jacob, D.J., Sulprizio, M., Estrada, L.A., Downs, W.B., Shen, L., Hancock, S.E., Nesser, H., Qu, Z., Penn, E., Chen, Z., Lu, X., Lorente, A., Tewari, A. and Randles, C.A., 2022. Integrated Methane Inversion (IMI 1.0): a user-friendly, cloud-based facility for inferring high-resolution methane emissions from TROPOMI satellite observations. *Geoscientific Model Development*, 15, pp.5787–5805. https://doi.org/10.5194/gmd-15-5787-2022.

Vaughn, T.L., Bell, C.S., Pickering, C.K., Schwietzke, S., Heath, G.A., Pétron, G., Zimmerle, D.J., Schnell, R.C. and Nummedal, D., 2018. Temporal variability largely explains top-down/bottom-up difference in methane emission estimates from a natural gas production region. *Proceedings of the National Academy of Sciences*, 115(46), pp.11712-11717.

Vigano, I., Van Weelden, H., Holzinger, R., Keppler, F., McLeod, A. and Röckmann, T., 2008. Effect of UV radiation and temperature on the emission of methane from plant biomass and structural components. *Biogeosciences*, 5(3), pp.937-947.

Wang, H., Fan, X., Jian, H. and Yan, F., 2024a. Exploiting the Matched Filter to Improve the Detection of Methane Plumes with Sentinel-2 Data. *Remote Sensing*, *16*(6), p.1023.

Wang, Y., Fang, M., Lou, Z., He, H., Guo, Y., Pi, X., Wang, Y., Yin, K. and Fei, X., 2024b. Methane emissions from landfills differentially underestimated worldwide. Nature Sustainability, 7(4), pp.496-507.

World Bank (2025) Commodity Markets. The World Bank. Available at: https://www.worldbank.org/en/research/commodity-markets [Accessed: 14 March 2025].

XGBoosting, 2024a. *Configure XGBoost "reg:squarederror" Objective*. [online] Available at: https://xgboosting.com/configure-xgboost-regsquarederror-objective/ [Accessed 20 Apr. 2025]

XGBoosting, 2024b. XGBoost Standardize Numerical Input Features. [online] Available at: https://xgboosting.com/xgboost-standardize-numerical-input-features/ [Accessed 20 Apr. 2025].

Zandbergen, S.R., Duren, R., Giuliano, P., Green, R.O., Haag, J.M., Moore, L.B., Shaw, L. and Mouroulis, P., 2022, September. Optical design of the Carbon Plume Mapper (CPM) imaging spectrometer. In Imaging Spectrometry XXV: Applications, Sensors, and Processing (Vol. 12235, pp. 47-53). SPIE

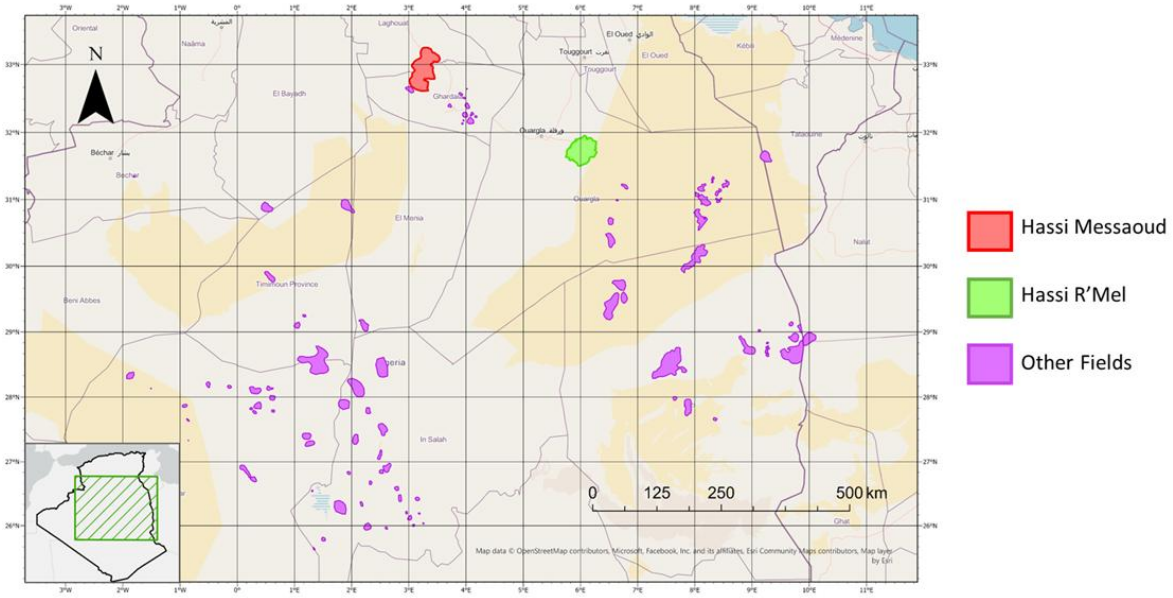
Zhang, L., Tian, H., Shi, H., Pan, S., Qin, X., Pan, N. and Dangal, S.R., 2021. Methane emissions from livestock in East Asia during 1961−2019. *Ecosystem Health and Sustainability*, 7(1), p.1918024.

Zhang, S., Ma, J., Zhang, X. and Guo, C., 2023. Atmospheric remote sensing for anthropogenic methane emissions: Applications and research opportunities. *Science of The Total Environment*, 893, p.164701.

Zhao, Y., Saunois, M., Bousquet, P., Lin, X., Hegglin, M.I., Canadell, J.G., Jackson, R.B. and Zheng, B., 2023. Reconciling the bottom-up and top-down estimates of the methane chemical sink using

multiple observations. *Atmospheric Chemistry and Physics*, 23(1), pp.789-807.

Appendix 1: Map of field locations in Algeria produced by manual survey of satellite data.



Appendix 2: Comparison of Q_{IME} to Study Q

Site	Active date	Lat	Long	Study Q	Q ± kg h−1	IME Q kg−1	Within ± of study Q?	Source
1	2019-11-20	31.6585	5.9053	8497	5700	5259.42	Yes	Varon et al., 2021
1	2020-07-30	31.66	5.8986	14800	7400	10536.46	Yes	Naus et al. 2023
2	2020-01-02	31.6172	5.9674	3600	1800	2091.09	Yes	Naus et al. 2023
3	2020-02-28	31.7569	5.9422	2100	1000	-	Could not detect	Naus et al. 2023
4	2020-02-28	31.7341	5.9677	2700	1400	-	Could not detect	Naus et al. 2023
5	2020-08-14	31.7692	5.9987	3400	1700	3270.14	Yes	Naus et al. 2023
6	2020-08-06	31.7776	5.9917	4800	2400	2359.05	Yes	Naus et al. 2023
8	2020-01-14	31.7571	6.1684	3700	1800	2688.63	Yes	Naus et al. 2023
9	2021-08-31	31.8066	6.1545	5453	2200	5384.88	Yes	Gorroño et al., 2023
10	2020-01-04	31.8647	6.1736	21000	6000	18800.22	Yes	Pandey et al., 2023