GHGSAT

DATA.AIR Gen 1

Description of Technology

Document No. GHG-21850-6101

30 May 2024

GHGSAT PRODUCT/DOCUMENT APPROVAL

Title:	**DATA.AIR Gen 1** **Description of Technology**
Document No.:	**GHG-21850-6101-a**
Submission Date:	**30 May 2024**
Contact Information:	GHGSat Inc. 1130 rue Sherbrooke Ouest, 15ème étage, Montréal, Québec, Canada, H3A 2M8 info@ghgsat.com +1 438 500 6700

CHANGE HISTORY

Version	Release Date	Notes
a	30 May 2024	Initial submission

CONTENT

GHGSAT PRODUCT/DOCUMENT APPROVAL 2
CHANGE HISTORY 2
CONTENT 3
ACRONYMS 4
1 TECHNOLOGY OVERVIEW 5
2 DETAILED DESCRIPTION 7
2.1 Optical system 7
2.2 Measurement Concept 8
2.3 Retrieval Method 11
2.4 Forward Model 11
2.5 Atmospheric Model 12
2.6 Inversion Procedure 13
2.7 Scene-wide Retrieval 14
2.8 Spatially Resolved Column Retrieval 14
2.9 Emissions Identification 15
2.10 Emission Rate Retrieval 15
3 DATA HANDLING, STORAGE AND DELIVERY 16
3.1 Data Delivery to End-Users 16
4 METHOD PERFORMANCE 19
4.1 Methodology 19
4.2 Probability of detection 19
4.3 Quantification Accuracy 20
4.4 Limitations 21
5 REFERENCES 23

Acronyms

AMDT	Advanced Methane Detection Technology
F-P	Fabry-Pérot
FSR	Free Spectral Range
GIS	Geographic Information System
GHG	Greenhouse Gas
GSD	Ground Sampling Distance
LFM	Linear Forward Model
OSF	Order Sorting Filter
TOASR	Top Or Atmosphere Spectral Radiance
WAF-P	Wide Angle Fabry-Pérot

1 TECHNOLOGY OVERVIEW

GHGSat provides remote sensing of methane emissions, based on the principle of optical absorption spectroscopy. Broadband light from the sun traverses the atmosphere, is backscattered by the ground, travels back up, and is then collected by the instrument. Methane absorbs specific wavelengths of the light, a spectral signature defined by molecular vibrational energy levels. A reduction in the light intensity at those wavelengths compared to a normal atmosphere indicates the presence of excess methane. Our spectrometers target methane absorption features in the shortwave infrared (SWIR) portion of the electromagnetic spectrum.

Instruments can be carried by microsatellites (DATA.SAT) or small aircrafts (DATA.AIR), which survey the target facilities as they orbit or fly above them. Both platforms share the same core of technology and processing methods. This document focuses on airborne instruments, which are flown at a typical elevation of 10,000 ft above ground level, providing a swath width of around 750 m. They operate in "nadir" mode, continuously capturing images while their line-of-sight points downwards.

The measurement concept is illustrated in Figure 1. Instruments are based on a wide-angle Fabry-Perot (WAF-P) imaging concept (Jervis et al., 2021; Sloan et al., 2016). The WAF-P element consists of two optical flats mounted within a mechanical enclosure, with the two optical surfaces positioned such that the inner surfaces form an optical cavity (Figure 1.a). This component acts as a wavelength bandpass filter (with possibly more than 1 transmitted mode) which depends on the angle of the light rays inside the cavity (Figure 1.b). During the observation sequence, the ground target traverses the field-of-view, sampling the full extent of the spectral information contained in the images (Figure 1.c).

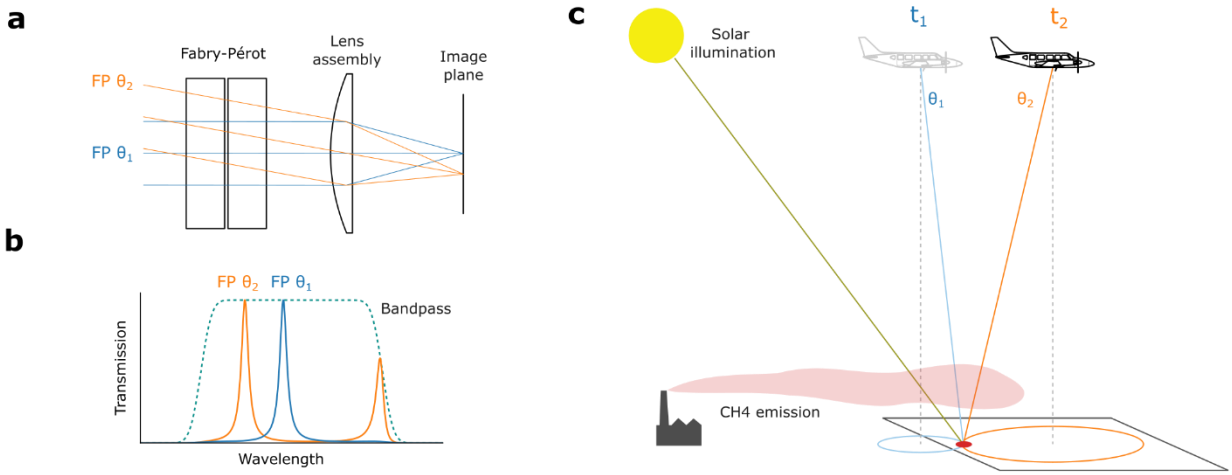

Figure 1. GHGSat DATA.AIR measurement concept. (a) Simplified representation of the optical system showing the FP element, lens, image plane, and light rays from two angles of incidence θ_1 and θ_2. (b) Transmission spectrum of the system for the two angles of incidence. (c) Measurement concept: a given location on the ground (red marker) is observed from different angles to provide information from different wavelengths.

The light intensity signal measured by the instrument can be converted to a methane concentration by using a variant of standard inverse methods. A mathematical forward model is constructed, which describes the wavelength-dependent transmission of light through the

atmosphere and the instrument, given a set of known state parameters (methane concentration, ground reflectance, F-P geometry, etc.). Nonlinear regression is then used to solve the inverse problem of retrieving the state parameters from measured light transmission. A spatially resolved methane concentration map is obtained from this algorithm, which is used to identify emissions which appear as high-concentration hotspots, and to calculate the emission rate. Example concentration maps measured by GHGSat DATA.AIR are provided in Figure 2 below, which shows two methane emissions from US oil & gas facilities.

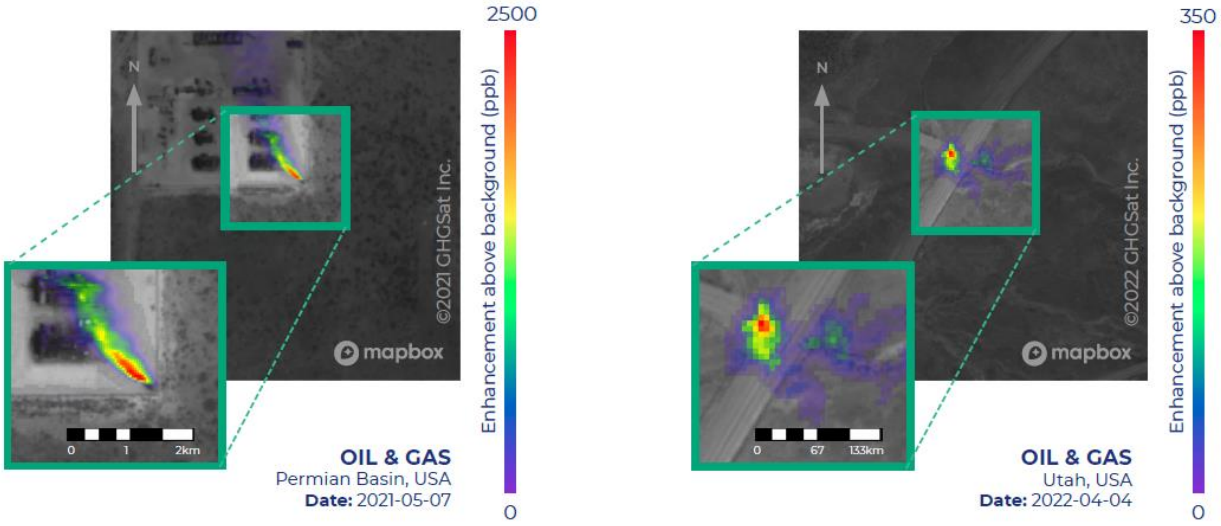

Figure 2. Sample methane concentration maps from GHGSat DATA.AIR for two US oil & gas facilities. Methane concentration enhancement above background is represented in colored heatmap, overlaid over standard aerial imagery.

2 Detailed Description

This section is adapted from peer-reviewed publications describing the measurement technology: Jervis, D et al: The GHGSat-D imaging spectrometer, Atmos. Meas. Tech., 14, 2127–2140, 2021, and D. Varon et al, Atmos. Meas. Tech., 11, 5673–5686, 2018.

A separate *visual workflow* document is provided attached to this application and can be referenced to follow the flow of data and information through the processes described in some of the following sub-sections.

2.1 Optical system

The GHGSat instrument optical system (Figure 3) is composed of three lens assemblies with focal lengths f_{t1}, f_{t2}, and f_{im}: the first two lenses, in confocal arrangement, constitute the telescope. The last lens, the imaging assembly, forms a two-dimensional image of the ground on the detector. The F-P is placed in the Fourier plane of the optical system, between f_{t2} and f_{im}. An order sorting filter (OSF), placed between the lenses that make up the f_{t2} assembly, defines the spectral bandpass region. The choice of focal lengths must balance several considerations simultaneously. First, the angular magnification ratio of the telescope $|f_{t1}/f_{t2}|$ must be consistent with the mechanical constraints on input aperture and F-P size, and the spectroscopic constraints that dictate the desired range of incident ray angles on the F-P. Second, the effective focal length $f = f_{im}|f_{t1}/f_{t2}|$ is constrained by the choice of spatial resolution.

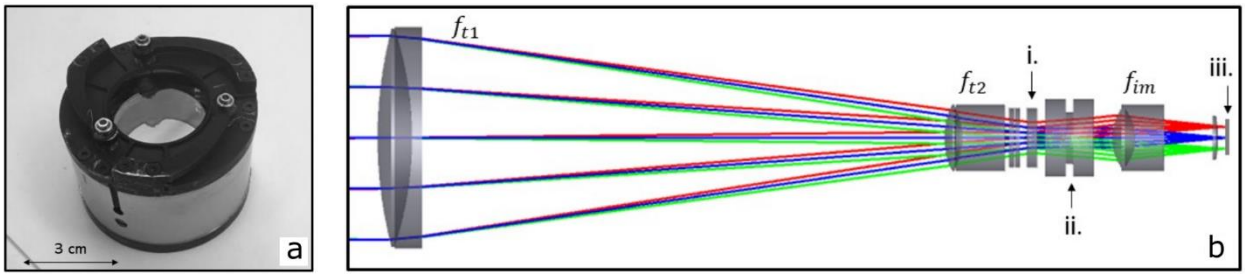

Figure 3: (a) The mounted Fabry-Perot interferometer. (b) Schematic of the unfolded optical system with the i.) OSF, ii.) F-P, and iii.) detector identified. The red, blue, and green rays originate from different ground locations.

For a given image, a polychromatic light ray originating from a specific ground location enters the optical system through the input aperture at some angle pair (ψ, ϕ), where ψ is the small elevation angle and ϕ is the azimuthal angle. In the paraxial approximation, the light ray emerges from the telescope with angle pair $(\left|\frac{f_{t1}}{f_{t2}}\right|\psi, \phi)$ and is incident on the F-P. The imaging assembly then focuses the light ray to detector pixel

$$(i, j) = (i_0 + b\left(\left|\frac{f_{t1}}{f_{t2}}\right|\psi\right)\cos(\phi), j_0 + b\left(\left|\frac{f_{t1}}{f_{t2}}\right|\psi\right)\sin(\phi)), \quad (1)$$

where b is a proportionality constant relating angle to pixel radius. The optical axis intercepts the 1280 x 1024 temperature controlled InGaAs detector array at pixel (i_0, j_0).

The GHGSat instrument F-P element consists of two optical flats mounted within a mechanical enclosure. The two optical surfaces are positioned such that the inner surfaces, with reflectivity R and spaced a distance d apart, form an optical cavity. Light with wavelength λ and incident angle $\theta = \left|\frac{f_{t1}}{f_{t2}}\right|\psi$ with respect to the F-P surface normal is transmitted according to:

$$T_{FP}(\theta,\lambda) = \frac{1}{1+\left(\frac{2\mathcal{F}}{\pi}\right)^2 \sin^2\left(\frac{2\pi n d\cos(\theta)}{\lambda}\right)} \quad (2)$$

where n is the index of refraction of the medium within the optical cavity and $\mathcal{F} = \pi\sqrt{R}/(1-R)$ is the reflectivity finesse. For each value of θ, the transmission spectrum of the F-P is a series of peaks that are spaced in wavelength by the free spectral range $FSR = \lambda^2/(2d)$ with a spectral width characterized by the full-width half-maximum $FWHM = FSR/\mathcal{F}$. Because the F-P accepts a continuum of θ values it samples a continuum of wavelengths within the passband.

In applications using F-P based spectrometers, the F-P gap spacing d is often dynamically scanned during the measurement (Reay et al., 1974). In contrast, the GHGSat instrument uses a fixed gap spacing and exploits the angular dependence of the m^{th} F-P transmission mode's spectral position $\lambda_m = 2\pi n d\cos(\theta)/m$ to measure the spectrum of the incident light. This approach simplifies the mechanical design and makes it much easier to meet stringent stability requirements. The FSR is chosen so that there are enough F-P transmission modes within the spectral bandpass to sample the entirety of the spectral bandpass within the imaging field of view. The F-P is temperature-controlled to keep thermal mechanical drift to a minimum.

Table 1: Instrument and operational parameters

Parameter	Value	Comments
Spectral range	~1.6 – 1.7 µm	Exact range within these bounds
Spectral sampling	0.0001 – 0.1 nm/pixel	Spectral sampling is nonlinear across the detector due to F-P transmission mode behaviour. Spectral sampling is finer near the center of the detector and coarser at greater radii.
Ground sampling distance (GSD)	~1.5 m	Depends on altitude and viewing angle
Transverse field of view	~750 m	Illuminated portion of each circular frame defines the swath width
Flight speed	~140 knots	
Flight altitude (Above ground level)	~10,000 ft	

2.2 Measurement Concept

The GHGSat spectrometer is based on a wide-angle Fabry-Perot (WAF-P) imaging concept (Sloan et al., 2016). A programmable number of closely overlapping two-dimensional images are taken in which the atmospheric absorption spectrum is "imprinted" on the images in the form of spectral rings due to the angle-dependent Fabry-Perot transmission spectrum. During the observation sequence, the ground target traverses the field-of-view, sampling the full extent of the spectral information contained in the images.

Figure 4 illustrates how the F-P samples the backscattered solar radiance spectra to generate spectral rings in the image. Multiple F-P transmission modes are allowed through the OSF

bandpass. Because the F-P transmission function depends only on θ, it is circularly symmetric and so can be expressed as a function of radius $r = \sqrt{(i - i_0)^2 + (j - j_0)^2}$. The $r = 0$ F-P transmission spectrum is shown in Figure 4(a) alongside the OSF transmission function and the normalized backscattered spectral radiance observed by the instrument (referred to as top-of-atmosphere spectral radiance or TOASR). For larger radii – and thus larger θ – the F-P spectrum shifts to lower wavelength with a $\cos(\theta)$ dependence, allowing us to sample different regions of the TOASR. Figure 4(b) shows the location (in wavelength space) of the F-P transmission peaks as a function of radius overlaid on the normalized TOASR. The instrument signal is shown in Figure 4(c). At each radius, the signal on the detector array is the wavelength integral of the TOASR multiplied by the F-P and OSF transmission spectra, i.e. the result of integrating Figure 4(b) along the vertical axis. A mathematical description of the forward model is given in Sect. 2.4.

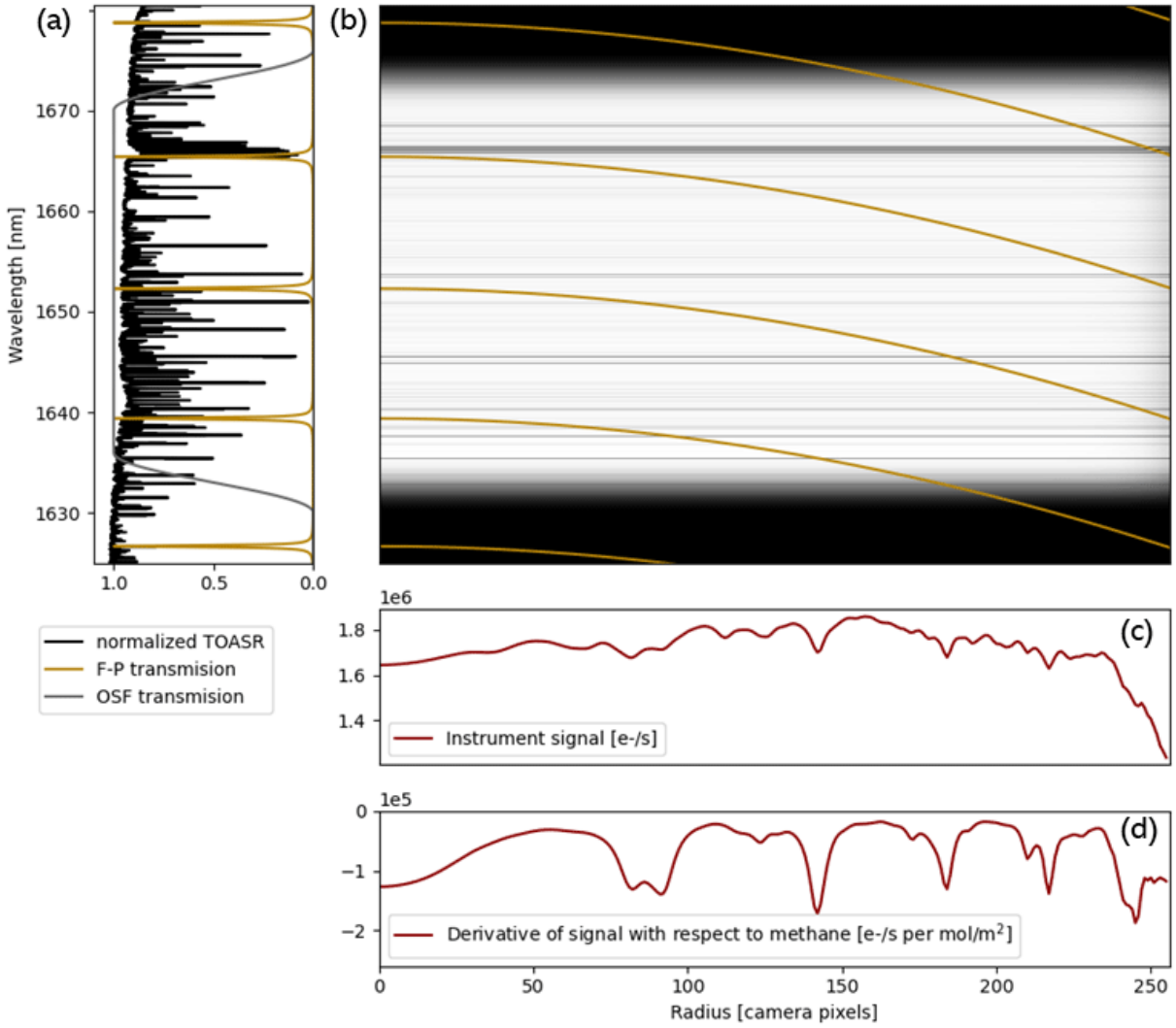

Figure 4: The plots (a) and (b) illustrate how the instrument signal in (c) is produced. In (a), the F-P and OSF transmission spectra are shown alongside the normalized top-of-atmosphere spectral radiance (TOASR) for the case of light rays incident on the F-P at normal incidence (i.e. $(\boldsymbol{r}, \boldsymbol{\theta}) = \mathbf{0}$). In (b), the location of the F-P transmission peaks are shown as a function of radius (gold lines) overlaid on the normalized TOASR (grey-scale background image). The horizontal dark bands at the top and bottom of (b) illustrate wavelengths where the OSF transmission is reduced to zero. By sampling a continuum of incident angles $\boldsymbol{\theta}$, the transmitted F-P transmission peaks measure a continuum if wavelengths within the passband. The instrument signal (c) results from integrating the multiplied signals in (b) along the vertical (wavelength) axis. The TOASR and instrument signal are calculated assuming a target elevation at sea level, a solar zenith angle of $40°$, and vertical column densities of 0.68 mol m^{-2} (methane), 160 mol m^{-2} (CO_2), and 830 mol m^{-2} (water vapor).

In order to measure the spectrum of solar radiation backscattered from a specific ground cell, the location of the ground cell in each image must be known. This is done with an image co-registration algorithm. We then construct a spectrum for each ground cell along the image frame axis by recording the measured signal as a function of the ground cell's radial position with respect to the spectral ring center. Figure 5 (a)-(d) show an observation where the location of an example ground cell has been tracked in each frame. The constructed spectrum is shown in Figure 5 (e). The colour of the data point represents the image frame from which the data was obtained. We construct approximately 200,000 of these spectra in order to retrieve the methane column density for each ground cell within the retrieval domain.

2.3 Retrieval Method

The goal of the retrieval algorithm is to estimate the instrument and atmospheric state vector $\mathbf{x}$ from a measurement vector $\boldsymbol{y}$. This is done by constructing a combined forward model $\boldsymbol{F}(\mathbf{x})$ of the instrument and atmosphere and making the association:

$$\boldsymbol{y} = \boldsymbol{F}(\mathbf{x}) + \boldsymbol{\epsilon}_y + \boldsymbol{\epsilon}_F \tag{3}$$

where $\boldsymbol{\epsilon}_y$ represents the measurement error and $\boldsymbol{\epsilon}_F$ represents error in the forward model. A retrieval of $\mathbf{x}$ requires that we have accurate knowledge of both the forward model and the errors in the measurement system. Because $\boldsymbol{F}(\mathbf{x})$ is a nonlinear function of $\mathbf{x}$, we must solve for the state vector iteratively. This requires knowledge of the Jacobian $\boldsymbol{K}(\mathbf{x}) = \frac{\partial \boldsymbol{F}(\mathbf{x})}{\partial \mathbf{x}}$ to weight the state vector step $\boldsymbol{\Delta}\mathbf{x}^i$ taken during the i^{th} iteration. In this section, we describe the instrument and atmospheric forward model and outline the inversion procedure used to estimate $\mathbf{x}$.

2.4 Forward Model

The forward model represents our best knowledge of the instrument and atmosphere, with approximations used to evaluate the model more efficiently when performing retrievals. The camera signal $F_{i,j}$ at detector pixel (i,j) in photocurrent units [$e^- s^{-1}$] is given by:

$$F_{i,j}(\mathbf{x}) = \int L(\mathbf{x},\lambda) \cdot C(\lambda) \cdot QE(\lambda) \cdot T_{OSF}(\lambda) \cdot T_{FP}(\theta,\lambda) d\lambda \tag{4}$$

where $L(\mathbf{x},\lambda)$ is the spectral radiance as a function of the state parameter $\mathbf{x}$ and wavelength λ, $C(\lambda)$ is the radiometric conversion factor that converts spectral radiance to the number of photons on a pixel per unit time, $QE(\lambda)$ is the quantum efficiency with which the camera converts a photon to electric charge, $T_{OSF}(\lambda)$ is the transmission of the order-sorting filter that defines the spectral bandpass region, and $T_{FP}(\theta,\lambda)$ is the F-P transmission function defined in Eq. (2). The camera signal in Eq. (4) is plotted as a function of radius in Figure 5 (c).

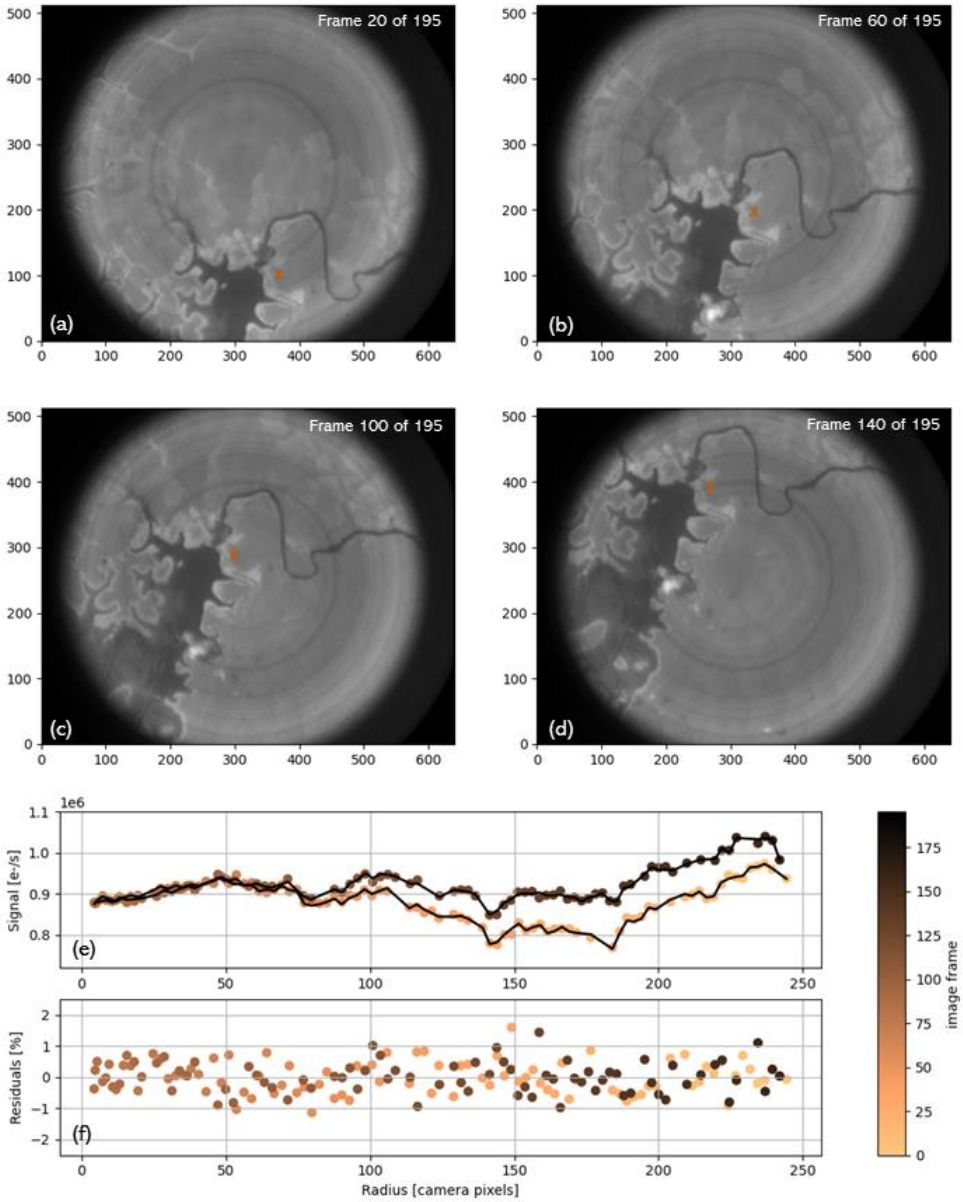

Figure 5. The images (a), (b), (c)., and (d). show a selection of frames from a satellite observation over the Lom Pangar hydroelectric reservoir in Cameroon taken on April 20th, 2017 with an example ground location (denoted by an orange "x") tracked in each frame. The same measurement concept is applicable to DATA.AIR acquired from an aircraft. The image axes are in pixels, with each pixel representing a 24 x 24 m² area on the ground. The plot in (e) shows the signal (circles) from the example ground location as a function of the image frame (circle colour) and radius from spectral ring center (horizontal axis). The forward model (black line) is plotted alongside the signal data and residuals between model and data are shown in (f).

2.5 Atmospheric Model

The spectral radiance $L(\mathbf{x}, \lambda)$ is calculated from the spectral irradiance $I(\lambda)$ assuming Lambertian surface reflectance:

$$L(\mathbf{x},\lambda) = \frac{a(\lambda)\cdot\cos(\theta_{sza})}{\pi R_{E-S}^2} I(\mathbf{x}',\lambda) \quad (5)$$

where $a(\lambda)$ is the spectrally-dependent surface albedo, θ_{sza} is the solar zenith angle, R_{E-S} is the relative Earth-Sun distance, $\mathbf{x}'$ is the state parameter vector without the albedo, and the spectral irradiance is the solution to a simplified radiative transfer equation where thermal emission, aerosol and molecular scattering have been neglected (Chandrasekhar, 2013):

$$\mu \frac{\partial I(\mathbf{x}',\lambda)}{\partial z} = -\alpha_{abs} I(\mathbf{x}',\lambda). \quad (6)$$

This equation is integrated along the downwelling and upwelling light path. Here $\mu = \cos(\theta)$, θ is the angle that the light travels through the atmosphere with respect to the Earth's surface normal, z is the altitude, α_{abs} is the pressure, temperature, wavelength, and species dependent absorption coefficient calculated using the HITRAN API (Kochanov et al., 2016), and the solar irradiance is introduced through a boundary condition and generated from the AER solar irradiance model (Clough et al., 2005). We integrate the radiative transfer equation discretely assuming 100 atmospheric layers that are evenly spaced in pressure.

We justify excluding thermal emission and molecular scattering from the atmospheric model because both are small effects for the wavelengths within our spectral bandpass region. Previous studies of simulated carbon dioxide retrievals using only the 1563 - 1585 nm band have found that neglecting aerosol and molecular scattering can lead to a few percent error, depending on the surface albedo and aerosol optical depth (Aben et al., 2007). This error can be either positive or negative, depending on whether the presence of aerosols – in combination with the surface albedo - leads to an increase or decrease in the average optical pathlength. An analysis of SCIAMACHY retrievals in this same spectral region found that errors decrease for aerosol vertical distributions that are narrow and closer to the Earth's surface (Houweling et al., 2005). GHGSat retrievals are primarily intended to measure local plume enhancements. Therefore, we are especially concerned with any unmodeled effects with spatial structure on the length scales of emission plumes. This could potentially include aerosol scattering, such as aerosols that might conceivably be co-emitted with methane plumes. However, since the presence of these aerosol plumes would be much closer to the surface and narrower in vertical profile than the aerosol profiles retrieved in (Aben et al., 2007; Houweling et al., 2005), we expect that errors arising from neglecting scattering should be small compared with other sources of measurement error, similar to what is assumed in AVIRIS airborne methane retrievals (Thorpe et al., 2014). We also note a recent study in which the effect of neglecting aerosols in an AVIRIS airborne methane plume retrieval was shown to be < 5% for representative methane plume enhancements and a significant aerosol optical depth (Huang et al., 2020).

2.6 Inversion Procedure

For any ground cell (p,q) in a reference frame we can compare the observation data vector $\mathbf{y}^{(pq)} = \{y_{i_1,j_1}^{(pq)}, y_{i_2,j_2}^{(pq)}, \ldots, y_{i_k,j_k}^{(pq)}\}$ to the forward model vector $\mathbf{F}(\mathbf{x}^{(pq)}) = \{F_{i_1,j_1}(\mathbf{x}^{(pq)}), F_{i_2,j_2}(\mathbf{x}^{(pq)}), \ldots, F_{i_k,j_k}(\mathbf{x}^{(pq)})\}$ and infer the state vector $\mathbf{x}^{(pq)}$ using a variant of standard inverse methods described below. The subscripts refer to the pixel indices for this ground cell within the respective frames of the image sequence from 1 to k. The reference frame coordinate system can then be georeferenced with the appropriate rotational and scale transformation.

We use optimal estimation (Rodgers, 2000) to infer the posterior distribution of the state vector given the observation data, an error model, and a prior distribution for the state vector. Assuming a Gaussian form for the measurement and prior probability density functions, maximizing the joint probability density function amounts to minimizing the cost function:

$$\chi^2(\mathbf{x}) = \left(\mathbf{y} - \boldsymbol{F}(\mathbf{x})\right)^{\mathrm{T}} \mathbf{S_o^{-1}} \left(\mathbf{y} - \boldsymbol{F}(\mathbf{x})\right) + (\mathbf{x} - \mathbf{x}_a)^{\mathrm{T}} \mathbf{S_a^{-1}} (\mathbf{x} - \mathbf{x}_{\mathrm{a}}) \quad (7)$$

where $\mathbf{S_o}$ is the observation error covariance matrix, $\mathbf{S_a}$ is the prior covariance matrix, and $\mathbf{x}_{\mathrm{a}}$ is the prior state vector. The Gauss-Newton procedure for minimizing the cost function requires that we update the state vector at each iteration by a step:

$$\Delta\mathbf{x}^{i+1} = (\boldsymbol{K}^{\mathrm{T}}(\mathbf{x}^i)\mathbf{S_o^{-1}}\boldsymbol{K}(\mathbf{x}^i) + \mathbf{S_a^{-1}})^{-1} \left(\boldsymbol{K}^{\mathrm{T}}(\mathbf{x}^i)\mathbf{S_o^{-1}} \left(\mathbf{y} - \boldsymbol{F}(\mathbf{x}^i)\right) + \mathbf{S_a^{-1}}(\mathbf{x} - \mathbf{x}_{\mathrm{a}})\right) \quad (8)$$

where $\boldsymbol{K}(\mathbf{x}^i)$ is the Jacobian of the forward model evaluated at $\mathbf{x}^i$.

At each iteration of the Gauss-Newton procedure, the forward model and Jacobian must be evaluated. This is computationally expensive for a single cell and evaluating it for the ~200,000 ground cells in our field-of-view is impractical. Instead, we use a two-step procedure: (1) a scene-wide average retrieval using the full forward model to estimate the scene-wide average state vector $\hat{\mathbf{x}}$, and (2) a per-cell retrieval done using a linearized forward model (LFM) evaluated at the linearization point $\hat{\mathbf{x}}$. The separate retrieval steps are described in the following sections. As a note, full flight lines can consist of an arbitrarily large number of frames. For retrievals, they are broken down into overlapping "chunks" of around 200 frames which are processed independently.

2.7 Scene-wide Retrieval

The goal of the scene-wide retrieval is two-fold: to retrieve scene-wide averaged surface and atmospheric parameters such as albedo and molecular column density, and to retrieve the F-P gap spacing. Even though the F-P is thermally stabilized, residual drift in the F-P gap spacing can occur during a survey. Because the F-P gap spacing directly affects the signal on each detector pixel, we retrieve d for each chunk. The scene-wide retrieval uses the full instrument model $F_{i,j}(\mathbf{x})$ from Eq. (4) in the optimal estimation procedure. The data vector $\mathbf{y}$ in the scene-wide retrieval is the radial average of the average of all image frames. The result of the scene-wide retrieval is the state parameter estimate $\hat{\mathbf{x}}$ which includes retrieved instrument parameters and the scene-wide averaged methane column density.

2.8 Spatially Resolved Column Retrieval

A linearized forward model (LFM) is constructed at the linearization point $\hat{\mathbf{x}}$:

$$\mathbf{F}^{LFM}\left(\mathbf{x}^{(pq)}\right) = \mathbf{x}^{(pq)}(1)\left(\mathbf{K}_1(\hat{\mathbf{x}}) + \frac{1}{\hat{\mathbf{x}}(1)}\sum_{l=2}^{n}\mathbf{K}_l(\hat{\mathbf{x}})\left(\mathbf{x}^{(pq)}(l) - \hat{\mathbf{x}}(l)\right)\right) + \sum_{l=n}^{n+m}\mathbf{K}_l(\hat{\mathbf{x}})\left(\mathbf{x}^{(pq)}(l) - \hat{\mathbf{x}}(l)\right) \quad (9)$$

where the first element of the state vector is taken to be the surface reflectance. The methane column density and ground reflectance model parameters are retrieved for each ground cell (p, q). Here, we set the prior variances to be very large such that the retrieved parameters are almost entirely determined by the data. There are two terms in the LFM: one with n terms that includes state parameters whose Jacobians scale with the surface reflectance (for example the molecular column densities), and another with m terms for state parameters whose Jacobians are not scaled by the reflectance. In particular, the molecular components of the Jacobian are calculated using a column density enhancement in the lowest atmospheric layer that extends from the surface to approximately 100 m altitude. The primary advantage of using the LFM is that we only compute the forward model and Jacobians once at the beginning of the per-cell retrieval. A disadvantage is that for the parameters that are nonlinear in the forward model, a retrieval using the LFM will introduce systematic biases for deviations far from the linearization point. For the particular case of molecular column densities, this leads to an underestimation that is corrected in post-processing using a non-linear correction function determined from a comparison of $\mathbf{F}^{LFM}(\mathbf{x})$ with $\mathbf{F}(\mathbf{x})$ at the appropriate linearization point (Varon et al., 2019).

2.9 Emissions Identification

Emissions are identified by visual inspection of the spatially resolved methane density retrieved at the previous step. When an emission is detected, the operator identifies its origin based on wind direction and plume morphology. Additionally, a binary mask of the plume is computed using the flood-fill algorithm. This mask identifies which connected ground cells have a methane concentration above the measurement uncertainty and originate from the defined source.

2.10 Emission Rate Retrieval

The methane emission rate is calculated using the method of integrated mass enhancement (IME) (Varon et al., 2018). IME is defined as the total amount of detectable methane enhancement that originates from a given source. For an observed plume comprising N ground cells of area A_j ($j = 1 \dots N$) and column density enhancement $\Delta\Omega_j$ (output of step 2.8),

$$\mathrm{IME} = \sum_{j=1}^{N} \Delta\Omega_j A_j \tag{10}$$

IME measures the local accumulation of methane, which is driven by two factors: the emission rate Q, and the rate at which the methane dissipates in the surrounding atmosphere characterized by its residence time τ. One can express τ dimensionally in terms of an effective wind speed U_{eff} [m s^{-1}] and a plume size L [m]:

$$Q = \frac{1}{\tau}\mathrm{IME} = \frac{U_{\mathrm{eff}}}{L}\mathrm{IME} = \frac{U_{\mathrm{eff}}}{L}\sum_{j=1}^{N} \Delta\Omega_j A_j \tag{11}$$

U_{eff} and L would have simple physical meanings of wind speed and plume length if dissipation of the plume occurred by uniform transport to a terminal distance downwind of the source. But the actual mechanism for plume dissipation is turbulent diffusion, which takes place in all directions. U_{eff} and L must therefore be viewed as operational parameters to be related to observations of wind speed U and plume extent. These relationships take the following empirical form:

$$L = \sqrt{A_M} \tag{12}$$

$$U_{\mathrm{eff}} = \alpha_1 \mathrm{U}_{10} + \alpha_2 \tag{13}$$

In the equations above: A_M is the area of the plume mask (connected ground cells where the enhancement is above the measurement uncertainty level); U_{10} is the wind speed at 10 m above ground, obtained from the GEOS-FP database and interpolated to the time and coordinates of the emission; α_1 and α_2 are model parameters inferred from controlled release experiments.

More conceptual details about the methodology and validation against other quantification methods are provided in (Varon et al., 2018).

3 Data Handling, Storage and Delivery

DATA.AIR data is collected using GHGSat's airborne instruments. The data is collected when the aircraft flies over the customer-requested targets based on a scheduled plan defined with the customer. This data is stored on a solid state drive (SSD) during the flight and then uploaded to an Amazon S3 bucket after the flight for processing

The processing of the raw data to produce the products that are delivered to end users is detailed in Section 3.1, which occurs in an Amazon S3 bucket (cloud-based solution). The procedures for processing the data are consistent for each observation and are documented in onboarding documentation for the end-users. If there are any tangible changes to the procedures, an update will be sent to the end-users. Information on manipulations to the raw data for interpretation of the products can be found in the associated metadata.

3.1 Data Delivery to End-Users

Search, discover, and data access capabilities are provided through GHGSat's secure HTTPS online data system called SPECTRA, which provides both visibility to GHGSat's catalogue as well as access to user-specific observation products and data. Each approved user is provided with unique and password-protected access credentials to SPECTRA, and e-mail notifications for emissions detected in the data can be enabled. The SPECTRA interface enables customized data visualization in which users can import or export data layers through ArcGIS REST enpoints or link to their own GIS client software solutions. An application programming interface (API) is available for retrieving data products or parameters.

The file packages delivered to the user's SPECTRA account are detailed in Table 2 and Table 3.

Table 2 GHGSat Product Delivery Package

Item	Contents
Imagery / Data	The image or data product, order as per call-up, with unique ID for each image / data product. The details of these images or data products are found in Table 3.
Metadata	The metadata and ancillary information for each image are provided in the JSON file format. This metadata includes the size, description, units, spatial coordinates (latitude and longitude) and pixel size, which is sufficient to allow a qualified programmer to read the volume and transform its contents into an image.
Additional Information	Any associated processing files and documentation to help the user to understand the product and facilitate data use.

Table 3 GHGSat Data Files or Products

File	Description	Format
Albedo	The per-pixel short-wave infrared (SWIR) surface reflectance. The pixel values represent the fraction of light received on a surface that is reflected. A value of zero indicates that no light is reflected and a value of one indicates that all the light received on a surface is reflected.	GeoTIFF

Reduced resolution albedo and World file	Resolution of the image is reduced compared to the albedo layer. The World file provides reference information for the geographic location of the data.	PNG, WLD
Methane concentration measurement	The methane concentration values in each pixel represent how many methane particles are in a give volume of air compared to the local background. The background is set to a value of zero. The concentration of the methane particles is measured in parts per billion (ppb) and is a column averaged concentration. The methane concentration GeoTIFF shows methane concentration of the entire observation. Note that this layer can contain artifacts (false positive methane enhancements); enhancements associated to real methane emissions are identified in the methane plume layer.	GeoTIFF
Methane concentration measurement error	Per-pixel methane measurement uncertainty. The uncertainty is for a single flight pass and includes instrument errors and the quality of the model fitted on the data. The error is defined as the standard deviation.	GeoTIFF
Methane concentration measurement isolated for each emission plume	This layer shows the methane concentration reported in the methane concentration measurement layer only at the location where a plume was detected.	GeoTIFF
Methane emission concentration map	High readability pseudo-colour map combining background imagery and the isolated emission plume identified in the methane plume layer.	PNG
Flags	This is a layer that presents the quality of each pixel in the methane concentration measurement layer. All the values in the flag layer are 1, 2, or 3: 1. Good: The pixel is considered to have good quality data. The values obtained at this location can be trusted. 2. No Data: There was no data available in the raw data at this location. 3. Bad Fit: The pixel is considered to have poor quality data. Typically, this means that the error at this location is high and/or that the signal is low. The value obtained at this location cannot always be trusted and need to be interpreted with caution.	GeoTIFF
Information about an identified emission	This file includes the spatial coordinates (latitude, longitude), detection time, wind speed, estimated source rate and uncertainty on the	CSV

	estimated source rate of the identified plume. Multiple methane emissions can be detected in one observation and each emission has its own ID.	

4 METHOD PERFORMANCE

GHGSat uses controlled releases as its primary technique for calibration and validation. This approach is appropriate to quantify two key performance metrics: detection threshold and quantification accuracy. Both our satellite and aircraft technologies have been tested in a series of independent third party studies led by Stanford University (El Abbadi et al., 2023; Rutherford et al., 2023; Sherwin et al., 2023a, 2023b). The articles are attached to this application as supporting documentation. This section provides a summary of the method and results, as well as complementary analysis to determine the aggregate detection threshold of 13.6 kg/hr (90% probability of detection at a wind speed of 3 m/s).

4.1 Methodology

Campaigns were conducted during the periods of October to November 2021 (satellite) and October to November 2022 (aircraft). The study design is a single-blind, multi-staged experiment. The auditor (Stanford University) performs a series of controlled releases at various rates (including possible nulls). The measurement provider (GHGSat) is aware of the test campaign and of the approximate release location, but does not know the exact equipment configuration, if methane is being released at any given time, or the true flow rate.

Flow rate was held constant during individual passes of the instrument but varied between passes and from day to day, allowing to probe detection sensitivity over the dynamic range of the instrument and under varied wind conditions.

The data collected by the auditor includes precisely metered flow rate, local wind speed and direction, and their associated uncertainties. At stage 1, the measurement provider has no access to this information, and therefore only has data obtained through the measurement system under test. Auditor data is progressively revealed at later stages of the experiment, allowing finer analysis of quantification accuracy, for example related to wind speed uncertainty. Limit-of-detection results presented in this document are from stage 1 (fully blinded).

4.2 Probability of detection

The probability of detection model for DATA.AIR is inferred from a set of 140 fully blinded measurements, ranging from 1.05 to 1140 kg CH_4 / hr (El Abbadi et al., 2023). Individual emission assessments (detection or no-detection) are presented as dark circles in Figure 6 below.

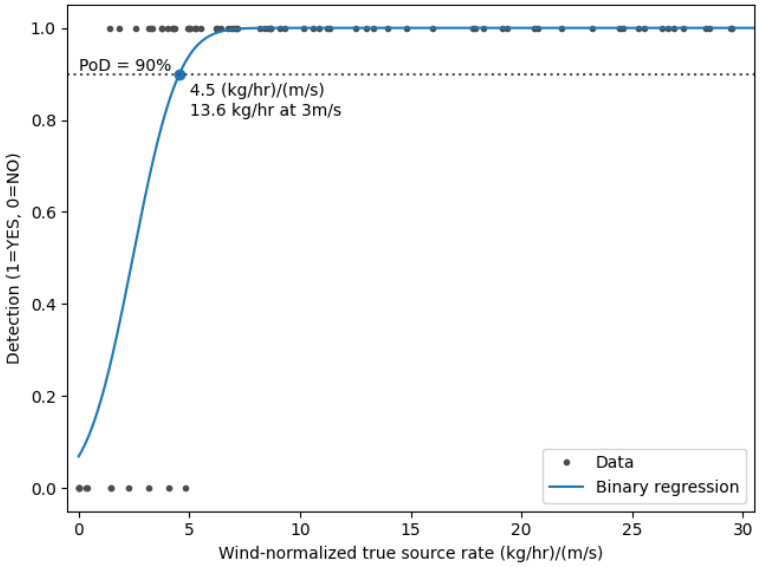

Figure 6. DATA.AIR emission detections from blind controlled releases and probability of detection model fit.

Emission detection is probabilistic by nature. There is no "hard cutoff" emission rate defining the boundary between successful and unsuccessful detection. Rather, there is a smooth curve expressing probability of detection (PoD) vs emission rate. The PoD is inferred by fitting a probit model (full blue line in Figure 6) to the detection data, defined by the equation:

$$p = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\beta_0 + \beta_1 \tilde{q}}{\sqrt{2}}\right)\right) \tag{14}$$

where p is the probability of detection erf is the error function, β_0 and β_1 are model parameters (fitted coefficients), and the independent variable $\tilde{q} = Q/U$ is the emission rate normalized by the wind speed. Wind speed has a strong influence on detection probability, which is why it must be accounted for explicitly. A linear relationship between Q and U is generally considered a good approximation for point sources (Jacob et al., 2016; Jervis et al., 2021), which the Q/U normalization derives from.

The parameters β_0 and β_1 are fitted to the results of the stage 1 (fully blinded) controlled releases, using a maximum likelihood optimization algorithm. To quote a single value for detection limit from this distribution, one must specify the desired probability of detection and wind speed. A PoD of 90% (dotted line in Figure 6) is used in this document for consistency with EPA definitions, at a reference wind speed of U_0 = 3 m/s. The corresponding detection threshold is 13.6 kg / hr.

4.3 Quantification Accuracy

Each emission quantification Q is associated with a measurement uncertainty estimate ΔQ, which accounts for variability in observation conditions such as ground reflectance, sun illumination, terrain, and wind data. The uncertainty is defined as the sum in quadrature of individual error components:

$$\Delta Q = \sqrt{\Delta Q_{\mathrm{meas}}^2 + \Delta Q_{\mathrm{wind}}^2 + \Delta Q_{\mathrm{mod}}^2} \tag{15}$$

where ΔQ_{meas} is the measurement error determined from the variance of retrieved methane concentration, ΔQ_{wind} is the error on the wind speed (whether it was locally measured or sourced from GEOS-FP), and ΔQ_{mod} quantifies the uncertainty of the IME to source rate conversion (Varon et al., 2018).

Additionally, the absolute accuracy of the quantification can be evaluated from the same controlled release data used to establish the detection threshold, by comparing the true emission rate with the GHGSat measurement. Results are shown in Figure 7 for two stages of the unblinding: (a) fully blinded, and (b) the locally measured wind speed is known. At stage 1, the results show the true source rate is generally contained within the measurement uncertainty ΔQ (expressed here as a 95% confidence interval), with a modest negative bias (under-estimation). This bias is removed at stage 2 when more accurate wind information is available.

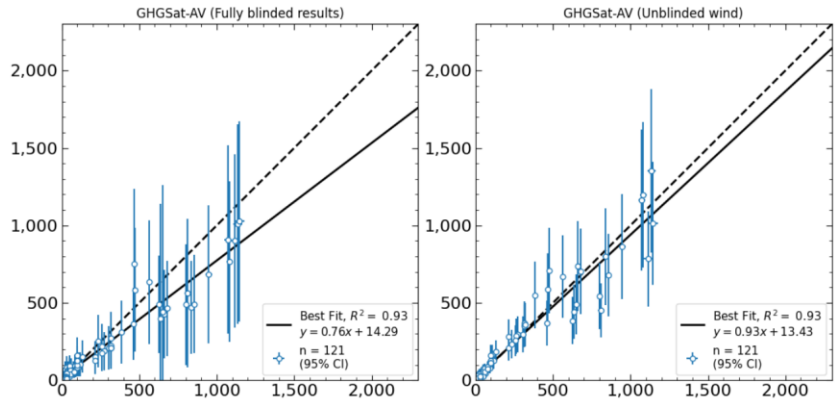

Figure 7. Comparison of the GHGSat emission rate measurement with ground truth at two stages of unblinding: (a) fully blinded, (b) locally measured wind speed unblinded. Markers and error bars represent individual measurements with the associated 95% confidence intervals. The full line is the linear regression, and the dashed line represents the ideal 1:1 ratio. Adapted from El-Abbadi et al.

4.4 Limitations

The measurement technique has limitations that are inherent to passive optical absorption spectroscopy, in that it requires the ground to receive direct sun illumination, as well as a clear line of sight between the instrument and the emission. In particular:

- Measurements during the night or in overcast weather are not possible. In partly cloudy conditions, obscured areas with insufficient signal will be flagged as invalid.
- Offshore measurements require a specific observing configuration to capture the direct sun glint reflection (predominantly specular reflection as opposed to diffuse reflection on land). At this moment, GHGSat only provides offshore measurements from its satellite platform (DATA.SAT), and not for aerial surveys (DATA.AIR).
- Underground leaks can be detected after they reach the surface.

Variations in wind speed affect the sensitivity and quantification accuracy, because the local concentration of methane from a leak is driven not only by the emission rate, but also the rate at which wind dissipates methane in the surrounding atmosphere. High wind speeds increase the detection threshold. On the other end, very low wind speed will favor detection of small leaks, but makes quantification more difficult as the relative uncertainty of the emission rate

is dominated by that of local wind. Both of those effects are accounted for in the methane quantification uncertainty that is delivered along with the source rate.

Retrieved methane concentration can be correlated with ground reflectance, resulting in some degree of systematic noise in the measurement. Homogenous and arid scene allows a better precision and lower detection threshold than mountain, farmland and urban scenes or areas with a complex topography.

5 REFERENCES

Aben, I., Hasekamp, O., Hartmann, W., 2007. Uncertainties in the space-based measurements of CO2 columns due to scattering in the Earth's atmosphere. J. Quant. Spectrosc. Radiat. Transf. 104, 450–459.

Chahine, M.T., Pagano, T.S., Aumann, H.H., Atlas, R., Barnet, C., Blaisdell, J., Chen, L., Divakarla, M., Fetzer, E.J., Goldberg, M., others, 2006. AIRS: Improving weather forecasting and providing new data on greenhouse gases. Bull. Am. Meteorol. Soc. 87, 911–926.

Chandrasekhar, S., 2013. Radiative transfer. Courier Corporation.

Clough, S.A., Shephard, M.W., Mlawer, E.J., Delamere, J.S., Iacono, M.J., Cady-Pereira, K., Boukabara, S., Brown, P.D., 2005. Atmospheric radiative transfer modeling: a summary of the AER codes. J. Quant. Spectrosc. Radiat. Transf. 91, 233–244.

Houweling, S., Hartmann, W., Aben, I., Schrijver, H., Skidmore, J., Roelofs, G.-J., Breon, F.-M., 2005. Evidence of systematic errors in SCIAMACHY-observed CO 2 due to aerosols. Atmospheric Chem. Phys. 5, 3003–3013.

Huang, Y., Natraj, V., Zeng, Z., Yung, Y.L., 2020. Quantifying the impact of aerosol scattering on the retrieval of methane from airborne remote sensing measurements. Atmospheric Meas. Tech. Discuss. 1–28.

Jervis, D., McKeever, J., Durak, B.O.A., Sloan, J.J., Gains, D., Varon, D.J., Ramier, A., Strupler, M., Tarrant, E., 2021. The GHGSat-D imaging spectrometer. Atmospheric Meas. Tech. 14, 2127–2140. https://doi.org/10.5194/amt-14-2127-2021

Kochanov, R. V, Gordon, I.E., Rothman, L.S., Wcisło, P., Hill, C., Wilzewski, J.S., 2016. HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data. J. Quant. Spectrosc. Radiat. Transf. 177, 15–30.

Rodgers, C.D., 2000. Inverse methods for atmospheric sounding: theory and practice. World scientific.

Roy, D.P., Wulder, M.A., Loveland, T.R., Woodcock, C.E., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., others, 2014. Landsat-8: Science and product vision for terrestrial global change research. Remote Sens. Environ. 145, 154–172.

Sloan, J.J., Durak, B., Gains, D., Ricci, F., McKeever, J., Lamoire, J., Sdao, M., Latendresse, V., Lavoie, J., Kruzelecky, R., 2016. Fabry-Perot interferometer based satellite detection of atmospheric trace gases.

Thorpe, A.K., Frankenberg, C., Roberts, D.A., 2014. Retrieval techniques for airborne imaging of methane concentrations using high spatial and moderate spectral resolution: application to AVIRIS. Atmospheric Meas. Tech. 7, 491–506.

Varon, D.J., Jacob, D.J., McKeever, J., Jervis, D., Durak, B.O.A., Xia, Y., Huang, Y., 2018. Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. Atmospheric Meas. Tech. 11, 5673–5686. https://doi.org/10.5194/amt-11-5673-2018

Varon, D.J., McKeever, J., Jervis, D., Maasakkers, J.D., Pandey, S., Houweling, S., Aben, I., Scarpelli, T., Jacob, D.J., 2019. Satellite discovery of anomalously large methane point sources from oil/gas production. Geophys. Res. Lett.