Evaluation of Emission Quantification Technologies

Final Report

February 7, 2024

Prepared for:

Petroleum Technology Alliance Canada

and Clean Resources Innovation Network

Prepared by:

Professor Kyle J. Daun, PhD., P. Eng., FASME
Professor Christiane Lemieux, PhD
Professor Audrey Béliveau, PhD
Daniel Blackmore
Nishant Narayanan
Augustine Wigle

University of Waterloo

Kevin Fritz

Arolytics, Inc.

Kirk Osadetz

CMC Research Institutes

Table of Contents

EXECUTIVE SUMMARY

In September 2021 the University of Waterloo (UW), Arolytics, Inc., and Carbon Management Canada (CMC) commenced a 3-year study to assess the potential of existing and emerging technologies for measuring methane emissions from upstream oil and gas sources, and develop uncertainties for these estimates.

Research commenced with a review of available technologies, which were categorized according to operating principle and applicability to various emission measurement scenarios. Manufacturer-specified performance characteristics were also summarized and tabulated. Based on these results, a subset of technologies was selected for detailed analysis: quantitative optical gas imaging (QOGI); truck-, drone-, and aircraft-based tunable diode laser-absorption spectroscopy (TDLAS); and aircraft-based near-wavelength and long-wavelength hyperspectral (HS) imaging.

This activity was followed by a more detailed assessment of the technologies. This was done in two ways: First, Monte Carlo simulations were performed using AroFEMP to calculate the expected reduction in methane emissions were the technologies incorporated into an alternative fugitive emissions management plan (alt-FEMP), relative to a standard OGI-based FEMP. Second, a laboratory-scale experimental study was carried out using QOGI cameras on a heated vent of methane to better understand the operating principles, characteristics, and capabilities of this technology.

The technologies were then deployed in three controlled-release field campaigns at CMC's field station in Newall County, Alberta[1]. The controlled releases were "semi-blind", meaning that the technology providers and operators knew the release locations, but not the release rates. A variety of measurement scenarios were considered, including stacks of various heights and an unlit flare; emissions from the top of a tank; and emissions from the side of a shed. The results highlighted how quantification accuracy was influenced by release rate, release scenario, and various environmental parameters (e.g., wind, ground temperature, and air temperature).

The final research phase is focused on developing techniques for quantifying the uncertainty attached to emission estimates. Uncertainty arises from a variety of sources, including measurement noise; uncertain model parameters (e.g., wind speed, air temperature); and structural errors caused by the approximations and simplifications made when deriving the measurement models. These errors introduce bias and variability in the estimates.

Two approaches have been developed to quantify uncertainty estimates, both based on Bayesian inference. The first approach is based on propagating measurement noise, model parameter uncertainty, and model error through the measurement equations. This approach focused on the truck based TLDAS system since the measurement equations (inverse Gaussian plume model) are straightforward and easy-to-use. In many cases, however, the model equations are highly complex and/or unavailable to a third-party operator. For this scenario, a second technique was developed that uses a set of controlled release measurements to develop an empirical likelihood model, which may be inverted to obtain uncertainties from subsequent measurements using the same technology. Finally, these models are being incorporated into a Monte Carlo simulation of an alt-FEMP, to assess the uncertainty attached to performance estimates for various methane emission management plans.

[1] While the Bridger GML was included in the analysis, they were unable to participate in field trials.

1. INTRODUCTION

1.1 Purpose and scope

As part of its strategy to address climate change, the Government of Canada has pledged to reduce methane emissions from the oil and gas sector by 40-45% below 2012 levels by 2025 [1, 2], with a further goal of reducing methane emissions by 75% by 2030 [3]. In December 2023 the Government of Canada announced further regulatory amendments that include enhanced screening and repair timelines that depend on leak rates and annual facility inspections conducted by an auditor [4].

To fulfill these obligations, Canada's oil and gas producers need tools that can both detect and quantify methane emissions. Regulators need quantification tools to assess compliance of industry to these regulations. Climate modelers need the data from these tools to understand how emissions from Canada's oil and gas industry contribute to climate change, information that provincial governments and the Federal government use to inform policy needed to fulfil Canada's international treaty obligations and to avoid the worst outcomes of climate change.

A diverse suite of candidate methane emission quantification tools is available to Canada's oil and gas industry. There is unlikely to be a "one-size-fits-all" solution: techniques designed to measure localized, persistent fugitive leaks from valves or gaskets may not be suitable for diffuse and intermittent emissions vented from storage tanks, surface casing vents, and CHOPS (cold heavy oil production with sand) wells. Persistent emissions may be identified reliably using periodic surveys, while highly variable or intermittent emissions may require continuous monitoring. Recent advancements in optoelectronic hardware that include mid-wavelength infrared (MWIR) cameras, tunable diode laser absorption spectroscopy (TDLAS) in a range of configurations, and hyperspectral imaging from ground-based, aerial, and orbital platforms have augmented and disrupted the field of traditional approaches based on extractive sampling (e.g., the Bacharach HI FLOW® sampler) although their accuracy, precision, and best practices under various measurement scenarios are still being refined.

Choosing the "right tool for the job" depends largely on the uncertainty associated with the methane emission estimates. This is particularly the case in a regulatory context. For example: How certain is it that an operator is complying with emissions? Is it practical to impose a limit on emissions that cannot be measured with reasonable accuracy? What is the most cost-effective technology to deploy for a given measurement scenario? To make these decisions, operators and regulators need to understand the uncertainty with which emissions may be quantified. Climate modelers and policy makers also need to understand the uncertainties attached to reported emission inventories to draft effective regulations that safeguard the environment without unduly penalizing oil and gas producers.

With these goals in mind, the Petroleum Technology Alliance Canada (PTAC) and Clean Resources Innovation Network (CRIN) engaged the University of Waterloo (UW), Arolytics, Inc. (Arolytics), and Carbon Management Canada, Inc. (CMC) to undertake a three-year research program entitled "Evaluation of Emission Quantification Technologies". The program is funded by the Alberta Upstream Petroleum Research Fund (AUPRF) through PTAC, the Government of Canada's Strategic Innovation Fund through CRIN, and the Natural Sciences and Engineering Research Council (NSERC) via the Alliance program.

The UW/Arolytics/CMC team brings together the diverse and complementary range of expertise needed to fulfill these objectives. The UW team is led by Professor Kyle Daun from the

Department of Mechanical and Mechatronics Engineering, an expert in optical gas imaging, laser-based diagnostics, remote sensing, and uncertainty quantification. Daun is joined by Professors Christiane Lemieux and Audrey Béliveau from the Department of Statistics and Actuarial Science. Lemieux's research interests focus on Monte Carlo simulations, while Béliveau studies Bayesian statistics and uncertainty quantification. Arolytics's participation is coordinated by Kevin Fritz, an expert in methane emissions quantification and FEMPs; Fritz also has extensive experience in methane emissions field measurements. Finally, CMC's participation is managed by Kirk Osadetz, a geoscientist and expert in methane emissions from Canada's oil and gas industry. Osadetz is CMC's Programs Development Manager and also manages CMC's Field Research Station, where candidate quantification technologies will be evaluated under industry-relevant conditions.

Research activities were organized into four project phases:

Phase 1: Quantification tool review interim summary table

Phase 2: Recommendation of viable quantification tools

Phase 3: Field campaign design and execution

Phase 4: Data analysis and summary

These research activities were guided with input and oversight from PTAC, CRIN, and members of the Industrial Steering Committee (ISC), a subcommittee of PTAC's Air Research Planning Committee (ARPC).

This final report summarizes the progress made in these four phases, discusses ongoing work, and points to future research directions. A particular focus is placed on activities that took place after the second interim report, including a third field trial that focused on airborne long-wavelength infrared hyperspectral imaging, and preliminary statistical analysis that took place in 2023. Some of these data analysis activities are ongoing and will be completed after the end date of this report.

2. PHASE 1: QUANTIFICATION TOOL REVIEW INTERIM SUMMARY TABLE

Research commenced in September 2021 with Phase 1. The outcome of this phase was a comprehensive survey of 23 candidate emission quantification technologies, along with an assessment of their capabilities, their technological readiness level, and their applicability to measuring emissions at the component, facility, and site level.

In general, each technology derives an emission measurement, Q (e.g., kg/hr) by combining a methane concentration measurement with an advection/transport model. In this regard, methane emission quantification technologies may be organized according to the nature of the data they collect, which largely determines their suitability for various emission scenarios and how uncertainty estimates may be derived. This is shown schematically in Appendix A.

Some technologies are based on point-concentration measurements, like short-path TDLAS, which provide a local concentration of methane, c (e.g., kg/m^3). These measurements are combined with a plume transport model (e.g., the Gaussian plume model [5, 6] or backwards Lagrangian stochastic model [7]) that incorporates local anemometry measurements to obtain an emission rate estimate. Other techniques are based on open-path measurements over long distances, like dual comb frequency spectroscopy [8], in which case the outcome is a column density, σ, integrated along the optical path (ppm·m or kg/m^2); this data may also be combined with a wind speed measurement to obtain an emission rate estimate.

A third category of imaging technology generates a 2D map of column densities; an example MWIR quantitative optical gas imaging (QOGI). In some cases, the emission rates are inferred by combining the column density map with a win measurement, while, for sufficiently high frame rates, a 2D intensity-weighted velocity field may be inferred by feature tracking between successive images. Finally, a fourth class of technology is used to measure emissions from confined flows (e.g., ducts) using gauges. The summary table and diagrams are in Appendix A of this report.

Based on this analysis, a subset of these technologies was identified for further analysis during subsequent project phases[2]:

1. QOGI
2. Truck-, drone- and aircraft-deployed TDLAS
3. Aerial near-wavelength infrared (NIR) hyperspectral imaging
4. Aerial long-wavelength infrared (LWIR) hyperspectral imaging.

A brief description of each technique is provided in the following subsections:

2.1 Quantitative Optical Gas Imaging (QOGI) using MWIR broadband cameras

This technique is based on imaging the thermal radiation emitted by a hydrocarbon plume hotter than its background (white plume) or the background thermal radiation that is absorbed by a plume colder than its background (black plume). The spectral intensity from the scene enters the camera aperture and is then imaged through a broadband filter and onto a focal plane array (FPA). The broadband "cold filter" and FPA are located behind a "cold shield" and maintained at cryogenic temperatures. This is done to enhance the thermoelectric conversion efficiency of the sensor and a

[2] Aerial TDLAS could not be included in the field trials for logistical reasons, but data from previous published field campaigns has been analysed through Phase 4 activities.

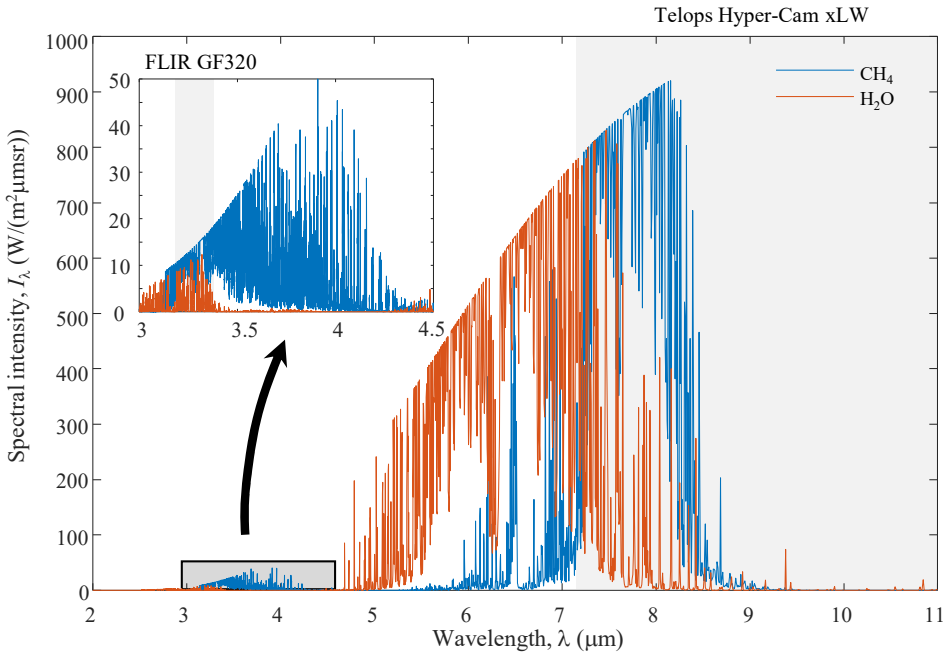

Figure 1: Emission spectrum for methane and water vapor at 300 K (a "white plume"). (CH_4: L = 20 cm, χ = 1, H_2O: L = 200 cm, χ =0.1.) Broadband QOGI measurements are constrained to the MWIR because of the water lines over the LWIR.

void blackbody radiation from the sensor, camera chassis, and the filter, which would otherwise contaminate the signal and preclude quantitative interpretation of the image.

The cold filter defines the measurement spectrum; a typical choice is 3.2-3.4 μm (FLIR GFx320), which is aligned with a primary vibrational-rotational absorption band corresponding to the C-H stretching mode for hydrocarbons (~3.34 μm, Fig. 1). The brightness of each pixel corresponds to the spectral intensity incident on the pixel, integrated over this wavelength band.

The column density of gas along each line-of-sight, σ (kg/m^2), is then inferred from the pixel intensity by inverting a spectroscopic model. This requires knowledge of: (1) the plume composition (usually taken to be methane); and, in the case of white plumes, (2) the plume and ambient temperatures. The plume temperature is usually taken to be that of the ambient air temperature, following the assumption that gas emerging from a leak will quickly reach thermal equilibrium with the surrounding air. In the case of gases other than methane, a conversion chart may be applied to correct the column densities.

The MWIR measurement spectrum is not the optimal choice based solely on the sensitivity of the pixel intensity to methane column density. Hydrocarbons have a secondary vibrational-rotational band at ~7.6 μm (LWIR) which is far more intense than the 3.34 μm feature because matter at ambient temperatures emits more thermal radiation at wavelengths closer to 10 μm. However, while broadband LWIR cameras may be used to visualize methane, quantitative analysis is precluded by significant emission features from ambient H_2O, as seen in Figure 1.

The 2D column density map is then combined with an advection model to obtain an overall mass flux according to

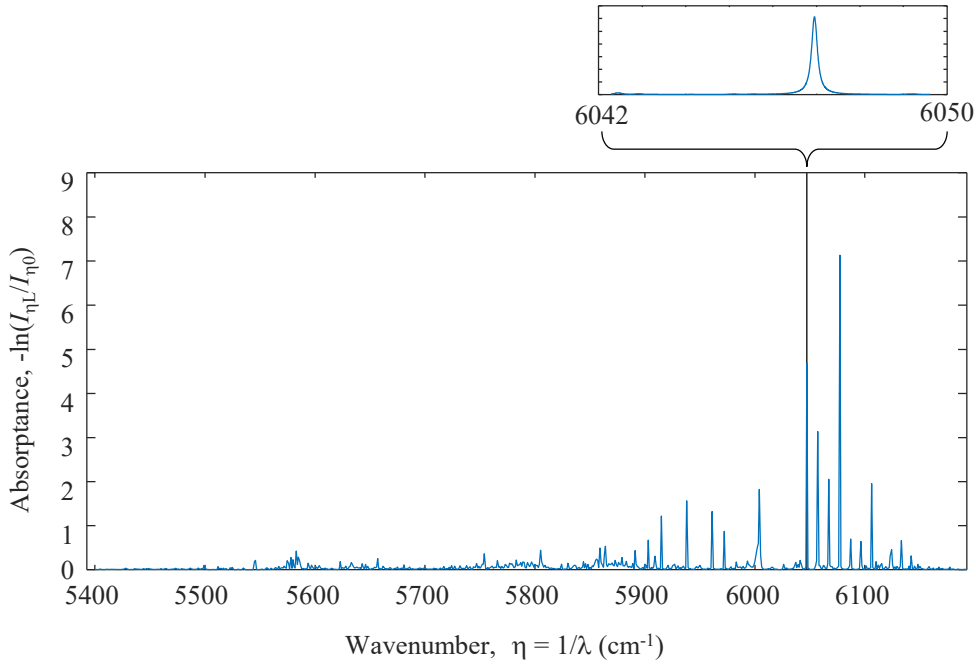

Figure 2: Near-Infrared (NIR) absorption spectrum for CH_4 used for TDLAS. Boreal Laser's GasFinder 3 instrument uses wavelength-modulated spectroscopy to infer the concentration based on an absorption line at 1654 nm (6046 cm^{-1}). This part of the spectrum is also the basis for the GHGSat instrument (1630-1655nm) based on absorbed sunlight.

$$Q = \int_0^{\Theta} \sigma(\theta)\left[\mathbf{u}(\theta)\cdot\mathbf{n}(\theta)\right]d\theta \tag{1}$$

where $\sigma(\theta)$ is the column density at a position θ over the control surface, **u** is the 2D projected velocity, and **n** is the 2D normal vector. The advection model is inferred from the apparent motion of the plume in successive images. Unlike the spectroscopic model, the advection model is a proprietary and largely "black box" aspect of commercial QOGI software. The advection calculation may be affected by artifacts like "pooling" and background motion due to, e.g. moving clouds or vegetation; these effects were explored in Phase 3 activities.

2.2 Truck-, aircraft-, and drone-based TDLAS

Tunable diode laser absorption spectroscopy is a low-cost and reliable way to measure the concentration of a particular gas, e.g., methane. A distinguishing feature of tunable diode lasers is that they may be "tuned" across a narrow range of wavelengths by adjusting the supply current; in the case of gas sensing, the wavelengths capture one or more spectral lines corresponding to defined transitions between quantized vibrational-rotational states that act like a "fingerprint" for the species of interest (Figure 2). In the case of a homogeneous gas over a path length L, the gas concentration may be inferred from the Beer-Lambert Law,

$$I_{\eta L} = I_{\eta 0}\exp\left(-\kappa_{\eta} L\right) \tag{2}$$

where η is the wavenumber (cm^{-1}, the inverse of wavelength), $I_{\eta 0}$ and $I_{\eta L}$ are the incident and transmitted laser intensity, and κ_{η} is the spectral absorption coefficient, which is proportional to the number density (molecules/m^3) of the species of interest. Inverting Eq. (2) results in a column

density, e.g. ppm·m or kg/m^2, which is proportional to the number of molecules along the path length.

This approach is called *direct absorption spectroscopy* (DAS) and is the simplest implementation of TDLAS. In *wavelength modulated spectroscopy* (WMS), the wavelength is tuned rapidly over a single line by harmonically varying the supply current, and the absorption coefficient/column density is then inferred from the AC and DC components of the transmitted signal; WMS provides superior signal-to-noise ratio (and hence a lower detection threshold) compared to DAS and is therefore the most common TDLAS approach for methane detection.

TDLAS may be deployed in a wide range of configurations to detect methane. For example, Bridger Photonics' Gas Mapping LiDAR system consists of a TDLAS laser operating in WMS mode mounted on a movable gimbal on an aircraft that overflies the methane plume [9]. The ground-reflected portion of the laser is detected by a sensor onboard the aircraft. The ground reflectivity and column density may be inferred from the backscattered signal. Column densities obtained at different view angles are then tomographically reconstructed to form a 3D concentration map of the plume. Finally, the concentration profile is combined with an advection model based on modeled (e.g., Meteoblue) or locally measured wind speeds to find mass flow rates. In this configuration, the laser path is defined by the laser, the reflected light, and the receiver. This system's main drawback is that it only works if enough laser light is reflected from the ground. Consequently, the Bridger system does not perform well on wet or snow-covered surfaces, which absorb most incident radiation.

Alternatively, the TDLAS sensor may be configured to measure the gas concentration over an enclosed path. In this approach, the volume is often terminated with a mirror to extend the path length and increase the instrument's sensitivity. In this configuration, the TDLAS instrument provides a time-resolved point concentration measurement (e.g., ppm, kg/m^3). A set of concentration measurements made by moving the sensor through the plume is then combined with an advection model to obtain a mass flow rate.

Two common approaches for traversing the sensor through the plume are: truck based TDLAS and airborne TDLAS (e.g., drone mounted.) In the case of the drone-mounted TDLAS sensor deployed in the first field campaign, the drone flies a descending helical path around the methane source (Figure 3); the helical path defines a control surface, A, which may then be combined with

Figure 3: Example helical path flown by the TDLAS-equipped drone in the first field campaign.

a wind measurement to determine the total release rate. In one approach, called the *cylindrical flux plane* approach [10], the emission rate is estimated from

$$Q = \int_A c(\mathbf{r})\mathbf{u}(\mathbf{r})\cdot\mathbf{n}(\mathbf{r})\,dA \tag{3}$$

where **r** is the position vector, **u**(**r**) and c(**r**) are the wind velocity vector and concentration at **r**, and **n**(**r**) is the unit normal vector. Alternatively, the emission rate may be found from an inverse Gaussian plume model, which is discussed below. The wind velocity may be found from anemometry (e.g., mounted on the drone) or based on the power required for the drone to maintain its position in a crosswind. This technique, along with other approaches for inferring methane release rates from drone measurements, are summarized in Ref. [10].

In the case of a truck mounted TDLAS system (Figure 4), concentration measurements are made at various plume transects. Inferring the emission source that caused these concentrations amounts to solving an ill-posed inverse problem, since an infinite set of candidate sources could account for the observed concentration measurements. Accordingly, it is necessary to impose additional information that connects the source to the concentration measurements; the simplest approach is to assume a Gaussian plume model [5], which assumes that the time-averaged concentration field from a steady release is given by

$$c(\mathbf{r}) = \frac{Q}{2\pi u\,\sigma_x\sigma_y}\left\{\exp\left[-\frac{(z-h)^2}{2\sigma_z^2}\right]+\exp\left[-\frac{(z+h)^2}{2\sigma_z^2}\right]\right\}\exp\left(-\frac{y^2}{2\sigma_z^2}\right) \tag{4}$$

where **r** = (x, y, z) is the position vector, u is the wind speed, h is the effective release height, and σ_y and σ_z are coefficients related to atmospheric conditions. In Eq. (4) the x-axis is aligned in the downwind direction. Equation (4) is rearranged to express Q in terms of the concentration distribution, which are approximated by the finite set of truck measurements, $\{c\}$. Other choices for inverting the concentration measurements are available, including the backwards Lagrangian stochastic approach [7].

The main drawback of both the drone and truck mounted TDLAS systems is that they assume that the release and wind are time-invariant, while, in reality, both change with time. The outcome of the first field trial showed the emission rates derived from the drone-based measurements were particularly susceptible to the errors introduced by this approximation.

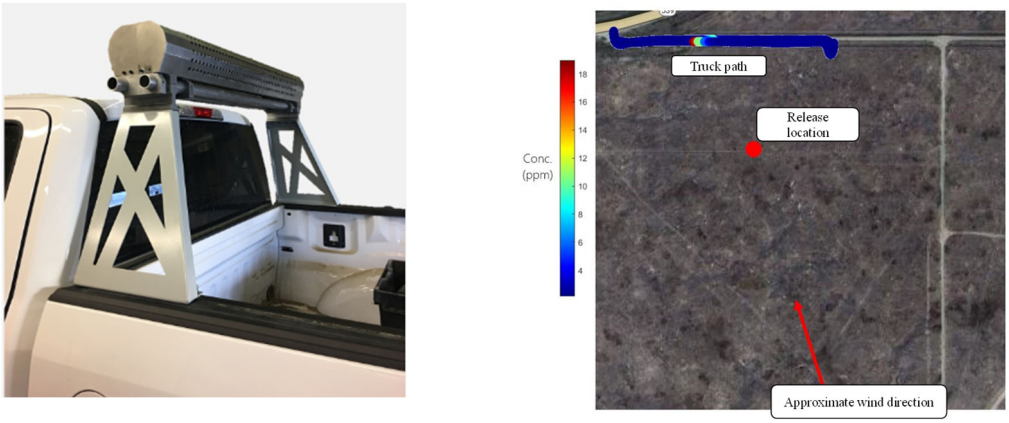

Figure 4: Truck-mounted TDLAS system used in the first two field campaigns.

2.3 Aerial NIR and LWIR hyperspectral imaging

As noted in Sec. 2.1., a key limitation of broadband QOGI cameras is that the detected spectral intensity is integrated over the measurement spectrum. Consequently, the camera can only be used to measure pure mixtures (or mixtures of known composition) and, in the case of "white plumes" the plume temperature must be specified. These limitations may be overcome through hyperspectral (HS) imaging.

In contrast to a broadband QOGI camera, which generates a single image aligned with the 3.34 λm V-R band of methane, HS cameras generate "data cubes" containing hundreds or thousands of images, each at a distinct wavelength. In many cases, this is done through interferometry; this type of HS camera is called an imaging Fourier transform spectrometer (IFTS). In this approach, the image entering the camera aperture is split. The two split images then travel slightly different distances and are then recombined; due to the different distances travelled, the recombined images are brighter or darker than the original image due to wave interference effects. Many images are collected for each optical path difference, and the resulting interferogram may be Fourier-transformed to produce the data cube.

Two types of IFTSs are analyzed in this study, each of which works over a distinct measurement spectrum and according to a different measurement principle. GHGSat has adopted their orbital package [11] into an airborne system; it consists of a Fabry-Perot IFTS that operates between 1630-1655 nm, over the NIR spectrum. Thermal emission from the gas and ground is negligible over this wavelength range; instead, the camera images sunlight transmitted through the atmosphere, reflected from the ground, and transmitted back to the camera. The methane column density is inferred from the attenuation of the transmitted light via a multilayer spectroscopic model, and then combined with an advection model to find the emission rate.

There are several schemes that may be used to convert the column density map into an emission rate. In the *integrated mass enhancement* approach the emission rate is approximated by [11, 12]

$$Q = \frac{u_{\text{eff}}}{L_{\text{eff}}} \int_A \sigma(x,y)\, dA \approx \frac{u_{\text{eff}}}{L_{\text{eff}}} \sum_{i=1}^{N} \sigma_i \Delta A_i \tag{5}$$

where u_{eff} is the wind speed and L_{eff} is a characteristic plume length, taken to be the square root of the plume area, found using a Boolean plume mask that distinguishes plume pixels from background pixels. A second approach, called the *cross-sectional method*, infers Q by measuring the mass flux across M column transects at various distances downwind from the source and parallel to the wind:

$$Q \approx \frac{u_{\text{eff}}}{M} \sum_{j=1}^{M} \int \sigma(x_j, y)\, dy \tag{6}$$

where x and y are oriented in the downwind and crosswind directions, respectively.

Telops and LiDAR Systems, Incorporated (LSI) also use a downwards looking IFTS (Telops Hyper-Cam xLW Airborne Mini), but one that operates at much longer wavelengths (7.4-12.5 μm). Instead of relying on transmitted sunlight, the camera exploits the thermal contrast between the gas and the background to infer column densities. As already noted, thermal emission over the LWIR spectrum is significantly greater compared to the MWIR spectrum used by broadband cameras; the spectral resolution of the IFTS makes it possible to disentangle contributions of methane and H_2O, which cannot be done using broadband cameras. The spectral resolution of the

camera also enables direct calculation of the ground temperature, which is a necessary boundary condition for the spectroscopic model.

In contrast to the Bridger airborne GFL system, which uses a laser to interrogate the plume, the GHGSat and Telops systems are entirely passive, but generally have lower detection thresholds. The GHGSat system is sensitive to ground reflectance and cloud cover, which affects the incident and reflected sunlight. The sensitivity of the Telops system depends on the thermal contrast between the plume and the background surface, among other factors.

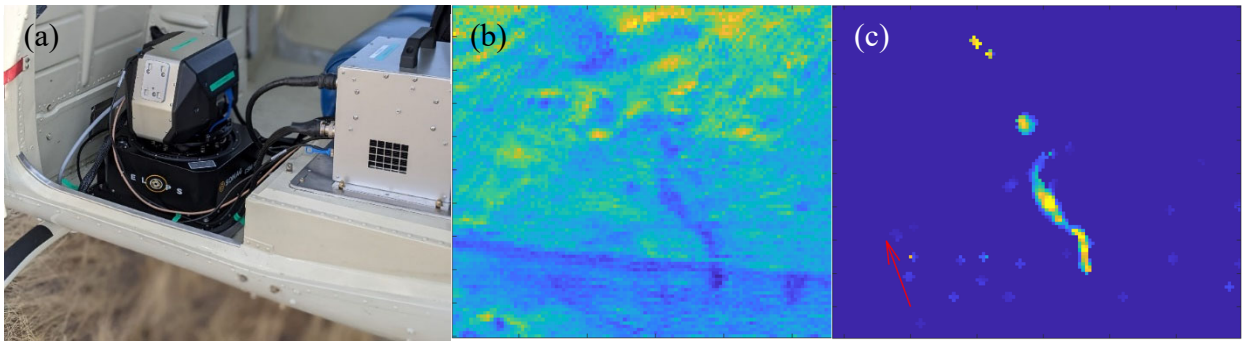

Figure 5: (a) Hyper-Cam xLW Airborne Mini; (b) spectral intensity map at 1305 cm^{-1}, which is converted into a pixel mass map (colors indicate the grams of CH_4 contained within each pixel).

3. PHASE 2: ANALYSIS OF CANDIDATE EMISSION QUANTIFICATION TECHNOLOGIES

The second project phase focused on an in-depth analysis of a subset of the technologies identified in Phase 1. This activity had two components: Phase 2a, led by Arolytics, examined how these technologies would perform when incorporated into an alternative fugitive emissions management program (alt-FEMP) relative to an OGI-based FEMP. Phase 2b, led by UW, focused on understanding the operating parameters of QOGI within a laboratory setting, and the model errors associated with the Gaussian plume modeling used to infer emission rates from the truck based TDLAS measurements using a computational fluid dynamics (CFD) simulation.

3.1 Phase 2a: Recommendation of the most viable quantification tools

Arolytics personnel used Arolytics' proprietary AroFEMP Monte Carlo simulation code to predict the performances of three technologies identified in Phase 1 were they to be incorporated into an alternative FEMP (alt-FEMP). These technologies were: airborne TDLAS (Bridger GFL); airborne NIR hyperspectral imaging (GHGSat); and truck based TDLAS (Boreal Laser). The performance of these technologies was assessed in terms of the reduction that would be realized through alt-FEMS based on these technologies compared to what would be obtained from the standard OGI-based FEMPs using the FLIR GF320/GL320 system. OGI survey time was derived from general time estimates reported by Canadian OGI service providers and the OGI detection capabilities were taken from the study conducted by Zimmerle et al. [13]. The detection capabilities and time information for the remaining technologies were obtained from each service provider.

This procedure considered various scenarios in which operators would use candidate technologies to detect and quantify methane emissions, and then action repairs over a calendar year based on the outcome of these surveys. The overall benefit of a given technology, in terms of the reduction in emissions (e.g., kg/year), were evaluated through a Monte Carlo simulation. Emission scenarios were sampled randomly from probability densities that represent operational conditions, and emissions were detected and quantified with probabilities derived from manufacturer-specified characteristics of the quantification technologies or previous field trials identified in Phase 1. Repeating this procedure multiple times amounted to an integration over the probability densities, with the outcome of the expected reduction in emissions.

Representative fugitive emissions probability profiles were derived from the FEMP-EA study dataset [14]. Modelling was performed using an anonymous set of infrastructures, drawn from the Petrinex database, representative of a typical Alberta upstream oil and gas producer. The criteria used to select the representative infrastructure were facility count, facility subtype distribution, and the ratio of facilities requiring triannual and annual surveys (per AER Directive 060 Table 4 [15]). A subset of 562 facilities was then chosen such that each of the criteria were equivalent to those of Petrinex's Alberta infrastructure. The FEMPs were modelled for a two-year timespan, which is the typical length of an alt-FEMP.

The OGI system was modelled as one would conduct the default FEMP [15] and the remaining technologies were modelled each in their own alt-FEMP. The frequency of screenings and the fraction of all facilities that received follow-up surveys were adjusted to design alt-FEMPs that were equivalent to or better than the default FEMP in fugitive emissions reductions.

The model results showed that all three alternative technologies could be used to design alt-FEMPs that are equivalent or superior to the default FEMP. These alt-FEMPs employed 2-3 screening

surveys each year with OGI follow-ups being conducted at 30% of all screened facilities, selecting those with the largest emissions. The alt-FEMP design and performance details are summarized in Table 1.

Table 1: Simulated performance of Alt-FEMPs relative to default FEMP (QOGI).

Technology	Conditions	Relative reduction
Airborne TDLAS (Bridger GFL)	Two screenings per year, 30% follow-up	8%
NIR HS imaging (GHGSat)	Three screenings per year, 30% follow-up	11%
Truck-based TDLAS (Boreal)	Three screenings per year, 30% follow-up	15%

These results demonstrate that these technologies can all be feasibly and practically employed as the screening technologies in alt-FEMPs that achieve fugitive emission reductions equivalent to a default FEMP for a "typical" Alberta oil and gas producer. In this modelling, each of the alt-FEMPs was able to perform better than the default FEMP, indicating that the alt-FEMPs could potentially achieve fugitive emissions reductions equivalent to the default FEMP using less follow-up or fewer screenings than were modelled.

It is important to note that while this modelling was conducted for a "typical" Alberta oil and gas producer, the infrastructure of a given oil and gas producer and the properties of that infrastructure can vary considerably between producers. In addition, each alternative technology has unique capabilities which may be more compatible to certain types of infrastructure and properties compared to others. Accordingly, the potential for successful implementation of an alt-FEMP is case dependent and should be evaluated for each unique scenario.

Finally, and crucially, the current Monte Carlo analysis employs a manufacturer-specified probability of detection, followed by prioritization based on emission rate and a probability of intervention to determine emission repairs and reductions. This analysis does not yet consider the emission rate measurement uncertainty, information that is relevant for assessing the compliance to regulations, allocating resources to reduce emissions, and for designing the alt-FEMPs themselves. The research carried out under this grant is focused on developing defensible uncertainties for these emission estimates, which may be incorporated easily into a Monte Carlo simulation to deliver an overall uncertainty in the emissions quantification and reductions achieved through a given alt-FEMP. This ongoing work is described in Sec. 5.3.

3.2 Phase 2b: Laboratory trials and development of digital twin system

In parallel, the UW team conducted laboratory-scale analysis and numerical simulations to further investigate the capabilities of the most promising technologies identified in Phase 1. Specifically, Phase 2b activities focused on developing and quantifying the performance of spectroscopic and optical flow models for QOGI, and the derivation of an inverse Gaussian plume model (IGM) for interpreting truck based TDLAS measurements.

Research into QOGI technology was led by Michael Nagorski, a MASc candidate under Daun. Laboratory measurements were carried out using a FLIR GFx320 camera and QL320 tablet system, as well as a four-channel Telops multispectral (MS) channel provided by CMC. Nagorski developed an "in-house" version of the QL320 tablet based in MATLAB® and validated its performance through simulated measurements generated using a CFD Large Eddy Simulation of

a methane plume, as well as plumes generated with a heated vent apparatus. The simulated measurements were used to validate the spectroscopic model used to infer the species column densities, and the optical flow algorithm used to obtain the velocity field [16, 17]. This information was particularly useful when interpreting the QOGI performance study carried out in the second field trial (Sec. 4.2).

In the heated vent experiments, infrared images captured using the GFx320 camera were interpreted using the QL320 tablet and UW's in-house code. In most cases, UW's code matched or exceeded the performance of the QL320 tablet with regards to the overall mass flow rate when using a bulk velocity estimated from the images, but the framerate of the GFx320 was too low for the in-house optical flow algorithm to provide robust velocity fields. A subsequent meeting with Providence Photonics, the originator of the QL320 system, held in late 2022, provided further insights into the proprietary aspects of this technology [18]. While the exact numerical methodology was not disclosed, it is believed that the algorithm is built upon one of the optical flow algorithms offered by the open-source computer vision library OpenCV [19], which are similar to those tested in-house.

While the Telops camera could visualize gas plumes at a high frame rate, quantitative emission estimates were not possible due to thermal emission from the uncooled filters. Further details of both the GFx320/QL320 and Telops test results were provided in the first interim report [20].

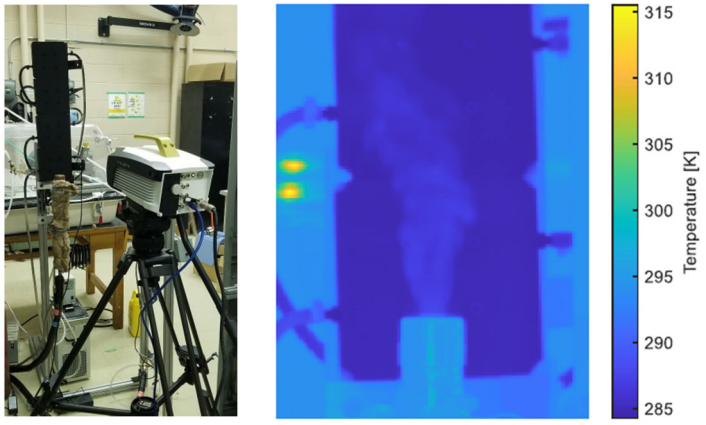

Figure 6: The heated plate apparatus at UW is used to assess the performance of QOGI cameras, including the FLIR GFx320 and the Telops Hyper-Cam.

Inverse Gaussian plume model (IGM) research is led by MASc student Daniel Blackmore with assistance from Paule Lapeyre, a postdoctoral fellow. The objective of this research is to quantify the ***model error*** inherent to the IGM. Model error is introduced by the simplifications and assumptions needed to derive Eq. (4), and, in most cases, is the main source of uncertainty when inferring emission rates from direct concentration and wind measurements. In the case of the IGM, the plume is assumed to be stationary (meaning that the wind is perfectly steady) and the plume concentration measurements are taken to be time averaged. In reality, both wind speed and direction fluctuate, and the methane concentration is measured near-instantaneously as the vehicle transects the plume. There are additional uncertainties introduced by the model parameters, including the wind speed, U, and plume spread constants, $\{\sigma_x, \sigma_y, \sigma_z\}$, which are imperfectly known.

To understand these errors and their effect on overall methane emission quantification uncertainty, an *in-silico* experiment was conducted using a computational fluid dynamics-large eddy simulation (CFD-LES). Simulations were carried out by Professor Jean-Pierre Hickey at UW using OpenFOAM. This ongoing analysis is described in Sec. 5.1 of the report.

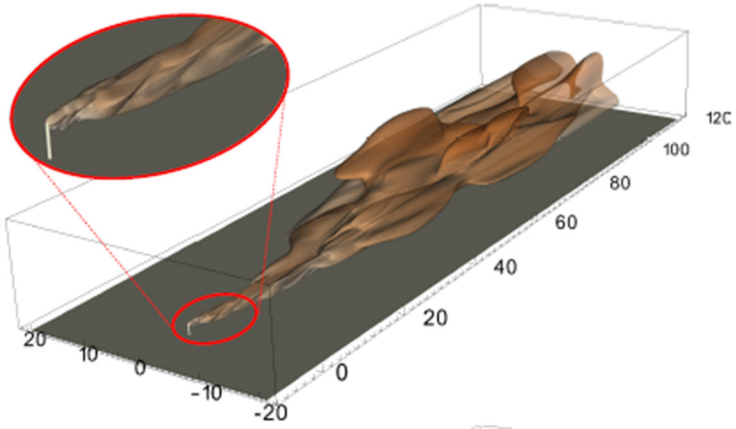

Figure 7: CFD-LES simulation of a methane plume, which is used to investigate model error involved in the truck-based TDLAS measurements.

4. PHASE 3: EXECUTION OF THE FIELD CAMPAIGNS

The main objective of this phase was to quantify the accuracy, precision, and uncertainty of emission estimates obtained from the subset of technologies identified in Phase 1. The project plan called for two field campaigns to be executed at the CMC Field Research Station (FRS) near Brooks, Alberta. The two planned field campaigns were held between April 20-26 and September 25-October 1, 2022, and were led by Kirk Osadetz, CMC's Programs Development Manager. The technologies deployed, and technology providers who participated in each field campaign are summarized in Table 2 and included: two QOGI systems operated by three different providers; a truck-mounted TDLAS system (Boreal); a drone-mounted TDLAS system (SAIT); an airborne NIR HS system (GHGSat); and an airborne LWIR HS system (Telops/LSI).

Unfortunately, the Telops/LSI system was unavailable for the first field campaign and was beset by technical issues during the second campaign. Accordingly, a third field campaign was held between September 19-25, 2023, that focused exclusively on the LWIR HS system. This campaign was funded by the Alberta Methane Emissions Program (AMEP) and the NSERC Alliance Missions Methane project.

The outcomes of the first two field campaigns have been presented in detail in the first and second interim reports [20, 21], respectively, so they are summarized here only briefly. The results of the third field campaign are presented in greater detail.

Table 2: Technologies and technology providers deployed at each field campaign. Numbers indicate the number of measurement events for each technology. Dashes denote no participation.

Technology/Provider	1st Campaign	2nd Campaign	3rd Campaign
QOGI (AGAT)	121	-	-
QOGI (Montrose)	-	97	-
QOGI (SAIT/UW)	15	116	-
TDLAS Truck (Boreal)	152	178	-
TDLAS Drone (SAIT)	15	-	-
Airborne NIR HS (GHGSat)	80	44	-
Airborne LWIR HS (LSI/Telops)	-	46	204[1]

[1]Multiple data cubes collected for each pass, multiple passes per release condition gives 600 passes and 17,000 independent estimates

4.1 Equipment, scenarios, and experimental procedures

Carbon Management Canada's FRS, located in Newall County AB (50°27'11.6"N 112°07'06.8"W), consists of a potable ATCO trailer, which serves as an operations center, meteorological instrumentation, and equipment for generating gas releases under industrially relevant scenarios, along with additional infrastructure focused on carbon sequestration research. The facility is in flat, unobstructed terrain, and is located immediately west of a large reservoir and south of Highway 539.

4.1.1 Meteorological equipment

The facility is equipped with instrumentation to measure wind speed, wind direction, ambient temperature, and dew point. Most weather parameters, especially wind speed and direction, were measured with multiple redundant systems. Sensors include a Davis WeatherLink Pro+ weather station affixed to the ATCO trailer that contains a cup-and-ball anemometer for wind speed; it also provides temperature, humidity, and solar irradiation data. Other meteorological data comes from a portable 81000-L, RM Young UVW 3D ultrasonic anemometer with a CR3000 measurement

and control datalogger set up next to the release locations, and a second 3D ultrasonic anemometer from the University of Colorado located next to the FRS ATCO trailer. A third set of wind data comes from an ultrasonic anemometer operated by Boreal laser during their truck mounted TDLAS measurements. Local meteorological measurements were also compared against predictions from Meteoblue.

The facility is surrounded by several feedlots, oil and gas batteries, and water wells, all of which are potential methane sources. For this reason, ambient methane was monitored continuously using a Picarro cavity ringdown spectrometer, located at the ATCO trailer, throughout the first two field trials. The ambient methane was found to be close to expected ambient levels throughout all testing except for a brief period during one day of the first field campaign when releases were conducted directly upstream of the ATCO trailer. Methane concentrations associated with the controlled releases far exceed this level, so the impact of ambient methane on the tests may be safely neglected.

4.1.2 Release rates, scenarios, and procedures

Liquefied natural gas was obtained from a commercial distributor in Lethbridge and brought to site in compressed gas cylinders within a trailer. A gas sample from the first field trial was taken offsite and analyzed to reveal a composition of 94.2% methane by volume, with the balance being ethane (approx. 3.4%), propane (1.1%) and 1.3% other minor components, predominantly N_2 and O_2 in atmospheric abundances. An assay on gas from the third field trial revealed 91% methane, with the balance being ethane (approx. 6%), propane (<1%), and similar minor components. Therefore, in general, the natural gas consists primarily of methane (90%).

The gas was discharged from the tanks through a regulator; as this occurs, the gas temperature drops dramatically due to the Joule-Thomson effect. The gas was then flowed through a conditioner/heat exchanger, shown in Figure 8, which raises the gas temperature to match that of the ambient air temperature. This was confirmed throughout the tests by measuring the gas temperature with a thermocouple as it was discharged to the atmosphere.

The gas was then flowed through an Alicat Model MCR-2000SLPM -D-PAR mass flow controller, before finally being vented out of the stack, tank, or shed, depending on the release type being investigated. The flow rate was inspected and logged digitally during the releases, revealing very little variation in the flow rate.

Releases ranged from 0.25 kg/h to 80 kg/h, along with some "null releases" to test the susceptibility of the quantification to false positives. The minimum release rates for each technology were chosen based on the corresponding minimum detection threshold (MDL). For QOGI, truck based TDLAS, and airborne LWIR HS imaging, a minimum non-zero release of 0.25 kg/hr was chosen, while the minimum non-zero release of the airborne NIR HS system was 5 kg/hr based on conversations with the provider. Release rates were quasi-randomized to avoid any consistent trend in release rate; this was done to both disguise the release rate from the provider and avoid any unintentional covariance between release rate and wind speed.

In the case of the third field trial, releases were programmed with guidance from the Telops minimum detection limit (MDL) tool (discussed in Sec. 4.3.2) to account for how changing wind speed and thermal contrast between the ground and air may influence the plume detection. This was done for two reasons: first, to "push" the diagnostic, while at the same time avoiding conducting a large number of releases that would be unlikely to be detected; and second, to test the predictive capability of the MDL tool.

Four different release types were used to investigate how different release configurations affect quantification accuracy. These included: a short, modular vent stack that could be adjusted to a release height of 1.7 m, 3.4 m, and 4.8 m above ground; a 13 m tall unlit flare stack; a 2.79 m tall storage tank; and a storage shed. The storage shed release was done through a perforation in the wall on the leeward side of the shed, approximately 1.5 m above the ground.

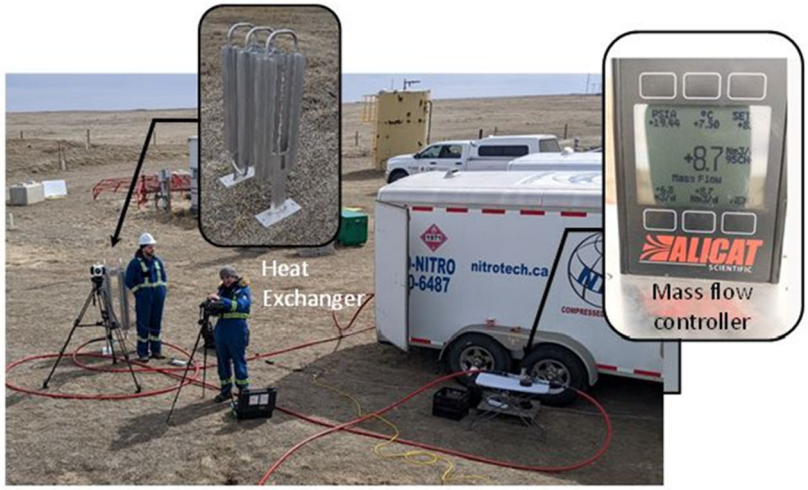

Figure 8: Compressed gas was released via a regulator and then conditioned to the ambient air temperature using a heat exchanger, before entering the mass flow controller. It was subsequently discharged to the atmosphere according to the particular release condition.

Most of the releases were performed using the 1.7 m stack and unlit flare. Release height is known to affect the quantification accuracy of airborne NIR HS imaging due to the way in which the plume interacts with the ground. It may also impact the performance of the truck-mounted TDLAS system via Eq. (4) (e.g., if the measurement path is beneath the plume). Release height also impacts ambient wind, which generally increases with height above ground. Finally, releases from the storage tank and shed were conducted to examine how aerodynamic interactions with the structure (e.g., turbulent wakes) and reflection from the metallic surfaces may impact the performance of QOGI systems.

Releases conducted during the third field campaign were done exclusively using the 1.7 m stack and the unlit flare. While the first two trials considered a single emission scenario at a given time, in the case of the third field campaign, releases took place simultaneously at two sites, each equipped with a trailer. The sites were located approximately 250 m apart on an approximately ENE-WNW track so that the helicopter could measure two scenarios in a single pass, as shown in Figure 9. This distance was deemed sufficient to avoid issues with plume overlap between the two sites, which was verified by examining infrared images. The east site releases were conducted with a 1.7 m stack while those from the west site were done from either a 1.7 m stack or 14 m unlit flare, although this was not known by the technology provider and data analysis team.

4.1.3 Emission quantification technologies and providers

Technology providers conducted measurements of the releases as they would in practice. The measurements were "semi-blind" in that the operators knew the release locations but did not know the release flow rate. Operators then processed the data and reported their release estimates to the research team, in most cases within three weeks of the measurement campaign. The operators were not provided any ancillary meteorology data from the FRS; instead, they either deployed their own

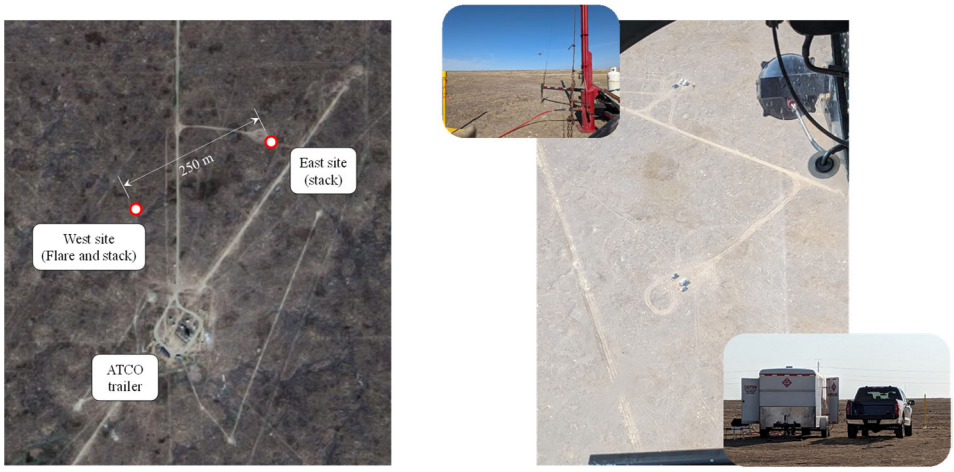

Figure 9: Location of releases for the third field trial, and view from the LSI helicopter. The helicopter flew repeated east-west and west-east passes, with pauses of 2-3 minutes each time the release rate was changed to allow the plume to develop.

instrumentation (Boreal) or employed a third-party weather service, e.g., Meteoblue (GHGSat) or Meteomatics (LSI/Telops).

Details of each technology/provider are summarized below:

Quantitative optical gas imaging (QOGI): QOGI measurements were conducted by three participants using two different systems, as summarized in Table 3. During the first field campaign, measurements were conducted by an operator from AGAT using a FLIR GF320 camera coupled to a Providence Photonics QL320 tablet (v 3.0.0.5), and an operator from UW (Nagorski) using SAIT's FLIR GFx320 camera connected to a FLIR QL320 (v. 1.4.1) tablet. The AGAT operator was highly experienced and proficient with their camera, while the UW operator was a novice. Notably, the SAIT system was considerably newer than the AGAT system and included key upgrades in the quantification software. Unfortunately, the SAIT camera was only available while SAIT was on site, so only a small number of simultaneous measurements were carried out by both operators at the same time.

During the second campaign, the UW operator was joined by an operator from Montrose. The UW operator used the same system as he used in the first campaign, while the Montrose operator used the OPGAL EyeCGas system. While the Montrose operator had extensive QOGI experience with the FLIR system, this was the first time they used the OPGAL system, and they reported some challenges due to his unfamiliarity with the system during the field trial. This lack-of-experience is reflected in some of the results shown in Sec. 5.2.

Table 3: Summary of OGI operators and equipment

Operator	Experience	System	Field trial
AGAT	Professional	FLIR GF320, Providence Photonics QL320 (v 3.0.0.5)	1
UW	Novice	FLIR GFx329, FLIR QL320 (v. 1.4.1)	1 and 2
Montrose	Professional, but new to system	OPCAL EyeCGas (v.1.0.24)	2

Drone-mounted TDLAS: The drone flew a descending helical flight pattern around each release, as shown in Figure 3. The altitude and radius of this pattern were adjusted heuristically based on wind speed and emission height. Raw data consisted of concentration, drone altitude, and GPS coordinates. Emission rates were then inferred from mass balance conducted on a control surface defined by the helical flightpath of the drone. The drone team remarked that the number of drone transits through each plume, which was limited by the interval with which the release rate was adjusted, was considerably lower than what they would normally use when conducting independent measurements.

Truck-mounted TDLAS: Methane releases were also quantified using a truck-mounted TDLAS system (Boreal Laser GasFinder 3VB). At the beginning of each day, the operators would position a 3D ultrasonic anemometer near the release point. During the measurements, the truck would search for the plume at distances ranging from 50 to 100 m downwind of the release point, driving over both roads and unprepared ground. The TDLAS system returned a path-average concentration (ppm) at one-second intervals (see Sec. 2.2), which is recorded along with the truck location, obtained from an onboard GPS unit. This information was displayed instantaneously to the truck operator and driver as a colored line that shows the truck path and concentration. The operators used this information to align their subsequent plume transects, in order to account for wind shift.

The data was post-processed by the operator using WindTrax [7], a backwards Lagrangian stochastic wind model. The operator also provided detailed concentration/GPS location to the UW team for subsequent uncertainty analysis.

Airborne NIR HS: Measurements were done at an altitude of approximately 1500 m AGL and 240 km/hr using a Piper Navajo equipped with a downward-looking NIR IFTS. The provider had the GPS location of the release, and flew repeated passes around the release, as shown in Figure 10. The flight path was roughly aligned with the wind direction. The provider used an internal algorithm (see Sec. 2.3) with wind speeds obtained from Meteoblue to obtain the emission estimates. The measurement swath is approximately 500 m [22].

Airborne LWIR HS imaging (LSI/Telops): Measurements were done at an altitude of approximately 350 m AGL and 35 m/s using a downward-looking LWIR IFTS mounted in a Bell 206-B helicopter (Figure 10). At this altitude and speed, the camera has a field-of-view of approximately 30 m × 30 m, which is considerably smaller than that obtained from the airborne NIR HS camera, principally due to the higher operational altitude of the GHGSat measurements, but it is more similar to the swath of the Bridger Gas Mapping Lidar (GML) system (128 m [9]). The Hyper-cam mini-LW airborne may be operated in "mapping" or "target" mode. In mapping mode, the downward view of the camera is fixed, allowing for a large swath of area to be inspected for leaks. In target mode, which was used for these tests, the GPS location of a known facility is entered into the system, which "locks" the camera on the target; this allows for multiple data cubes of a leak to be collected over a single pass, which is useful for quantification. The IFTS was operated with a spectral resolution of 10 cm^{-1}; at this resolution and airspeed, on average 15 data cubes could be captured of each site per pass. The system provides the operator with a real-time view of the infrared scene, along with a preliminary real-time mapping of the plume and ground temperature. The preliminary mapping is "conservative", meaning that a plume may not be visible to the operator, but may be detected and quantified during data processing after the measurement campaign.

Unfortunately, LSI/Telops were unable to participate in the first field campaign due to aircraft maintenance issues. During the second field campaign, the aircraft became unavailable for all but two days of the field trial, and, unfortunately, the camera targeting mode was set incorrectly, resulting in limited measurements for one day, while data from the second day was inadvertently deleted during post-processing at Telops. No such system issues were encountered during the third campaign, although the UW crew adjusted stack position was moved approximately 10 m to the perimeter of the imaging area for 1-2 hours without realizing the small size of the IFTS swath. Consequently, for these measurements much of the plume fell outside of the camera swath. These measurements were removed from the data set.

The third field campaign generated an enormous amount of data; approximately 17,000 data cubes had to be post-processed manually, which required approximately two months of dedicated analysis.

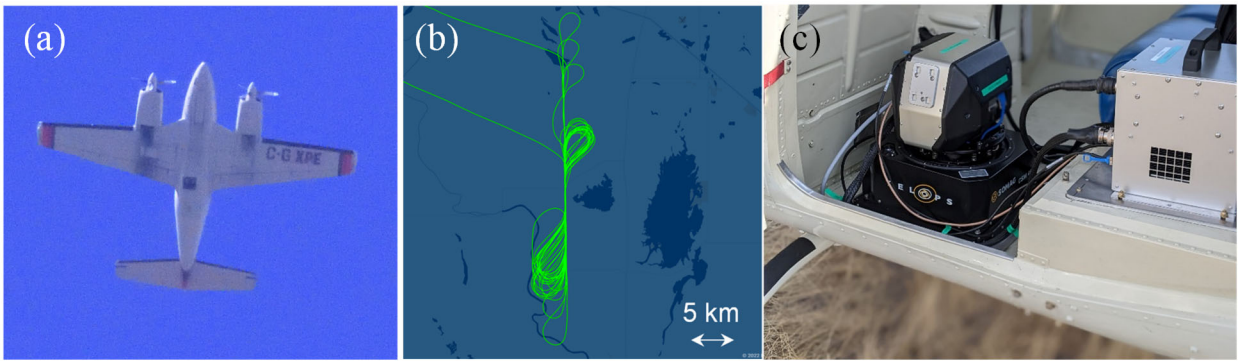

Figure 10: (a) Piper Navajo and (b) sample flight path used by GHGSat. (c) Telops Hyper-Cam Mini Airborne installed in LSI's Bell 206B.

4.2 Key outcomes of the first two field campaigns

Details of the first and second field campaigns are provided in the previous interim reports [20, 21]. Key findings include:

Truck-mounted TDLAS (Boreal): This technique provided the most accurate and unbiased estimates and appeared least-susceptible to environmental factors. On average, true emission rates were underestimated, and the bias appears to be proportional to the underlying release rate. Both the IGM and BLS models provided estimates of similar quality. Estimates from the unlit flare and stack were more accurate than emissions from the top of the storage tank. Measurement accuracy appeared resilient to changes in wind speeds.

Drone-mounted TDLAS (SAIT): Only 15 measurements were attempted by the SAIT team, and a majority of these were considered "missed detections" due instrument failure or unfavorable measurement conditions (e.g., unsteady wind.) The release estimates obtained from the drone were the least accurate of those from the first field trial. The SAIT team remarked that the short duration of the measurement events differed from the procedures they normally follow, which involved making a larger number of measurements per release.

QOGI: The second field campaign featured an extensive QOGI component, which systematically examined how a large set of parameters impact QOGI quantification accuracy. These results are

summarized here, but the reader is referred to the second field campaign report for a more detailed discussion [21].

QOGI estimates had a slight positive bias at low emission rates, and a negative bias at larger rates. In general, the UW QOGI estimates were slightly more accurate than those obtained from AGAT, which was attributed to upgrades made to the QOGI software between the Providence Photonics and FLIR versions of the QL320 (Table 3). Estimates from the Montrose operator were significantly less accurate than the other two operators, which was attributed to the Montrose operator's lack of familiarity with the OPGAL EyeCGas system.

Estimates of stack releases were more accurate than those made of emissions from the top of the tank, and significantly more accurate than those from storage shed. In the case of the tank, this was attributed to the distance between the emission location on the tank top and the ground based QOGI operator, and the complex aerodynamics at the top of the tank. The QOGI-inferred emission rates made on the shed were complicated due to a "pooling" of the gas in the aerodynamic wake of the shed, and reflection of the gas from the reflective surface of the shed (Figure 11).

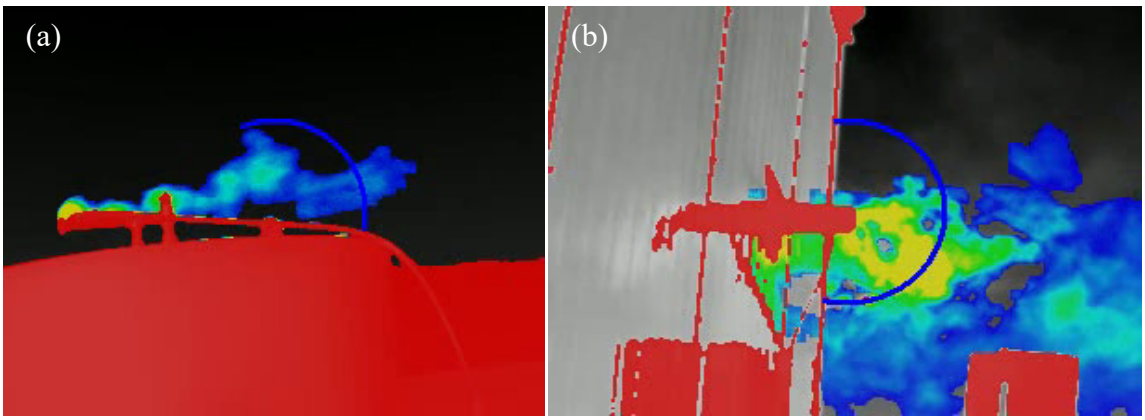

Figure 11: QOGI images of emissions from (a) the top of a tank and (b) the side of a shed.

There was no clear relationship between QOGI quantification accuracy and windspeed. This was a surprising result, since QOGI operators often use a wind-speed cutoff of 4.5 to 16 m/s as an operational limit beyond which the QOGI-inferred emission rate may be suspect [13]. In some specific measurements, an abrupt increase in wind speed led to a pronounced drop in measurement accuracy. However, very low wind speeds may also lead to "pooling" of the methane at the emission point, which also impacts quantification measurements. In general, the complexity and opacity of the feature tracking model used to connect the spectroscopically derived column densities to the emission rates made it difficult to develop a simple relationship between wind speed and accuracy, which motivates the development of empirical uncertainty quantification techniques (Sec. 5.2).

The effect of measurement distance on QOGI accuracy was examined systematically during the second field campaign (Figure 12). In general, increasing the distance between the emission source and camera impairs accuracy because it becomes more difficult to resolve the turbulent features that are tracked by the velocimetry algorithm. Emission accuracy also drops at very short distances to the release because: (1) it becomes difficult to contain the entire plume across the quantification

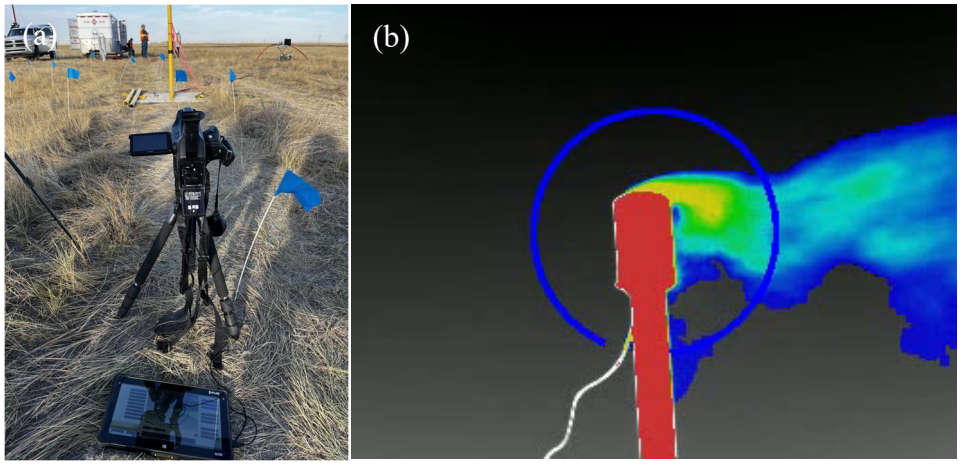

Figure 12: Procedure for measuring QOGI accuracy as a function of distance. (a) Blue flags denote various measurement radii from the stack base. (b) The blue circle shows the control surface used to measure the emission rate.

control line (Figure 12 (b)); and (2) the rate at which flow features travel across the scene approaches the limitations imposed by the camera framerate.

Plume background and thermal contrast were also found to impact QOGI accuracy. Cloud cover may degrade thermal contrast relative to clear sky, and the velocimetry algorithm may be impacted by cloud motion. Under some conditions, nearly horizontal measurements made with a clear sky background also gave poor thermal contrast, presumably because the effective sky background is due to emission along a long atmospheric path length, which, on average, is close to the local air temperature.

Airborne NIR HS imaging (GHGSat): The operator reported conditions during the first field campaign as marginal due to partial cloud cover; 32 out of 80 measurements were reported by the provider as "missed detections". Overall, the technology mostly underestimated the true emission rates, and bias was generally proportional to emission rate. There was no apparent correlation between wind speed and estimate accuracy; rather, cloud cover appeared to be the main factor that limited accuracy. The provider reported that cloud cover can affect the estimate by attenuating the incident sunlight used to measure the methane, and by casting a "mottled" pattern on the ground that confounds plume identification. In contrast to the other providers, GHGSat provided uncertainty estimates for their emission rates, although these underestimated the true error in the estimates.

Conditions during the second field campaign were nearly ideal for the airborne NIR HS imaging system, and it performed markedly better; only 2 of 46 measurements were reported as "missed detections". The estimates had an average error of 62% and, in contrast to the first field campaign, systematically overestimated the true emission rate.

4.3 Third field campaign

The third field campaign, which took place between September 18-23, 2023, focused exclusively on the airborne LWIR HS system. The objective of this field campaign was twofold: first, to assess the predictive capability of the Telops MDL tool, which is used to plan and execute missions (e.g., identify conditions where measurements should not take place or inferred detections may not be reliable); and second, to assess the accuracy of the emission estimates obtained from the data cubes. The aircraft operated at its nominal flight condition (~350 m AGL, ~35 m/s) for the first five days and the first half of the sixth day. On the second half of the sixth day the helicopter operated at different altitudes and airspeeds to explore how these parameters affected the MDL and quantification accuracy. These results are excluded from this report.

The total flight time of the Bell 206 was dictated by its fuel capacity, which depended on payload. With a pilot, equipment operator, and hyperspectral camera, the mission flight time was approximately 3 hours (excluding required factor-of-safety). The helicopter was based out of Brooks Airport, about a 10-minute flying time from the FRS, giving a useful measurement time of approximately 2 hours and 40 minutes. On each day, except for Sep 18, Sep 20, and Sep 23, the helicopter flew two missions: one in the AM, and one in the PM, with a lunch break and refueling between missions. A total of 204 individual release scenarios were considered during this time; for each scenario the helicopter flew five overhead passes, and collected, on average, 15 data cubes per pass for a total of 16,883 data cubes.

Table 4: Measurement times for the Telops HyperCam Mini-LW Airborne. Start and end times refer to the first and last measurement times for each mission in MST.

Date	**First Mission**		**Second Mission**	
	Start	End	Start	End
Sep 18	10:14	12:42	Scrubbed (high wind)	
Sep 19	8:30	10:57	11:45	14:33
Sep 20	8:55	11:25	Scrubbed (poor thermal contrast)	
Sep 21	8:24	11:09	11:54	14:43
Sep 22	10:00	12:50	13:53	16:03
Sep 23	10:07	12:10	End of campaign	

The procedure for quantifying emissions from the hyperspectral data has been detailed in Sec. 2.3, but the MDL tool merits some further discussion. The LWIR HS camera detects methane based on the intensity measured from the spectral bin corresponding to the 1306 cm^{-1} methane line to that of neighboring spectral bins (Figure 13). Methane may only be detected when the difference between these two spectral bins is sufficiently larger than the noise equivalent spectral radiance (NESR) of the camera, which is due to a combination of photonic, thermal, and electronic noise. The difference between the spectral intensities is predicted by a spectroscopic model that accounts for the methane column density (i.e., the number of methane molecules between the camera and the ground, in ppm·m), the thermal contrast between the air and ground, and the relative humidity, which affects the intensity of the neighboring pixels due to the instrument line-shape of the spectrometer, which "smears" the methane line with neighboring water lines at nearby

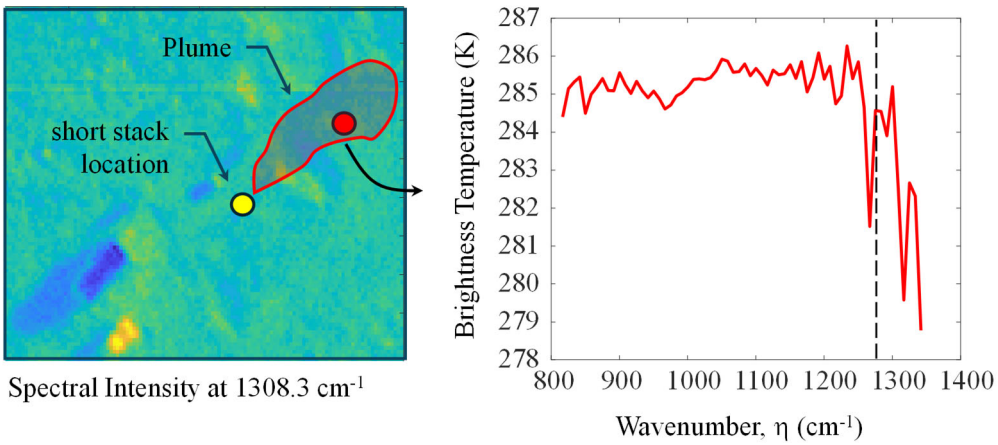

Figure 13: (a) Methane plumes are identified based on the spectral intensity at 1306 cm^{-1}, which is aligned with a strong CH_4 line. The plume outline is shown in red. (b) Brightness temperature (scales with spectral intensity) highlights the presence of methane corresponding to the red dot in (a).

wavenumbers. The methane column density, in turn, is related to the leak rate and wind speed through a Gaussian plume dispersion model.

4.3.1 Weather conditions and emission rates

The weather conditions during the field trials are summarized in Table 5. In general, a wide range of conditions were encountered, which elucidated how meteorological conditions affected both MDL and quantification accuracy. The weather conditions were generally favorable except on Monday, September 18, when measurements were suspended due to wind speeds outside of the safe operating envelope of the aircraft.

Table 5: Weather conditions during the third field campaign, taken from the Davis WeatherCenter located at the FRS. Minimum and maximum conditions pertain to operational times.

Date	**Min/Max Temp [°C]**		**Min/Max Wind [m/s]**		**General weather conditions**
	Min	Max	Min	Max	
Sep. 18	20.4	27.1	1.8	8.5	Poor (high wind at noon)
Sep. 19	8.1	16.7	0.9	6.3	Good
Sep. 20	7.3	11.1	1.8	6.7	Poor (low thermal contrast)
Sep. 21	2.4	18.3	0	3.6	Mixed
Sep. 22	11.4	22.3	0.4	5.8	Good
Sep. 23	14.9	19.8	4.5	9.4	Good (morning only)

As noted in Sec. 4.1.2, the MDL tool was used to inform the choice of emission rates to ensure that the system was "pushed to its limits" without having too many scenarios in which it would be impossible to detect emissions. Accordingly, in addition to the meteorological instrumentation described in Sec. 4.1.1, the UW team used a FLIR T650sc LWIR infrared camera to measure the ground temperature throughout the tests, which changed continuously with solar irradiation, cloud cover, and wind. The team then input the temperature into the MDL tool and obtained a real-time estimate of the MDL throughout the tests. The ground temperature was found to vary significantly

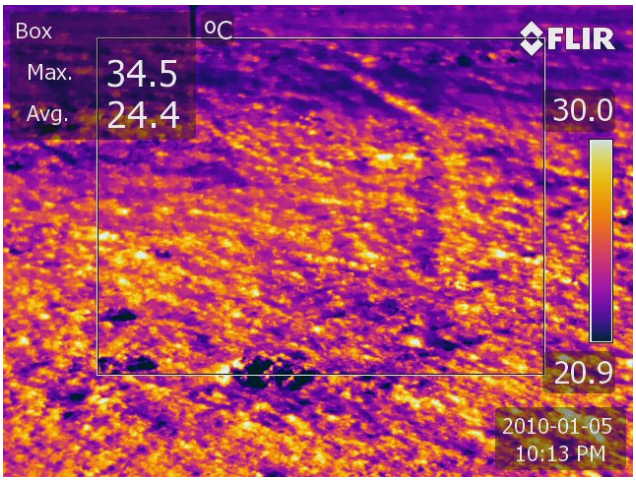

Figure 14: Example infrared image of the ground taken by the UW team during testing.

in a single image due to differences in ground cover, e.g., bare earth versus grass (Figure 14). The UW team interpreted the ground temperature as the average within the measurement box.

The hyper-cam operator may also calculate the MDL in real time but using the radiometric ground temperature found from the hyper-cam. Periodically the hyper-cam operator would report his ground temperature estimate, which the UW team would compare with their measurements (without sharing the results with the hyper-cam operator.)

The technology providers remarked that they normally started their missions at around 10 AM, to maximize the amount of useful measurement time in the context of the MDL. As indicated in Table 4, measurements started between 8:30 and 9 AM on Sep 19, 20, and 21 to explore how the MDL varied throughout the day. As an example, Figure 15 shows how the ground temperature, air temperature, wind speed, and predicted MDL varied throughout the measurements during Sep 19, based on ground temperature measurements from the hyperspectral camera. The MDL is initially very large due to the low thermal contrast between the ground and air temperatures. The thermal contrast initially grows as the ground temperature increases with solar heating, and the air temperature lags due to less efficient heat transfer between the ground and the air. As the ground is heated by sunlight the thermal contrast increases, although it lags, and the MDL generally drops as the day progresses. The increase in MDL in the afternoon is attributed mainly to a temporary drop in ground temperature, caused by cloud cover. Wind speed can also affect the MDL, as a higher wind speed dilutes the plume and lowers the column density for a given pixel, although in our tests the thermal contrast appeared to be the most significant factor affecting MDL.

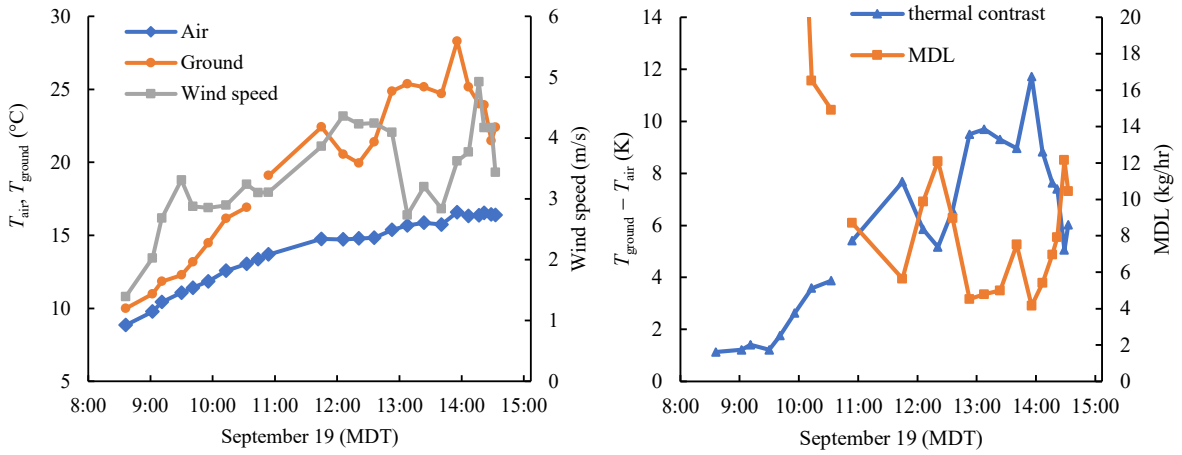

Figure 15: (left) Wind speed, ground temperature, and air temperature; and (right) temperature contrast and minimum detection limit (MDL) for September 19. On this day, the MDL was effectively infinite due to low thermal contrast before 10 AM MST.

4.3.2 Predictive performance of MDL tool

During post-processing, Telops recalculated the MDL based on the radiometric ground temperature, along with air temperature, wind speed, and relative humidity from the Meteomatics website. After analyzing all 16,883 data cubes, Telops detected more than 93% of the emission rates, and, overall, the detection capabilities of the system met or exceeded what was predicted using the MDL tool. Notably, there were no "false positive" detections (i.e., detecting methane during a null release). While Telops considered all 16,883 data cubes individually, Figure 16 shows detected releases vs. missed releases for each measurement event, assuming that a "detection" constitutes detected methane for 50% of the data cubes for that release. The results clearly demonstrate the predictive capability of the MDL tool. A more detailed summary of the performance of the Telops MDL tool will be made available via the AMEP program.

4.3.3 Quantification accuracy of the airborne LWIR HS system

While the Telops MDL tool appears to predict the minimum detectable limit, Figure 17 shows that the quantification tool systematically underpredicts the true emission rate, by a factor of approximately 2.5. The UW team is currently working with Telops to understand what aspect of their measurement model is responsible for this underprediction. While there is clearly a systematic bias, it should be noted that this performance is not atypical of other instrumentation, as discussed in Sec. 5. Future work regarding the airborne LWIR HS system is discussed in Sec. 6.

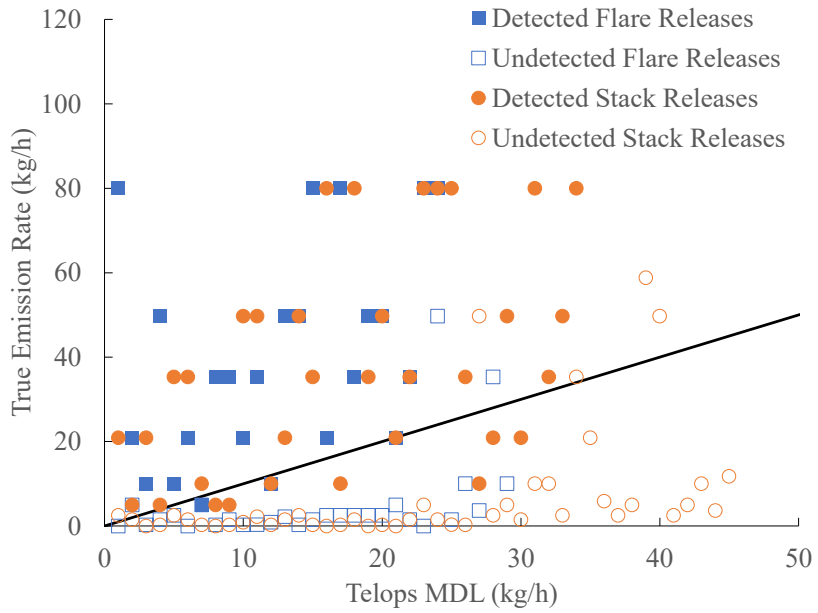

Figure 16: Predictive performance of the MDL tool. An emission is considered detected if it is detected in over 50% of the data cubes collected for that scenario. Emissions above the solid line should be detected, according to Telops's MDL tool.

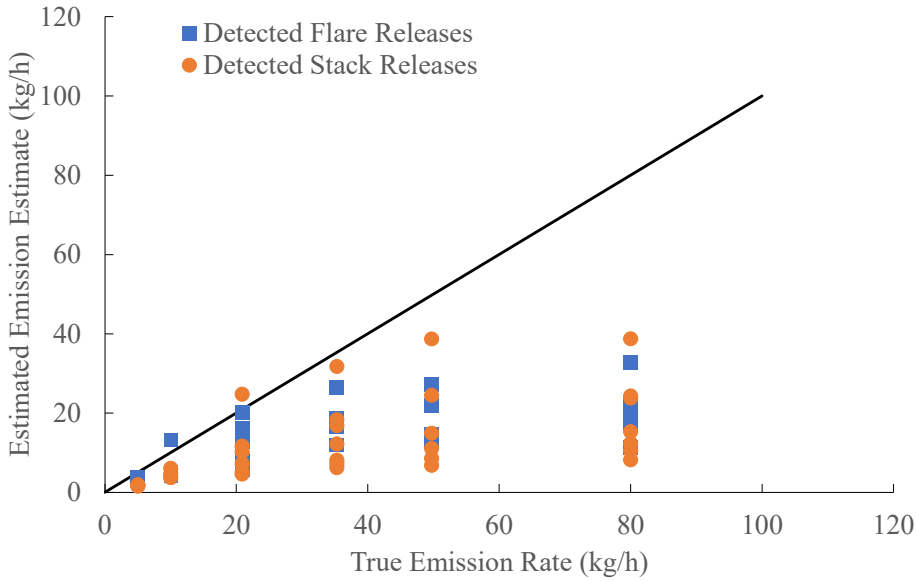

Figure 17: Performance of the quantification tool for the detected emissions.

5. PHASE 4: DATA ANALYSIS AND SUMMARY

The overarching objective of this project is to understand the uncertainty attached to emission estimates from the systems identified in Phase 1, and, more generally, develop generalized techniques that can be applied to any methane emissions quantification technology. This research phase has three main directions: (1) device-specific uncertainty quantification from a detailed instrument model; (2) a generalized uncertainty analysis using empirical models derived from the field campaign data; and (3) implementation of uncertainty estimates in a Monte Carlo model (e.g., AROfemp or LDAR-Sim [23]) to assess the expected uncertainty in overall emissions reductions and the cost-effectiveness of various FEMPs.

5.1 Instrument-level uncertainty estimation

In scenarios in which a detailed instrument model is available, as is the case for the truck based TDLAS, it may be possible to derive uncertainty estimates by accounting for measurement noise, model parameter uncertainty, and model error within the Bayesian framework. An instrument model has the general form

$$\mathbf{b}^{\text{meas}} = \mathbf{b}^{\text{true}}\left(Q^{\text{true}}\right) + \boldsymbol{\delta}\mathbf{b}^{\text{noise}} \tag{7}$$

where $\mathbf{b}^{\text{meas}}$ is a vector of measurements collected from the instrumentation, $\mathbf{b}^{\text{true}}$ is the true but unknown parameter, that is related to the quantity-of-interest Q by some problem physics. While the true physics is unknown, a simplified model is used in its place, but with some additional model error

$$\mathbf{b}^{\text{true}}\left(Q\right) = \mathbf{b}^{\text{model}}\left(Q,\boldsymbol{\Phi}\right) + \boldsymbol{\delta}\mathbf{b}^{\text{model}} \tag{8}$$

where $\boldsymbol{\Phi}$ is a vector of additional unknown model parameters. The overall measurement equation becomes

$$\mathbf{b}^{\text{meas}} = \mathbf{b}^{\text{model}}\left(Q,\boldsymbol{\Phi}\right) + \boldsymbol{\delta}\mathbf{b}^{\text{noise}} + \boldsymbol{\delta}\mathbf{b}^{\text{model}} = \mathbf{b}^{\text{model}}\left(Q,\boldsymbol{\Phi}\right) + \boldsymbol{\delta}\mathbf{b} \tag{9}$$

A key observation is that the modeled data is not expected to "exactly" match the measurements due to the $\boldsymbol{\delta}\mathbf{b}$ term, even if the "true" model parameters are known, due to two factors: (1) the measurement error, $\boldsymbol{\delta}\mathbf{b}^{\text{noise}}$, e.g., sensor noise; and (2) the model error, $\boldsymbol{\delta}\mathbf{b}^{\text{model}}$, which come from the simplifying assumptions and approximations needed to derive the model equations, which causes the model to differ from "the real world."

Both measurement noise and model error may be envisioned as "random variables" that obey underlying probability density functions (PDFs). This treatment is obvious in the case of the measurement noise, which comes from a random process, but is less clear for the model error, since the difference between the model and true physics should be systematic, i.e., a bias. However, the analyst's *state-of-knowledge* about this error may be considered as random, which justifies this treatment. The model parameters, $\boldsymbol{\Phi}$, are also imperfectly known, and are therefore treated as random variables that obey PDFs. In conventional analysis these parameters are often specified, but in Bayesian inference they are treated as *nuisance variables* and inferred from the measurements along with the leak rate.

Since $\boldsymbol{\delta}\mathbf{b}^{\text{noise}}$ and $\boldsymbol{\delta}\mathbf{b}^{\text{model}}$ are random variables, both the measurement data, $\mathbf{b}^{\text{meas}}$, and the estimated emission rate, Q, are also random variables described by PDFs. These PDFs are related by Bayes' equation

$$p\left(Q,\boldsymbol{\Phi}\middle|\mathbf{b}\right)=\frac{p\left(\mathbf{b}\middle|Q,\boldsymbol{\Phi}\right)p_{\mathrm{pr}}\left(Q,\boldsymbol{\Phi}\right)}{p\left(\mathbf{b}\right)}=\frac{p\left(\mathbf{b}\middle|Q,\boldsymbol{\Phi}\right)p_{\mathrm{pr}}\left(Q\right)p_{\mathrm{pr}}\left(\boldsymbol{\Phi}\right)}{p\left(\mathbf{b}\right)} \tag{10}$$

where $p(\mathbf{b}|Q, \boldsymbol{\Phi})$ is the likelihood of observing the instrument data for a given leak rate and set of model parameters (i.e., the "likelihood"), $p(Q, \boldsymbol{\Phi}|\mathbf{b})$ is the posterior probability of a leak rate and set of model parameters given the observed data (the "posterior"), and $p_{\mathrm{pr}}(Q)$ and $p_{\mathrm{pr}}(\boldsymbol{\Phi})$ are PDFs that define what is known about the leak rate and model parameters before the measurement takes place (the "priors"). The denominator in Eq. (10) is called the "evidence", and is used to normalize the numerator so that the posterior probability satisfies the Law of Total Probability,

$$\int_{Q,\boldsymbol{\Phi}} p\left(Q,\boldsymbol{\Phi}\middle|\mathbf{b}\right)dQd\boldsymbol{\Phi}=1 \tag{11}$$

Note that, in the absence of prior information, $p_{\mathrm{pr}}(Q, \boldsymbol{\Phi}) = 1$ and $p(Q, \boldsymbol{\Phi}|\mathbf{b}) = p(\mathbf{b}|Q, \boldsymbol{\Phi})$, which means that the likelihood of observing the data conditional on Q and $\boldsymbol{\Phi}$ equals the posterior probability of a given $\boldsymbol{\Phi}$ and Q conditional on the observed data. More specifically, the set of (Q, $\boldsymbol{\Phi}$) that maximizes the likelihood of observing the data also maximizes the posterior probability density; this is the maximum likelihood estimate (MLE), and, for independent and identically-distributed noise, the MLE is often found by least-squares minimization of the residual between $\mathbf{b}$ and $\mathbf{b}^{\mathrm{model}}(Q, \boldsymbol{\Phi})$. Importantly, however, the uncertainty in these variables is conveyed by the width of the posterior distribution, which can be very large for ill-posed models. Accordingly, additional information about Q and $\boldsymbol{\Phi}$ should be incorporated via prior PDFs, which narrows the posterior PDF compared to having an uninformative prior, $p_{\mathrm{pr}}(Q, \boldsymbol{\Phi}) = 1$.

If $\boldsymbol{\Phi}$ contains m parameters, $p(Q, \boldsymbol{\Phi}|\mathbf{b})$ is a multidimensional joint PDF having m+1 dimensions. A univariate PDF for Q may be found by "marginalizing out" the nuisance variables in the PDF,

$$p\left(Q\middle|\mathbf{b}\right)=\int_{\boldsymbol{\Phi}} p\left(Q,\boldsymbol{\Phi}\middle|\mathbf{b}\right)d\boldsymbol{\Phi} \tag{12}$$

which may be interpreted as the probability density function for the leak rate given a set of measured values, for all possible model parameters. This distribution is a comprehensive description of what is known about the emission rate conditional on the measured data and prior information. The marginalized posterior may be summarized using Bayesian credibility intervals,

$$Q\in\left[Q_a,Q_b\right] \quad s.t.\int_{Q_b}^{Q_b} p\left(Q\middle|\mathbf{b}\right)dQ=\alpha \tag{13}$$

where α is some probability, e.g., 90%. Accordingly, this interval is interpreted as containing the "true" leak rate with 90% probability. A representative marginalized posterior PDF and credibility intervals is shown in Figure 18.

In order to derive the likelihood PDF, $p(\mathbf{b}|\ Q, \boldsymbol{\Phi})$, the model for $\mathbf{b}(Q, \boldsymbol{\Phi})$ must be available, and the probability density function of the model error, $\boldsymbol{\delta}\mathbf{b}^{\mathrm{model}}(Q, \boldsymbol{\Phi})$, must also be known. The first condition is difficult to satisfy for many commercial quantification systems, since the model equations are often highly complex and proprietary; as a case in point, the velocimetry model in most QOGI systems is proprietary and opaque, as is the technique used to convert the TDLAS column densities into emission rates used by Bridger and GHGSat. On the other hand, the truck-mounted TDLAS system used an open-source backwards Lagrangian solver (BLS) to convert

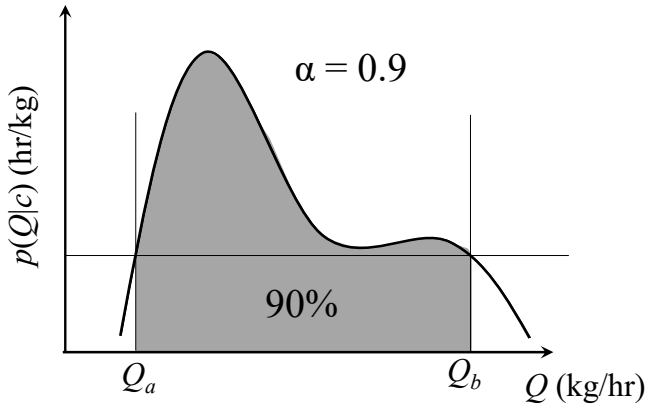

Figure 18: Calculation of 90% Bayesian credibility intervals from marginalized posterior PDF.

concentration measurements into emission rates, and for this reason we focus on this system for developing an instrument-specific uncertainty quantification methodology. For similar reasons, Zhou et al. [24] used the BLS model as a basis for developing a Bayesian uncertainty estimate for methane emission rates inferred from drone-based TDLAS concentration measurements. In this work we use the inverse Gaussian plume model, derived from Eq. (4), which was also used as a basis for the Bayesian uncertainty quantification in Ref. [25].

In this scenario **b** contains concentration measurements provided by the onboard TDLAS system at one second intervals, while the model parameters in $\mathbf{\Phi}$ consist of $\{\sigma_x, \sigma_y, \sigma_z\}$, which depend on the atmospheric stability class, and the wind speed, U. (In this semi-blind measurement, the release height and location are perfectly known, as is the relative location of the truck from the onboard GPS sensor.) These parameters are treated as "unknowns", but with prior PDFs that are developed from the literature for σ_x, σ_y, and σ_z, and, in the case of the wind speed, from meteorological data. Notably, the wind speed prior can be chosen depending on the source of the wind speed; in the case of an on-site ultrasonic anemometer, it may be very narrow, while if an online data source is used (e.g., Meteoblue) it would be broader, reflecting the fact that this parameter is known with less certainty.

It is also possible to derive a prior PDF for the release rate, the form of which depends on the specificity of prior information. At a bare minimum, it is known that Q must be strictly non-negative, and oftentimes there is an estimate of how large the release could be, Q_{est}. In this case, an appropriate prior (one that satisfies the maximum entropy criterion [26]) is

$$p_{\text{pr}}(Q) = \frac{1}{Q_{\text{est}}} \exp\left(-\frac{Q}{Q_{\text{est}}}\right) \tag{14}$$

which is defined so that there is a 50% probability that $Q \in [0, Q_{\text{est}}]$ and 50% probability that $Q \in [Q_{\text{est}}, \infty)$.

A second challenge concerns the underlying distribution of $\boldsymbol{\delta}\mathbf{b}$, which is dominated by the difference between the Gaussian plume model and the true turbulent advection physics. The Gaussian plume model is often a reasonable approximation of the time-averaged concentration profile of a plume emanating from a point source, provided the wind is constant and steady, and the terrain is homogeneous and flat. However, the truck measures a near instantaneous concentration distribution, and even were the wind to be constant, meandering and relative

Figure 19: Instantaneous snapshot of tracer smoke from the first field campaign, highlighting that the instantaneous concentration field differs from the time-averaged profile, even under constant wind conditions.

dispersion creates a complex instantaneous picture of the concentration profile [27] (Figure 19). The situation becomes even more complex when accounting for changes in wind speed and direction over the timescale of the plume transect.

Zhou et al. [24] assume that **δb** obeys either a normal or lognormal distribution, giving rise to a normal or lognormal likelihood function

$$p\left(b\middle|Q,\mathbf{\Phi}\right)=\begin{cases}\dfrac{1}{\sigma_e^{\mathrm{G}}\sqrt{2\pi}}\exp\left[-\dfrac{1}{2}\left(\dfrac{b-b\left(Q,\mathbf{\Phi}\right)}{\sigma_e^{\mathrm{G}}}\right)^2\right] \\ \dfrac{1}{\sigma_e^{\mathrm{LN}}b\sqrt{2\pi}}\exp\left[-\dfrac{1}{2}\left(\dfrac{\ln\left(b\right)-\ln\left(b\left(Q,\mathbf{\Phi}\right)\right)}{\sigma_e^{\mathrm{LN}}}\right)^2\right]\end{cases} \tag{15}$$

where b is cross-plume integrated mass concentration. The parameters σ_e^{G} and σ_e^{LN} are estimated from the residual between measured concentration and modeled concentration obtained under controlled-release experiments. Zhou et al. then estimate the posterior PDF using concentration measurements obtained through sequential plume transects by a drone equipped with a TDLAS sensor; the posterior PDF derived from each transect becomes the prior for the one estimated from next transect through a Bayesian updating process [24]. In this way, the posterior PDF becomes progressively narrower from successive transects.

In this ongoing work, we will derive the likelihood function using the CFD-LES results shown in Figure 7, and then apply the results to truck-based concentration measurements obtained from isolated transects conducted during the first and second field campaigns. The success of this technique will be gauged based on whether the true release rate lies within the derived credibility intervals with the expected probability.

5.2 Generalized empirical uncertainty estimation using field campaign data

In most cases, however, the detailed information needed to derive a model-based likelihood PDF is unavailable, particularly to third-parties who may wish to construct their own, independent, uncertainty estimates for emissions measurements. Accordingly, we developed a flexible empirical

framework that can be applied to any "black box" technology provided there is sufficient controlled release data for this technology.

We begin by asking the question: for a given true emission rate, what range of measurements could be expected given the observed controlled release data, considering model error and measurement noise? Let Q_i^{true} be the true emission rate corresponding to the *ith* observation in the field campaign, and Q_i^{est} be the estimated emission rate for the *ith* observation. As indicated in the field campaign data, the relationship between Q_i^{true} and the bias and variability usually change with the emission rate, and not necessarily in a linear way. Also, any empirical model must account for the fact that all technologies may report a "false positive", that is, a non-zero Q_i^{est} when Q_i^{true} is actually zero.

A flexible likelihood that allows the bias and variance of Q_i^{est} to vary with Q_i^{true} and incorporates the possibility of false positive is given by

$$\log\left(Q_i^{\text{est}}\right) = \log\left(\phi_i\right) + \varepsilon_i \tag{16}$$

where

$$\varepsilon_i \sim \mathcal{N}\left(0, \sigma_i^2\right) \tag{17}$$

and

$$\phi_i \sim \text{median}\left(Q_i^{\text{est}}\right) \tag{18}$$

This likelihood is normal on the log scale, which makes it lognormal on the measurement scale. The piecewise function ϕ_i has the form

$$\phi_i = \begin{cases} \alpha_0 + \alpha_1 Q_i^{\text{true}} + \alpha_2 \left(Q_i^{\text{true}}\right)^2 & Q_i^{\text{true}} \le \gamma \\ \alpha_0 + \beta_0 + \left(\alpha_1 + \beta_1\right)\left(Q_i^{\text{true}}\right)^2 & Q_i^{\text{true}} > \gamma \end{cases} \tag{19}$$

which transitions from a quadratic function for values of Q_i^{true} below a threshold γ to a linear one above this threshold. Continuity at this threshold is imposed by enforcing $\beta_0 = \alpha_2\gamma^2 - \beta_1\gamma$. The model parameters are adjusted to give the best prediction results and fit to the data, following a Bayesian procedure. The likelihood function covers the expected or "average" behavior of the estimates given a true emission rate; in a perfect scenario all the coefficients in Eq. (17) equal zero except for α_1, which should equal unity. However, this is never the case due to systematic offsets and biases in the measurements caused by imperfectly known model parameters, and real-world physical phenomena that are not captured by the model. For example, α_0 accounts for "false positives", in which the technology detects an emission rate when the true emission is zero. The variance represents random errors that include measurement noise and random aspects of the model error, e.g., time-dependent turbulent fluctuations in the plume.

In many cases, the true emission rate may be related to the median measured emission rate by a linear function, quadratic function, or blended linear-quadratic function. In this context, α_1 and α_2 are the coefficients for Q_i^{true} and $(Q_i^{\text{true}})^2$, describing the quadratic measurement between the median measurement and the true emission rate when the true emission rate is less than the threshold, γ, and sums $\alpha_0 + \beta_0$ and $\alpha_1 + \beta_1$ are the slope and intercept between the median measurement and true emission rate when the true emission rate exceeds γ.

The physical interpretation of this model is aided by exponentiating both sides of Eq. (16), leading to

$$Q_i^{\text{est}} = \phi_i \exp\left(\varepsilon_i\right) \tag{20}$$

where ϕ_i is the median measurement of Q_i^{est}. Thus, in general the random error manifests as a multiplicative error that may arise from an error in the advection sub-model, the sub-model used to infer concentration or column density, and measurement error. A similar scheme was recently proposed by Conrad et al. [28], although they assume that the error term has a constant variance for all release rates. In this work, the model defined by Eqs. (16)-(19) accounts for the fact that the variance may (and usually does) change with release rate and accommodates false positives: specifically, we either assume that the variance in Eq. (17) is either constant,

$$\sigma_i^2 = \sigma^2 \tag{21}$$

following Conrad et al. [28], or

$$\sigma_i^2 = \left(\tau + Q_i^{\text{true}} / \eta\right)^{-1} \tag{22}$$

A constant variance means that the expected randomness in the measured release rate should be proportional to the true release rate via Eq. (20), while the model in Eq. (22) means that the randomness increases sub-linearly with the release rate.

As a concrete example, consider the problem of recovering the emission rate from a downwind concentration measurement at a precisely known location and one instant in time, $c(t)$. In principle, the concentration may be related to the emission rate as

$$c = Q \cdot D\left(U\right) \tag{23}$$

where D is an advection model that depends on the wind speed, U, which is idealized as constant. The inverted measurement model is then

$$Q^{\text{est}} = D^{-1}\left(U\right) \cdot c^{\text{meas}} \tag{24}$$

which is the product of a concentration measurement and an advection term. In the case of the Gaussian plume model, for example, D^{-1} would be found by inverting Eq. (4) to solve for Q explicitly. In principle, there is a "true" but unknown advection model that relates the true release rate to the downstream concentration,

$$c^{\text{true}}\left(t\right) = Q^{\text{true}} D^{\text{true}}\left[t, U\left(t\right)\right] \tag{25}$$

and, furthermore, the measured downstream concentration is related to the true downstream concentration by an error term, $c^{\text{meas}} = c^{\text{true}}(t) + \delta c$, e.g., instrument noise. Substituting this result into Eq. (24) gives a relationship between the true and estimated emission rates

$$Q^{\text{est}} = Q^{\text{true}} D^{-1}\left(U\right) D^{\text{true}}\left[t, U\left(t\right)\right] + D^{-1}\left(U\right)\delta c \tag{26}$$

In an ideal world, D^{true} and D would be identical, and $D^{-1}D^{\text{true}}$ would equal unity. However, the simplifications used to derive $D(U)$ introduce both a bias and a random error, e.g., the turbulent fluctuations of the unsteady plume, wind gusts, and variations in surface roughness. In this

scenario, and assuming that δc is small, the median estimate should be related to the true release rate by

$$\phi_i = \alpha_0 + \alpha_1 Q_i^{\text{true}} \tag{27}$$

and Eq. (21) would hold for the variance.

Deviations from this model may arise when the advection model error or concentration measurement error are functions of leak rate; for example, in many cases the accuracy and precision of spectroscopic concentration measurements may be influenced by the total column density of methane, which depends on the emission rate. In this scenario, δc would depend on Q^{true}, which would introduce a nonlinear dependence in the variance, and possibly the median.

The complete set of model parameters is then

$$\mathbf{\Phi} = \left[\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1, \gamma, \tau, \eta\right]^T \tag{28}$$

A subset of model classes may be realized by setting some of these parameters equal to zero.

The model classes, and PDFs for the model parameters are estimated from a set of controlled releases. Suppose the measurements are contained in **M**, where $M_i = Q_i^{\text{est}}$, and the true emission rates are in **Q**, where $Q_i = Q_i^{\text{true}}$. In this case the relevant form of Bayes' equation is

$$p\left(\mathbf{\Phi}\middle|\mathbf{M}\right) = \frac{p\left(\mathbf{M}\middle|\mathbf{\Phi}\right) p_{\text{pr}}\left(\mathbf{\Phi}\right)}{p\left(\mathbf{M}\right)} \tag{29}$$

where $p(\mathbf{\Phi}|\mathbf{M})$ is the posterior probability density for the model parameters based on the measurements obtained from the technology in question, along with any prior knowledge of the parameters as specified by $p_{\text{pr}}(\mathbf{\Phi})$. The likelihood PDF, $p(\mathbf{M}|\mathbf{\Phi})$, specifies how likely it is to observe the measurements conditional on the set of model parameters as well as the true emission rates, $\{Q_i^{\text{true}}\}$. Finally, $p(\mathbf{M})$ is the marginal distribution of the data, which is fixed and allows $p(\mathbf{\Phi}|\mathbf{M})$ to satisfy the Law of Total Probability. Details about how the prior PDFs are chosen for the parameters, and the way in which a model is chosen from a set of candidate model classes, is discussed in detail in Appendix A.

We exemplify this technique by considering QOGI measurements on controlled releases made during the second field campaign. Figure 20 and Figure 21 shows the uncertainty quantification model results for the Montrose and UW operators, respectively. The shaded areas denote a specified probability of measuring a given leak rate given a true release rate. For example, Figure 20 shows that there is a 50% probability that the UW operator would observe an emission between 5 and 13 kg/h for a true emission rate of 20 kg/h.

These results provide valuable insight into the bias and variability of the measurements, as well as some of the underlying causes. For example, in the case of the UW operator the posterior mean prediction (the "most likely" prediction) is a nonlinear function of the true emission rate at low emission rates, and transitions to a linear function of the true emission rate at higher emission rates. The initial nonlinearity is common to many quantification technologies, and may be due, in part, to the prevalence of false positive detections at low release rates.

In contrast, Figure 21 shows a much larger bias in the inferred emission estimates, which is atypical of other QOGI measurements. This is indicative of a problem in the measurements and is likely due to the operator's lack of familiarity with the OPGAL system.

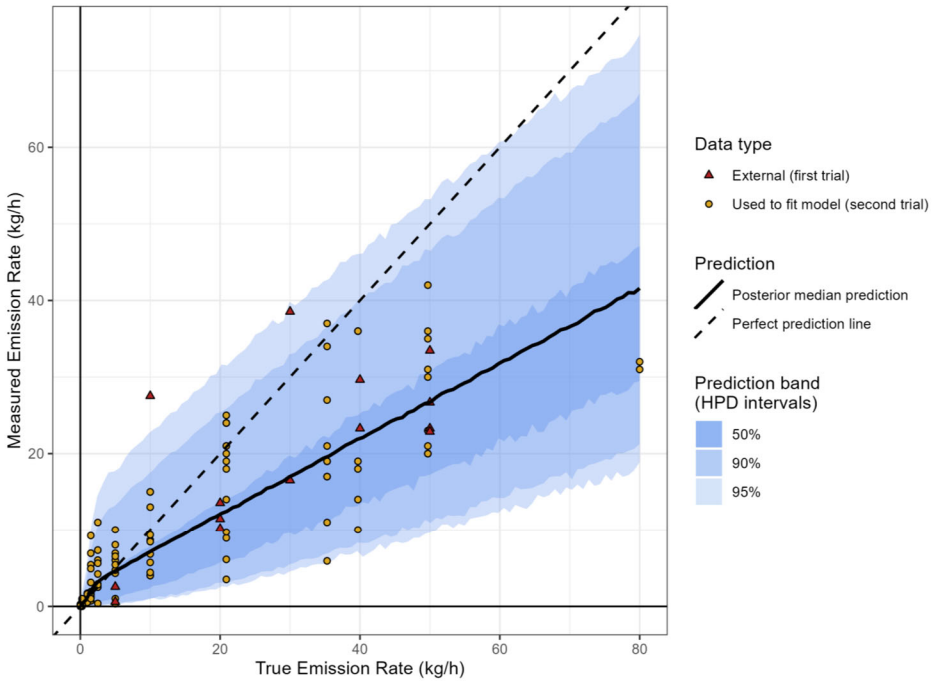

Figure 20: *Uncertainty quantification model results for the UW QOGI operator. Results from the first and second field campaigns are denoted by triangles and circles, respectively. The results from the first field campaign are used as external data to verify the adequacy of the model.*

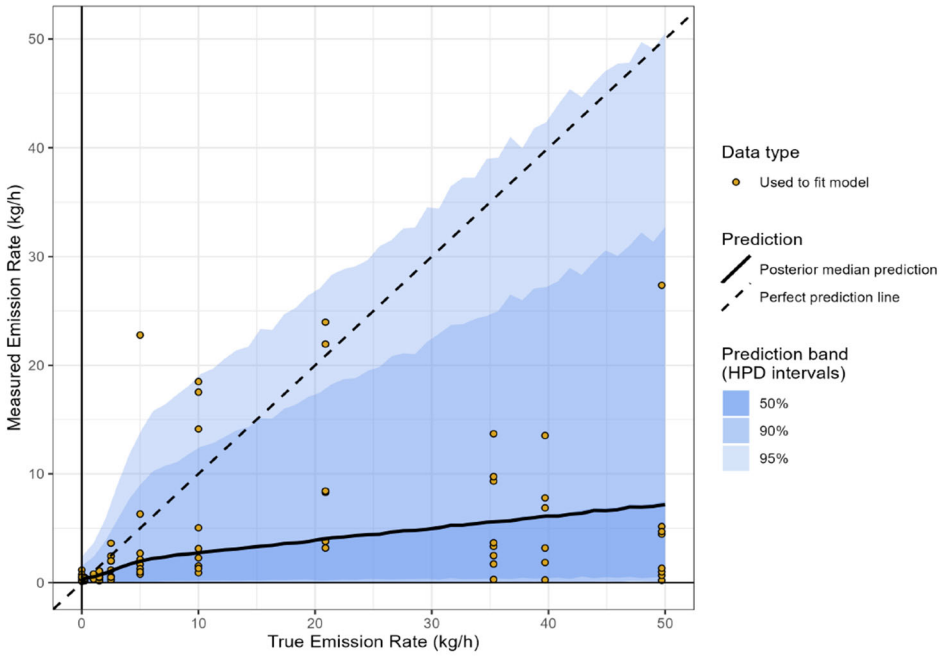

Figure 21: *Uncertainty quantification model results for the Montrose QOGI operator. The low slope highlights a systematic problem with the technology, likely due to the operator's inexperience with the camera.*

While this analysis provides key insights into the relative performance of quantification technologies, it can also be adapted to develop an uncertainty estimate for a new measurement made in an operational context. Suppose that Q^{new} represents the unknown true emission rate associated with a new measurement, M^{new}, made during a leak detection and repair survey, and we have available controlled release measurements made during a field campaign in vector **M**. We wish to obtain $p(Q^{\text{new}}|M^{\text{new}}, \mathbf{M})$, which through Bayes' equation is given by

$$p\left(Q^{\text{new}} \middle| M^{\text{new}}, \mathbf{M}\right) = p\left(M^{\text{new}} \middle| Q^{\text{new}}, \mathbf{M}\right) p_{\text{pr}}\left(Q^{\text{new}}\right) \tag{30}$$

where $p(M^{\text{new}}|Q^{\text{new}}, \mathbf{M})$ is the likelihood of observing the new measurement conditional on a hypothetical true release rate and the set of controlled release measurements, and $p_{\text{pr}}(Q^{\text{new}})$ is the prior probability of that emission rate, e.g., what is typical for a given piece of equipment. The likelihood is found by marginalizing out the influence of the model parameters in Φ,

$$p\left(M^{\text{new}} \middle| Q^{new}, \mathbf{M}\right) = \int_{\mathbf{\Phi}} p\left(M^{\text{new}} \middle| Q^{new}, \mathbf{\Phi}\right) p\left(\mathbf{\Phi}\middle|, \mathbf{M}\right) d\mathbf{\Phi} \tag{31}$$

where $p(\Phi|\mathbf{M})$ is the posterior PDF for the model parameters found from the controlled release data and $p(M^{\text{new}}|Q^{\text{new}}, \Phi)$ comes from the likelihood model in Eq. (16). Further details on this procedure are in Appendix A.

As an example, suppose that the UW QOGI operator measures an emission, which is expected to lie somewhere between 0 and 200 kg/hr with uniform probability. Further suppose that the true (but unknown) release rate is 25 kg/hr. Five sample measurements are drawn from the corresponding likelihood model shown in Figure 20, which are shown in Table 6. The posterior PDFs obtained from Eq. (31) are shown in Figure 22, for the case of a single measurement and using five measurements and reflect the complete state-of-knowledge about the releases, given the measurements, the prior information, and the set of controlled release data. Notably: (1) even though the estimates shown in Table 6 are significantly less than 25 kg/hr, because the QOGI technology systematically underestimates the release rate, the controlled release information "corrects" this bias, leading to a PDF that is roughly centered on the true release rate; (2) the posterior PDF obtained using five measurements is considerably narrower than the one found using a single measurement, since the additional measurements act to reduce the underlying variance of the measurements.

Table 6: Simulated QOGI-inferred emission rates, given a true emission rate of 25 kg/hr.

Measurement, i	M_i^{est} (kg/hr)
1	19.7
2	11.6
3	8.4
4	18.1
5	17.0

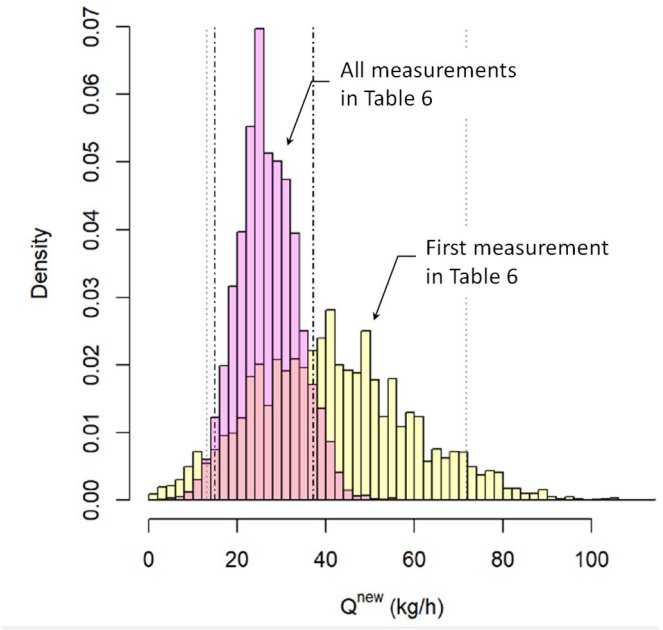

Figure 22: Posterior PDFs for the true emission rate, given the QOGI measurements, the prior information, and the controlled release data shown in Figure 20. One PDF corresponds to using the first measurement in Table 6, while the other uses all five measurements.

5.3 Incorporating uncertainty models into Monte Carlo FEMP simulations

One of the original goals of this study was to use the emission measurement models developed for different quantification technologies to study and compare different FEMPs using Monte Carlo simulation. A further goal was to incorporate measurement uncertainty for the different quantification technologies, in order to obtain an uncertainty estimate for these comparisons. Since several of the management decisions in a FEMP rely on emission measurements, incorporating uncertainty on those measurements allows for more realistic and informative comparisons and, eventually, better decision-making.

With the general empirical uncertainty model detailed in Section 5.2, we are now in a position to reach this goal. This work is currently being led by the UW team, with an expected completion date of May 2024. More precisely, we plan to compare a default FEMP (as per Alberta Energy Regulator Directive 060) and various alternative FEMPs involving different quantification technologies. We plan to simulate leaks and corresponding noisy measurements based on the models developed for the different quantification technologies, and then observe how this additional noise affect the deployment of the FEMP. In addition, we can propagate this uncertainty in the estimate of overall emission reductions. This approach allows one to perform comparisons between FEMPs that include a statistical assessment of the differences in emissions reduction observed and whether they are significant or not. In particular, using this approach, results such as those reported in Table 1 would be accompanied with an uncertainty estimate for the expected emission reduction.

Additionally, we can evaluate how the uncertainty in the measurements will affect the estimated cost and emission reductions of the various FEMPs being evaluated, compared to an idealized scenario where measurements are assumed to be exact. In turn, this may lead to the development of alternatives FEMP that are more robust to measurement uncertainty.

6. CONCLUSIONS AND ONGOING WORK

6.1 Conclusions

Canada's oil and gas industry urgently needs new technology that can detect and quantify methane emissions from its upstream operations. While considerable attention has focused on detection probability, the ability to quantify emissions is assuming new importance, particularly in the context of emerging policies and regulations and a growing understanding of how methane emitted by upstream oil and gas operations is accelerating climate change. However, these emission estimates can only be interpreted properly in the context of their uncertainty.

This research project focused on analyzing six existing and emerging methane quantification technologies: QOGI; truck-, drone-, and aircraft-based TDLAS; and airborne near-wavelength and long-wavelength hyperspectral imaging. The expected performance these technologies, were they to form the basis of an alternative fugitive emissions management plan (alt-FEMP), was estimated through a Monte Carlo simulation over two years of simulated emission events typical of an upstream site in Western Canada. The outcome of this simulation showed that alt-FEMPS based on airborne HS as well as airborne and truck-based TDLAS systems would realize significant emission reductions relative to a standard QOGI-based FEMP.

In parallel, the operational characteristics of the FLIR GF320 QOGI system was explored through a laboratory study using a heated vent apparatus. This study demonstrated the underlying measurement principles, capabilities, and limitations of this system, and shed light on how various real-world measurement conditions may impact the reliability of the inferred emission rates.

A subset of the technologies was then assessed through three measurement campaigns at CMC's field research station in Newall County, AB. Campaigns featured semi-blind controlled releases from stacks of various heights, an unlit flare, the top of a storage tank, and the leeward side of a structure. These measurements provided insight into the overall accuracy and precision of the technologies for a range of measurement scenarios, and their susceptibility to environmental parameters like wind speed and ground temperature.

The measurement campaign data was used to develop statistical models of the uncertainty attached to the emission estimates obtained from the technologies. In general, data from the measurement technologies, the underlying emission rate, and auxiliary measurement parameters are viewed as random variables that obey probability density functions (PDFs) related by Bayes' equation. In scenarios where a detailed measurement model is available, as is the case for truck based TDLAS measurements, the measurement noise, uncertain model parameters, and model errors may be propagated via Bayes equation to develop a posterior PDF for the true emission rate conditional on the measurement data and any prior information about the expected emission rate.

More generally, however, the measurement model may be highly complex or unavailable to the analyst. In this case, a likelihood function for observing an indicated emission conditional of an underlying true emission may be developed through using controlled release data obtained under relevant conditions. This likelihood may then be inverted and combined with any available prior information about the expected emission rate to obtain a posterior PDF for the emission rate. The posterior PDF amounts to a comprehensive statement about what is known about a given emission rate based on the measurement, uncertain model parameters, and any prior information.

6.2 Ongoing work

6.2.1 Incorporating uncertainty estimates into Monte Carlo simulations

Ongoing work is focused on incorporating the empirical statistical model in Sec. 5.2 into the Monte Carlo simulations used to develop emission inventories. For example, most Monte Carlo FEMP simulations only consider the emission rate, detection probability, and the probability and effectiveness of repair/intervention as statistical parameters to generate an expected emission reduction (e.g., kg/yr). Incorporating the uncertainty attached to the emission estimate into these simulations would provide an uncertainty in the expected emission rate over the duration of the simulation. This information is crucial when designing a FEMP, enabling the designer to determine an acceptable tradeoff between cost and the *certainty* with which the emissions will be reduced. Similarly, measurement uncertainty should be incorporated into Monte Carlo estimate of overall emission inventories, in a manner similar to Johnson et al. [29]. This information may partially explain disagreements between "bottom up" and "top down" inventories.

6.2.2 Development of the LWIR HS measurement system

While most of the technologies considered in this study have been deployed commercially for at least three years, this is not the case for the LWIR HS measurement system. The outcome of the third field trial (Sec. 4.3) highlight the capability of this system for detecting methane emissions, but also that the estimated emission rates are highly biased. UW personnel are currently working with Telops to understand and correct the origin of this bias, which could be due either to the spectroscopic model used to infer the methane column density (or "pixel mass") from each spectrum, or the advection model that combines the pixel mass map and wind speed into an overall emissions estimate.

6.3 Recommendations for future research

Further controlled release campaign data could also elucidate how some of the uncontrolled experimental parameters impact uncertainty. The controlled release data developed from the field campaigns was "pooled" to provide an overall likelihood estimate for a given measurement system, which amounts to marginalizing over the environmental conditions encountered at the site. Further field measurements could be used to develop narrower conditional PDFs, since, in general, incorporating more information into the estimate, e.g., a specific wind speed, should act to reduce the uncertainty interval.

While this study endeavored to generate industrially relevant scenarios, there were still significant differences between the controlled releases that could be conducted at CMC's Newall County facility and what may be encountered at an upstream site. Most importantly, the releases were "single-blinded" in that the emissions had a single source, and the operators knew where the source was located. More realistic and challenging scenarios could be developed at other facilities, including the METEC facility in Fort Collins CO, and the NGIF Emissions Testing Centre in Edson AB.

7. REFERENCES

[1] *Canadian Environmental Protection Act (CEPA) S.C. 1999, c. 33.*

[2] Environment and Climate Change Canada, "A healthy environment and a healthy economy: Canada's strengthened climate plan to create jobs and support people, communities, and the planet," ECCC, Ottawa, 2020.

[3] Environment and Climate Change Canada, "Review of Canada's methane regulations for the upstream oil and gas sector," ECCC, Ottawa, 2021.

[4] Environment and Climate Change Canada, "Proposed amendments to the Federal methane regulations for the oil and gas sector: Technical backgrounder," ECCC, Ottawa, 2023.

[5] D. R. Caulton, Q. Li, E. Bou-Zeid, J. P. Fitts, L. M. Golston, D. Pan, J. Lu, H. M. Lane, B. Buchholz, X. Guo, J. McSpiritt, L. Wendt and M. A. Zondlo, "Quantifying uncertainties from mobile-laboratory-derived emissions of well pads using inverse Gaussian methods," *Atmos. Chem. Phys.,* **18**, 15145, 2018.

[6] B. Kura and A. Jilla, "Feasibility of the inverse-dispersion model for quantifying drydock emissions," *Atmosphere,* **10**, 328, 2019.

[7] B. Crenna, "An Introduction to WindTrax", Department of Earth and Atmospheric Science, the University of Alberta, Calgary, 2006.

[8] C. B. Alden, S. C. Coburn, R. J. Wright, E. Baumann, K. Cossel, E. Perez, E. Hoenig, K. Prasad, I. Coddington and G. B. Reiker, "Single-blind quantification of natural gas leaks from 1 km distance using frequency combs," *Environ. Sci. Technol.,* **53**, 2908, 2019.

[9] M. R. Johnson, D. R. Tyner and A. J. Szerkes, "Blinded evaluation of airborne methane source detection using Bridger Photonics LiDAR," *Remote Sens. Environ.,* **259**, 112418, 2021.

[10] D. Hollenbeck, D. Zulevic and Y. Chen, "Advanced leak detection and quantification of methane emissions using sUAS," *Drones,* **5**, 117, 2921.

[11] D. J. Varon, J. McKeever, D. Jervis, J. D. Maasakkers, S. Pandey, S. Houweling, A. I., T. Scarpelli and D. J. Jacob, "Satellite discovery of anomalously large methane point sources from oil/gas production," *Geophys. Res. Lett.,* **46**, 13507, 2019.

[12] D. J. Varon, D. J. Jacob, J. McKeever, D. Jervis, B. O. A. Durak, Y. Xia and Y. Huang, "Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes," *Atmos. Meas. Tech.,* **11**, 5673, 2018.

[13] D. Zimmerle, T. Vaughn, C. Bell, K. Bennett, P. Deshmukh and E. Thoma, "Detection Limits of Optical Gas Imaging for Natural Gas Leak Detection in Realistic Controlled Conditions," *Environ. Sci. Technol.,* **54**, 11506, 2020.

[14] D. Risk, K. MacKay, M. Laboie and E. Bourlon, "Methane Emissions Data Aggregation and Analysis Project for FEMP-EA and Canadian context: Resources, patterns, and measurement methodology performance," PTAC, Calgary, Alberta, 2019.

[15] Alberta Energy Regulator, "Directive 60: Upsteam Petroleum Industry Flaring, Incinerating, and Venting," AER, Calgary, 2022.

[16] B. Horn and B. G. Schunck, "Determining Optical Flow," *Artif. Intell.,* **17**, 185, 1981.

[17] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. International Joint Conference on Artificial Intelligence*, Vancouver BC, 1981.

[18] Y. Zeng and J. Morris, Interviewees, *Virtual Meeting with Providence Photonics.* [Interview]. 19 December 2022.

[19] OpenCV, [Online]. Available: https://opencv.org/. [Accessed 25 January 2023].

[20] K. J. Daun, C. Lemieux, A. Beliveau, D. Blackmore, M. Nagorski, K.-M. Shim, P. Lapeyre, A. Wigle, J. Johnson, K. Fritz and K. Osadetz, "Evaluation of Emission Quantification Technologies: Interim Report," University of Waterloo, Waterloo, 2022.

[21] K. J. Daun, C. Lemieux, A. Béliveau, D. Blackmore, D. Nagorski, P. Lapeyre, A. Wigle, K. Fritz and K. Osadetz, "Evaluation of Emission Quantification Technologies: Second Interim Report," University of Waterloo, Waterloo, 2023.

[22] W. Shaw, "Measuring methane emissions using satellite and airborne instruments: Toward a tiered observation system," *Proc. EnviroTech*, Calgary, 2019.

[23] T. A. Fox, M. Gao, T. E. Barchyn, Y. L. Jamin and C. H. Hugenholtz, "An agent-based model for estimating emissions reduction equivalence among leak detection and repair programs," *J. Clean. Prod.,* **282**, 125237, 2021.

[24] X. Zhou, A. Montazeri and J. D. Albertson, "Mobile sensing of point-source gas emissions using Bayesian inference: An empirical examination of the likelihood function," *Atmos. Environ.,* **218**, 116981, 2019.

[25] X. Zhou, F. H. Passow, J. Rudek, J. C. von Fisher, S. P. Hamburg and J. D. Albertson, "Estimation of methane emissions fom the U.S. ammonia fertilizer industry using a mobile sensing approach," *Elementa-Sci. Anthrop.,* **7**, 19, 2019.

[26] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin, Bayesian *Data Analysis, 3rd Ed*., Boca Raton FL: CRC Press, 2014.

[27] M. Cassiana, M. B. Bertagni, M. Marro and P. Salizzoni, "Concentration fluctuations from localized atmospheric releases," *Boundary Layer Meterol.,* **177**, 461, 2020.

[28] B. M. Conrad, D. R. Tyner and M. R. Johnson, "Robust probabilities of detectuion and quantification uncertainty for aerial methane detection: Examples for three airborne technologies," *Remote Sens. Environ.,* **288**, 113499, 2023.

[29] M. R. Johnson, B. M. Conrad and D. R. Tyner, "Creating measurement-based oil and gas sector methane inventories using source-resolved aerial surveys," *Comm. Earth Environ.,* **4**, 139, 2023.

APPENDIX A

"Emission and Applications of Uncertainty in methane Emissions Quantification Technologies: A Bayesian Approach."

This is a draft of a journal manuscript that defines the empirical uncertainty quantification technique described in Sec. 5.2.

Estimation and Applications of Uncertainty in Methane Emissions Quantification Technologies: A Bayesian Approach

Augustine Wigle[a,*], Audrey Béliveau[a], Daniel Blackmore[b], Paule Lapeyre[b], Kirk Osaditz[c], Christiane Lemieux[a], Kyle J. Daun[b]

[a]*Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada*
[b]*Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, Canada*
[c]*Carbon Management Canada, Calgary, Alberta*

Abstract

An accurate understanding of uncertainty is needed to properly interpret methane emission estimates from the upstream oil and gas sector in a variety of contexts, from component-level measurements to yearly industry-wide inventories. One possibility is to derive an uncertainty estimate from the physical model that connects the measurement data to the emission estimates directly, but this information is often proprietary and thus unavailable to end users. Instead, we provide a method to develop probability distributions of measurements given a true emission rate empirically using controlled release data. This method is completely technology-agnostic, and provides a route to summarise uncertainty without the need to release proprietary modelling or data. To demonstrate the wide applicability of the method, we introduce an algorithm that can be used to synthesize the uncertainty model and measurement-based surveys to produce an uncertainty range for new measurements in the field.

Keywords: methane, uncertainty quantification

1. Introduction

Deep and rapid reductions in methane emissions from leading anthropogenic sources, especially upstream oil and gas activities, are crucial in order to avoid the worst outcomes of climate change [1], but doing this requires instrumentation that can reliably detect and quantify these emissions. Technologies for doing this include: quantitative optical gas imaging (QOGI) using mid-wavelength infrared (MWIR) cameras [2, 3, 4]; stationary [5] and mobile [6, 7, 8] methane

*Corresponding author
Email address: amhwigle@uwaterloo.ca (Augustine Wigle)

concentration sensors; and airborne [9, 10] and satellite-based [11] measurements. All of these systems utilize a measurement model that relates direct observations and auxiliary inputs to the methane emission rate. Often the measurement model consists of a spectroscopic sub-model that connects some radiometric measurement to a column density (ppm$\times$m) or path-averaged concentration estimate (ppm), and an advection sub-model, usually informed using anemometry data. The output of the inversion procedure is typically a point estimate of the methane emission rate from the source (e.g., kg/hr).

Emissions estimates can only be interpreted properly in the context of uncertainty. This aspect is particularly important in view of existing and emerging methane emissions regulations and reduction commitments [1, 12, 13], e.g., to answer the question "with what probability is this facility compliant with a particular regulation?" Methane leak detection and repair (LDAR) programs should also be optimized with quantification uncertainty in mind to give the best trade-off between cost and emissions reductions, as it has been shown that high quantification uncertainty contributes to certain types of LDAR programs being less cost effective [14]. Further, methane emissions measurements are used to develop broader jurisdiction-wide and global inventories [15, 16, 9], which are needed to assess progress towards emissions reduction targets and to inform policies and regulations, but these decisions can only be made in the context of uncertainty. Therefore, there is a need for transparent techniques for estimating emissions uncertainty that can be applied consistently in different contexts.

Approaches for quantifying methane emission uncertainty may be categorized as either physics-based or data-driven. Physics-based approaches address uncertainty associated with measurement noise, uncertain model inputs, and, especially, the model errors induced by the approximations and simplifications needed to derive a tractable measurement model, in an explicit way. As an example, Montazeri et al. [17] derive formulas for different error components of QOGI estimates, with the aid of virtual data generated from a computational-fluid dynamics large eddy simulation (CFD-LES). Caultron et al. [8] developed uncertainties for emission estimates obtained from a truck-mounted concentration sensor and inverse Gaussian plume model by accounting for uncertainty in the Gaussian model diffusion coefficient, emission source and height, and wind speed and stability class. Cambaliza et al. [18] developed uncertainties for emission estimates inferred from aircraft-based concentration measurements using different values for the background carbon dioxide and methane, depth, changes in the convective boundary layer height, and perpendicular wind speed parameters.

While physics-based approaches provide key insights into the uncertainty of methane emission estimates obtained from various technologies, and how they should be deployed to minimize these uncertainties, they also have several key drawbacks. First, they require detailed knowledge of the measurement model, which may be very complex or unavailable due to proprietary aspects of the technology. Second, the results of a physics-based uncertainty analysis are specific to a given technology and will not be broadly applicable, requiring cumbersome effort for every technology of interest. Third, results of a purely

physics-based uncertainty estimate may not agree with what is observed in real-world scenarios due to missing or inadequate modelling of uncertainty sources. Moreover, existing physics-based uncertainty analyses do not include methods or procedures for how the results should be applied in practice, and there is a lack of consistency in reporting of the results [8]. For example, most physics-based approaches do not show how their results should be used to derive a 95% confidence interval based on a given measurement from the technology.

Empirical approaches to uncertainty quantification rely on a statistical model which compares true and measured emission rates from controlled-release data. The statistical model can then be used to predict future measurements given a true emission rate, or inverted to give a confidence interval for the true emission rate given a measurement. Empirical approaches have two key benefits over physics-driven approaches: 1) they are data-driven, meaning the results will likely resemble what is actually observed in the field, and 2) the statistical framework can be leveraged to provide unified and consistent guidelines for how the results of the uncertainty analysis should be used in practice.

Empirical approaches require data from controlled-release field trials. To this end, many single-blinded or double-blinded field trials have been conducted with the goal of assessing the performance of methane emissions quantification technology, e.g., [19, 20, 21, 10, 22]. However, empirical approaches employed on these data have mainly been limited to simple linear regression approaches which do not allow for quantification uncertainty to vary with the emission rates [19, 20, 21, 10], with the exception of [22], who provide an approach to derive the distribution of the true emission rate given a measurement from an airborne methane detection and quantification technology.

In this work, we introduce a flexible empirical framework to elucidate quantification uncertainty that can be applied to any technology modality and illustrate its use using field trial data from two campaigns carried out using four methane detection and quantification technologies as well as controlled release data reported by [22]. The empirical framework allows for the derivation of two important probability distributions: 1) the distribution of measurements given the true emission rate and 2) the distribution of the true emission rate given a measurement. The first distribution is a building block to the second distribution, and has the potential to be incorporated into simulation software that models LDAR programs such as FEAST, LDAR-Sim, and AROFemp [23, 24, 25]. The second distribution is an important input to simulation methods used to derive measurement-based inventory estimates such as [26]. Our approach to deriving the second distribution also provides the opportunity to incorporate context-specific information into the analysis, such as knowledge of the emission rate distribution in a given region. The results of the analysis are data-driven and the design of the field trials allows for the assessment of the potential real-world effectiveness of the uncertainty quantification results.

Figure 1: Technologies studied in the field trials. Top row, left to right: QOGI A, QOGI B, QOGI C. Bottom row, left to right: airborne NIR HSI, truck-based TDLAS.

2. Materials and Methods

2.1. Methane Quantification Technologies

We demonstrate the analysis procedure using controlled release data from four methane quantification technologies, three of which were evaluated in the field trials described in Sec. 2.2: QOGI; truck-mounted tunable diode laser-absorption spectroscopy (TDLAS); and airborne near-infrared hyperspectral (NIR HS) imaging. We also consider an airborne TDLAS system ("Gas Mapping LiDAR™ " (GML) from Bridger Photonics, Inc.) based on data reported in Conrad et al. [22]. Examples of the technologies investigated in the field trials are shown in figure 1.

2.1.1. Quantitative Optical Gas Imaging (QOGI)

QOGI systems are almost exclusively based on a mid-wavelength infrared (MWIR) camera that contains a cold filter centered on the 3.34 μm methane vibrational-rotational band. The intensity entering the camera aperture is imaged through the cold-filter and onto a focal plane array (FPA) that produces a pixel intensity. The cameras are usually calibrated to generate a spectrally-averaged absolute intensity along each pixel line-of-sight. The camera data is then analyzed in near real-time by software on a peripheral tablet. The measurement model is composed of a spectroscopic sub-model that generates a column density map of the gas, and an advection model that infers a 2D projected velocity field from the apparent plume motion between successive images. These quantities are then combined to obtain a mass flow rate (e.g, kg/s).

The reliability of QOGI-derived emission estimates depends on factors that include measurement distance between the plume and the camera, thermal

contrast between the plume and the background, wind speed, and leak rate [27, 19, 28]. Identifying favorable measurement scenarios draws considerably on operator experience [29]. Three QOGI systems were deployed by three operators of varying experience, as summarized in Table 1.

Table 1: QOGI operators and equipment

Operator	**Experience**	**System**	**Field Trial**
A	Professional	FLIR GF320 with Providence QL320 (v. 3.0.0.5)	1
B	Professional, new to system	OPGAL EyeCGas (v. 1.0.24)	2
C	Novice	FLIR GFx320 with FLIR QL320 (v. 1.4.1)	1 & 2

QOGI Operator A used a FLIR GF320™ camera with a Providence QL320™ tablet (v. 3.0.0.5); QOGI Operator B used the OPGAL EyeCGas™ (v. 1.0.24) and QOGI Operator C used a FLIR GFx320™ camera with the FLIR QL320™ Tablet (v. 1.4.1). Notably, while Operator B was an experienced QOGI operator, they were unfamiliar with the OPGAL system during the field trial. QOGI Operator A was highly experienced and familiar with their equipment, while QOGI Operator C was a novice, having less than six months of experience with the system.

2.1.2. Truck-mounted TDLAS

Methane releases were also quantified using a truck-mounted TDLAS system (Boreal Laser GasFinder 3 VB™). The absorptance, and therefore methane column density (e.g., ppm·m), is inferred through wavelength-modulation spectroscopy (WMS) [30] and then converted to a path-average concentration (ppm). The truck traversed the plume at distances ranging from 50 to 100 m downwind of the release point. Methane concentrations were measured at one second intervals; these concentrations and wind speeds obtained from an ultrasonic anemometer operated by the service provider were then processed using a backwards Lagrangian stochastic quantification algorithm [31, 32] to obtain a release estimate for each plume transect.

2.1.3. Airborne NIR HS imaging (GHGSat-AV™)

The airborne NIR HS system (GHGSat-AV™) consists of a downward-looking wide-angle Fabry-Perot imaging Fourier transform spectrometer that operates between 1630-1655 nm [33], mounted inside an aircraft [34]. The aircraft overflew the releases at an approximate altitude of 250 m above ground level and airspeed of 240 km/hr. Thermal emission from the gas and ground is negligible over this wavelength range; instead, the camera images sunlight transmitted through the atmosphere, reflected from the ground, and transmitted back to the camera. The methane column density is inferred from the attenuation of the transmitted light via a multi-layer spectroscopic model, and then combined

with an advection model [35] using wind data from an online weather model to find the emission rate.

2.1.4. Airborne TDLAS

Bridger's airborne GML™ system consists of two tunable diode lasers, and a sensor that detects the ground-reflected laser light. One laser is used for range finding and determining ground reflectivity, while the other scans the 1651 nm CH_4 absorption line to determine column density. The lasers move in a conical pattern, which forms an ellipsoidal swath on the ground. Reflected light from the range-finding and methane-absorbing lasers are combined to form a column density via WMS. The column density estimates across the swath are used to form a 3D plume concentration map, which is combined with an advection model using wind speed from online weather data to obtain a release rate [20].

2.2. Field Trial Design and Execution

The QOGI, truck-mounted TDLAS, and airborne NIR HS imaging systems were evaluated through two controlled release field campaigns executed at Carbon Management Canada (CMC)'s Newall County Research Station near Brooks, Alberta, the first during April 20-26 2022 and the second during September 25-October 1, 2022. Technology developers and service providers were invited to attend the field trials and quantify emissions of natural gas in a variety of industrially relevant scenarios, including 1.7 m, 3.4 m, and 4.8 tall stacks, and a 14-m tall unlit flare, following their standards-of-practice.

An assay showed that the natural gas consisted of 94.2 % methane, 3.4 % ethane, 1.1 % propane, and 1.3 % minor components, predominantly N_2 and O_2 in roughly atmospheric abundances. The gas was released from a compressed cylinder via a regulator valve and flowed through a heat exchanger to condition the gas to atmospheric temperature. The conditioned gas then passed through a mass flow controller (Alicat MCR-2000SLPM - D-PAR) and discharged to the atmosphere in a manner that depended on the release scenario as shown in figure 2.

Local meteorological conditions were measured using a portable 81000-L RM Young 3D ultrasonic anemometer and a Davis WeatherLink Pro+™ weather station. The portable anemometer was located 2.25 m above the ground. Background measurements of methane and other relevant species were monitored throughout the tests using a Picarro cavity ringdown spectrometer located in a structure approximately 250 m from the release locations. Background methane measurements were between 2 to 2.5 ppm throughout the tests. Meteorology and controlled release data for these field trials are provided in the supplemental information (SI).

Service providers did not have access to meteorology data; instead, they conducted their own on-site measurements or relied on third-party weather models, as they would when deploying the technology in a practical scenario. Service providers then compiled their own estimates and provided them to the academic team.

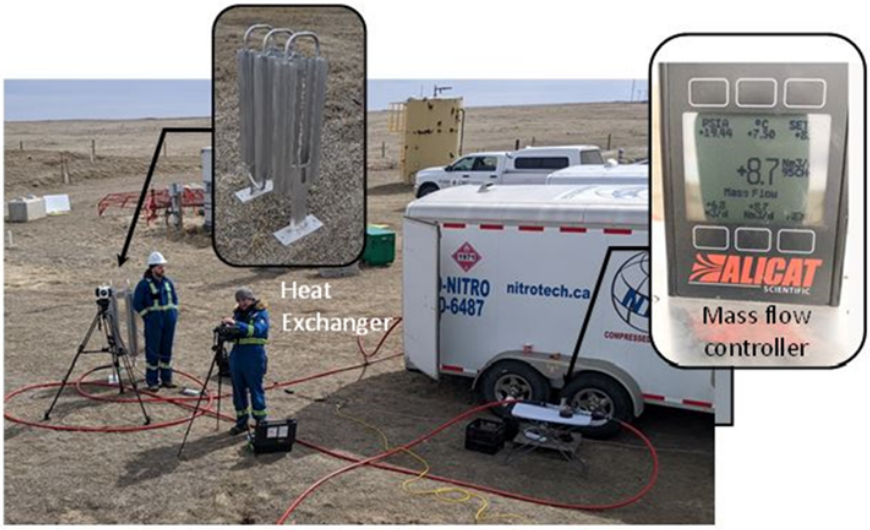

Figure 2: Controlled release set-up including heat exchanger and mass flow controller.

Additional data was taken from controlled release studies reported in the literature: Bridger GML system was taken from the controlled, fully-blinded release study reported by Conrad, et al. [22].

Table 2: Summary of technologies, providers, and available data from the field trials and external sources. N_1 and N_2 refer to the number of observations collected for a given technology during the first field trial and second field trial, respectively."Other" refers to data from Ref. [22].

Technology	N_1	N_2	Other
QOGI Operator A	117	0	NA
QOGI Operator B	0	71	NA
QOGI Operator C	14	106	NA
Truck TDLAS	142	125	NA
Aerial TDLAS	NA	NA	405
Aerial NIR HSI	46	37	NA

3. Uncertainty Quantification

3.1. Models for Uncertainty Using Controlled Release Data

We propose a statistical model that answers the following question: For a given true emission rate, what range of measurements could be expected, given the observed controlled release data, in the context of model error and measurement noise? We take a Bayesian approach to fitting the model and thus in Sec. 3.1.2 we give a brief overview of Bayesian analysis, followed by technical details

that include the prior distributions, Sec. 3.1.3, and model selection methods, Sec. 3.1.4.

3.1.1. Novel Flexible Model

Let Q_i be the true emission rate corresponding to the ith observation in the field trial, and M_i be the emission rate estimated by the technology for the ith observation, $i = 1, \ldots, n$ where n is the total number of observations for the given technology. The relationship between Q_i and the bias and variability of M_i can be complicated, since both the bias and the variability may change over the range of Q_i. Additionally, the relationship between Q_i and M_i may not be strictly linear, as shown in e.g. figure 3. The model must also account for the fact that all technologies may report a "false positive", that is, estimating a non-zero M_i when $Q_i = 0$.

A flexible likelihood which allows the bias and variance of M_i to vary with Q_i and incorporates the possibility of false positives is given by

$$\log(M_i) = \log(\phi_i) + \epsilon_i, \tag{1}$$

where

$$\epsilon_i \sim \mathrm{N}(0, \sigma_i^2)$$

and

$$\phi_i = \mathrm{median}(M_i).$$

A specification for ϕ_i is

$$\phi_i = \begin{cases} \alpha_0 + \alpha_1 Q_i + \alpha_2 Q_i^2 & Q_i \leq \gamma; \\ \alpha_0 + \beta_0 + (\alpha_1 + \beta_1) Q_i & Q_i > \gamma. \end{cases}$$

To ensure that the function is continuous at $Q_i = \gamma$, we impose the restriction that $\beta_0 = \alpha_2 \gamma^2 - \beta_1 \gamma$. In this likelihood, there is a quadratic relationship between ϕ_i and Q_i for values of Q_i smaller than a threshold γ and a linear relationship for larger values of Q_i. This likelihood is normal on the log scale, which corresponds to a log-normal likelihood on the measurement scale.

The specification of ϕ_i can be modified to give the best prediction results and fit to the data. For example, the threshold parameter γ, β_0 and β_1 could be removed, which would give a quadratic relationship over the whole range of Q_i. Table 3 summarizes the parameters that may be removed and Section 3.1.4 shows how the likelihood is chosen.

The model can be rewritten to facilitate interpretation by exponentiating both sides of Eq. (1):

$$M_i = (\phi_i) \times e^{\epsilon_i}, \tag{2}$$

where ϕ_i is the median measurement for a true emission rate of Q_i.

The likelihood in Eq.(1) is an extension of the scheme proposed by Conrad et al. [22]. That is, their model is a special case of our likelihood where $\alpha_0 = 0$, $\alpha_2 = 0$, $\beta_0 = 0$, $\beta_1 = 0$, and $\sigma_i^2 = \sigma^2$ for all $i = 1, \ldots, n$. Briefly, they assume the median value of M_i has a multiplicative relationship with Q_i, that is,

$\phi_i = \alpha_1 \times Q_i$, and the multiplicative error term has a constant variance over the range of Q_i. Our model expands on this in three ways. First, we allow for linear and quadratic relationships between the median measurement and Q_i rather than a strictly multiplicative one. This is useful in modelling more complex relationships. Additionally, it allows false positives to be modelled via α_0. Second, the inclusion of the threshold parameter t allows more flexibility in modelling the relationship between the median of M_i over the range of Q_i rather than assuming a common median function for all Q_i. Third, we investigate different variance structures for ϵ_i which can allow the variance to change with Q_i to more accurately model the patterns observed in controlled data from some instrumentation rather than assuming a constant variance. Another difference between this approach and that of [22] is that they investigate different distributions for the error term, whereas we restrict ourselves to the log-normal distribution, but investigate different forms for the median and variance which are motivated by the data. Finally, we take a fully Bayesian approach to estimation and inference discussed in section 3.1.2 whereas [22] uses a (frequentist) maximum likelihood approach.

It is important to note that in this model the errors are additive on the log scale, which implies multiplicative errors on the raw measurement scale as shown in Eq. (2). The simplest way to model the variation is to set the variance of ϵ_i to a constant, $\sigma_i^2 = \tau^{-1}$ for all i, where τ is referred to as the precision parameter. Multiplicative errors may be suitable for lower and moderately-sized emission rates, but for large values of Q_i purely multiplicative errors may overestimate variability for some technologies. A possible explanation for this is that for smaller emission rates, both the error in raw concentration or column density estimates and error in the advection model are significant, leading to product uncertainty. For larger emission rates, however, either raw measurement error or advection model error dominates, leading to sub-multiplicative errors in this range. To accommodate this, we also propose using $\sigma_i^2 = (\tau + Q_i/\eta)^{-1}$ as an alternative variance structure for ϵ_i which allows the variability of the error terms to decrease with increasing Q_i. The approach to choosing an appropriate likelihood, including the variance specification, is described in Sec. 3.1.4.

This section concludes with interpretations of the parameters in the likelihood. Threshold parameter γ allows the linear relationship to change for larger values of Q_i. α_0 represents the median measurement when the true emission rate is zero, which accounts for false positives. Parameters α_1 and α_2 are the coefficients for Q_i and Q_i^2, describing the quadratic relationship between the median measurement and the true emission rate when the true emission rate is less than the threshold. Sums $\alpha_0 + \beta_0$ and $\alpha_1 + \beta_1$ are the slope and intercept of the linear relationship between the median measurement and Q_i when Q_i exceeds the threshold. Some technologies exhibit simpler relationships between Q_i and the median of M_i, in which case some or all of γ, α_2, β_0, and β_1 may be dropped from the model. Finally, two different variance specifications were used in the likelihood: either $\sigma_i^2 = \tau^{-1}$ or $\sigma_i^2 = (\tau + Q_i/\eta)^{-1}$. In the former case, τ^{-1} is the variance of all measurements on the log scale. In the latter specification, τ^{-1} is the variance of the measurements when $Q_i = 0$ and c controls how

much the variance changes as Q_i increases from zero. A larger value of η corresponds to a milder reduction in variance as Q_i increases. Table 3 summarizes the parameters in the likelihood, their units, and scenarios in which they may be included or removed.

3.1.2. Bayesian Analysis

The model parameters are estimated using a Bayesian approach. This is done for several reasons: (i) the method is flexible, allowing the likelihood to be tailored to the data; (ii) data can be synthesized seamlessly from multiple sources (e.g., multiple measurement campaigns or different measurement modalities); (iii) it explicates the use of prior information; and (iv) it provides the full probability distribution of measurements given a true emission rate, amounting to a comprehensive definition of what is known about the emission rate (see Sec. 4.2).

Let the unknown parameters of a statistical model for the measurement M_i given a fixed true emission rate Q_i be represented by vector $\underline{\theta}$. In the Bayesian framework, these parameters are envisioned as random variables defined by probability distributions that are related by Bayes' equation,

$$p(\underline{\theta} \mid \underline{M}) = \frac{p(\underline{M} \mid \underline{\theta}) p(\underline{\theta})}{p(\underline{M})}, \tag{3}$$

where $\underline{M} = (M_1, M_2, \ldots, M_n)'$. The probability distribution of interest is the *posterior* distribution, $p(\underline{\theta} \mid \underline{M})$, which summarises the information about $\underline{\theta}$ in the observed measurements along with any external or "prior" knowledge we may have about $\underline{\theta}$. The likelihood distribution, $p(\underline{M} \mid \underline{\theta})$, is the probability distribution of the set of measurements for a fixed value of $\underline{\theta}$ and is also a function of the true emission rates $\underline{Q}$, which is fixed and thus suppressed in the notation in Eq. (3). It describes how likely it is to observe $\underline{M}$ for a given value of $\underline{\theta}$ and $\underline{Q}$ in the context of measurement noise and model error. Under the assumption that measurement errors are independent between measurements, we also have that $p(\underline{M} \mid \underline{\theta}) = \prod_{i=1}^{n} p(M_i \mid \underline{\theta})$. The prior distribution, $p(\underline{\theta})$, describes what is known about the unknown parameters before data is collected. Prior distributions for the proposed models are discussed in Sec. 3.1.3. Finally, $p(\underline{M})$ is the marginal distribution of the data, which is constant for fixed data.

The posterior $p(\underline{\theta} \mid \underline{M})$ is estimated using Markov Chain Monte Carlo (MCMC) sampling, which generates a set of samples from the posterior distribution of $\underline{\theta}$ [36]. These samples are readily used to derive quantities of interest, such as credibility intervals. MCMC sampling is done using Just Another Gibbs Sampler (JAGS)[37] via the runjags R package [38].

We wish to understand the distribution of a measurement given a fixed value of Q_i. Let $\tilde{M}_i$ represent a new, unobserved measurement and Q_i be an accompanying true emission rate. Then we wish to find the posterior predictive distribution of $\tilde{M}_i$, given by $p(\tilde{M}_i \mid \underline{M}, Q_i) = \int p(\tilde{M}_i \mid \underline{\theta}, Q_i) p(\underline{\theta} \mid \underline{M}) d\underline{\theta}$, which is the integral over $\underline{\theta}$ of the likelihood of M_i given fixed $\underline{\theta}$ and Q_i times the posterior distribution of $\underline{\theta}$. This distribution can be obtained via simulation using samples from the posterior.

3.1.3. Prior Distributions

As with any Bayesian model, an appropriate prior distribution depends on the context of the problem at hand, including pre-existing knowledge and the scale of the data. We summarise the units of each parameter in Table 3.

Parameter	Units	Exclusion criteria
α_0	$[M_i]$	*
α_1	Unitless	None
α_2	$[M_i]^{-1}$	†
β_0	$[M_i]$	*,†
β_1	Unitless	†
γ	$[M_i]$	†
τ	Unitless	None
η	$[M_i]$	†

Table 3: Summary of the units for each parameter in the likelihood, important to keep in mind when specifying prior distributions. Exclusion criteria: * = may be removed if technology does not have false positives in controlled release data, † = may be removed if its removal leads to a simpler model which has adequate fit and predictive performance.

The parameters of the likelihood to be estimated are α_0, α_1, τ, along with optional parameters α_2, γ, β_1, and/or η. All parameters are assumed to be independent, that is,

$$p(\theta) = p(\alpha_0, \alpha_1, \tau, \alpha_2, \gamma, \beta_1, \eta) = p(\alpha_0)p(\alpha_1)p(\tau)p(\alpha_2)p(\gamma)p(\beta_1)p(\eta), \qquad (4)$$

so we can specify individual prior distributions for each parameter.

We can derive empirical prior distributions by considering the role of each parameter in the model and their units. Since α_0 is the median measurement when the true emission rate is zero, it should be small and non-negative. It is parameterized by a gamma distribution, where the shape and rate parameters can be chosen so the mean of the distribution is similar to the mean false positives observed in the data. For example, for the data reported by QOGI Operator A, the average false positive is 0.27 kg/h so we use a gamma distribution with shape parameter = 0.5 and rate parameter = 2, which has a mean of 0.25 and variance of 0.125. Note that β_0 is a function of other parameters to ensure that the piece-wise function is C^1 continuous. A prior is not specified for this parameter.

In a simple linear model, α_1 is the slope of Q_i so for every one unit increase in Q_i, the median measurement increases by α_1 units. A perfect technology would have $\alpha_1 = 1$. Thus we use a prior for α_1 which has a median of 1 and is non-negative. Further, we seek a distribution with the property that for any constant $k > 1$, the probability that $\alpha_1 > k$ should be the same as the probability that $\alpha_1 < 1/k$, or in other words, the probability that the technology over-estimates by a factor of k is the same as the probability that it under-estimates by a factor of $1/k$. This property is desirable in the prior because information about under-

or over-estimation should only come from the data. Thus we use a standard log-normal distribution (shape parameter equal to one, location parameter equal to zero, and scale parameter equal to zero [39]) because it has this property. For example, if α_1 follows the standard log-normal distribution the probability that $\alpha_1 < 1/2$ = probability that $\alpha_1 > 2 = 0.244$.

Coefficient α_2 is associated with Q_i^2 when $Q_i \leq \gamma$. Similarly to α_0 and α_1, we restrict this parameter to be non-negative to avoid taking the log of zero or a negative number. Due to the units of α_2, we use a prior with a relatively small variance. For example, if we wanted the variance to be equivalent to 10 kg/h, that equals 0.1 $[\text{kg/h}]^{-1}$. Thus we use a half-normal distribution with variance parameter equal to 1.

Parameter β_1 (unitless, nonnegative) represents the change in α_1 when $Q_i > \gamma$. As with α_1, we use a standard log-normal distribution as a prior for β_1.

The threshold parameter γ represents the value of Q_i for which the relationship with M_i changes from quadratic to linear. The parameter γ must be somewhere in the range of the Q_i data. We use a uniform prior distribution on (0, max) where max is determined by the largest value of Q_i observed in the data for a given technology.

Parameter τ represents either the inverse of the variance of measurements on the log scale in a constant variance model, or the inverse of the variance of measurements when $Q_i = 0$ on the log scale and is referred to as the precision parameter. We use a vague non-negative prior of a half-normal with variance parameter set to 100 on $\tau^{-1/2}$, as suggested in [40].

Finally, if the more complicated variance model is used, a prior must be chosen for η. Little external information is known about η except it must be non-negative. We use a half-normal with variance set to 100.

The sensitivity of results to prior specification was checked for all models. Results were obtained for the stated priors. Next, the model was refit with priors where the range and/or variance was changed for some parameters. The posterior distribution of each parameter was then compared between the two models. The resulting 90% prediction bands were also compared between the models. Unless otherwise stated in the results section, the model results were insensitive to the prior specification.

3.1.4. Model Selection

As discussed in Sec. 3.1.1, a variety of candidate models may be formed by adding or removing likelihood parameters, each of which may result in different implications for measurement bias and variability. For example, removing Q_i/η from the variance expression leads to a simpler model which has constant variance on the log scale. In general, a model with more parameters will fit the controlled release data better but may also be prone to over-fitting, which can result in poor predictive performance. Therefore, we use a combination of Deviance Information Criteria (DIC), prediction bands, and residual plots to select a model that provides a good trade-off between goodness of fit and complexity.

The Deviance Information Criterion, or DIC, combines goodness-of-fit to the training data and model complexity to provide an overall assessment of the

model [36]. It is analogous to the AIC, a frequentist model selection tool used by Conrad et al. [22] in the context of uncertainty modelling of methane quantification technologies. When comparing multiple models, a lower DIC value indicates a better balance between model fit and complexity, with differences of two or more considered meaningful [41].

We are also guided by plotting prediction bands derived from the posterior predictive distribution of measurements for different values of Q_i over a scatterplot of the data used to fit the model. If the prediction bands show a much wider or narrower spread than the data used to fit the model, this is an indication that the variance is not modelled well. Prediction bands can also be compared to additional data that was excluded from the model fit ("external data"), which indicates the generalizability of the model predictions. If the model predictions look similar to the external validation data, this is an encouraging sign that the model is suitable to be used under different conditions. Investigating the residuals, defined as the difference between the model-predicted value ($\hat{M}_i$) and the observed data point (M_i), that is, $\text{residual}_i = M_i - \hat{M}_i$ provides still more insight into areas of improvement for the model.

The DIC, prediction bands, and residual plots were used for model selection as follows: First, the simplest model possible with constant variance was fit to the data (a multiplicative model with α_1 if there are no false positives in the data or a linear model with α_0 and α_1 otherwise). The DIC was calculated using JAGS. Prediction bands were compared to the data used to fit the model and residual plots were inspected. If the prediction bands were much wider or narrower than the spread of the data, this indicated that the variance model should be explored. If the residual plots showed systematic problems, this indicated that the median specification should be explored. Models were then augmented as suggested by the diagnostic plots, refit, and DIC was re-calculated. This process was repeated until the diagnostic plots were satisfactory and the DIC was at least three less than that of the previous model.

4. Results and Discussion

4.1. Uncertainty Results

In this section, we present the selected likelihoods for five different methane quantification technologies/operators, discuss the performance, and comment on the generalizability of the model if applicable. The chosen models are summarized in Table 4.

4.1.1. QOGI Technologies

Prediction bands and posterior median predictions are shown for QOGI Operators A, B, and C in Figures 3, 4 and 5, respectively. All QOGI technologies underestimate emissions on average. The likelihoods from QOGI Operators B and C are best modelled using a quadratic function for the median below a small threshold, then a linear function. QOGI Operator A is best modelled using a quadratic function for the median. Data provided by Operator A has a more

Table 4: Summary of selected models for each methane quantification technology provider.

Technology	Selected likelihood	
	ϕ_i	σ_i^2
QOGI A	$\alpha_0 + \alpha_1 Q_i + \alpha_2 Q_i^2$	$(\tau + Q_i/\eta)^{-1}$
QOGI B	$\begin{cases} \alpha_0 + \alpha_1 Q_i + \alpha_2 Q_i^2 & Q_i \leq \gamma; \\ \alpha_0 + \beta_0 + (\alpha_1 + \beta_1) Q_i & Q_i > \gamma. \end{cases}$	τ^{-1}
QOGI C	$\begin{cases} \alpha_0 + \alpha_1 Q_i + \alpha_2 Q_i^2 & Q_i \leq \gamma; \\ \alpha_0 + \beta_0 + (\alpha_1 + \beta_1) Q_i & Q_i > \gamma. \end{cases}$	$(\tau + Q_i/\eta)^{-1}$
Truck TDLAS	$\alpha_0 + \alpha_1 Q_i$	$(\tau + Q_i/\eta)^{-1}$
Aerial TDLAS	$\alpha_1 Q_i$	$(\tau + Q_i/\eta)^{-1}$
Aerial NIR HSI	$\alpha_0 + \alpha_1 Q_i$	τ^{-1}

limited range than the other QOGI technologies - with a max Q_i value of 30 kg/h, compared to 80 kg/h for Operator C and 50 kg/h for Operator B. For QOGI technologies in general, the likelihood has more curvature in the lower range of Q_i while a linear relationship on the log scale is suitable for higher release rates.

The results for QOGI Operator B are quite distinct from those of Operators A and C. This may be attributed to this operator's lack of familiarity with the camera settings during the testing, as reported by the operator. This lack-of-familiarity manifests as an additional factor that influences (broadens and biases) the likelihood.

QOGI Operator C was present for both field trials. Only 14 measurements were made for this technology at the first field trial, which we use as external data. These data points fall within the 95% prediction band, suggesting that the model is generalizable.

4.1.2. TDLAS

Results from the selected models for truck and aerial TDLAS systems are shown in Figs. 6 and 7. The truck-based TDLAS tends to underestimate emissions, while the aerial technology overestimates on average.

For truck-based TDLAS, the model was fit using data from the second field trial, while data from the first field trial was used as external data to assess the model's generalizability. Most of the external data points fall within the prediction bands. However, the median trend appears different for the external data. A possible explanation for this is that weather conditions may have differed between trials one and two in such a way that estimates were systematically larger during the first trial than the second.

Figure 8 shows the predictions resulting from the Bayesian uncertainty model derived for the airborne TDLAS compared to the one presented by Conrad et al. [9]. The Bayesian model gives narrower prediction bands than the other model, which is particularly noticeable in the upper range of Q_i. This is likely due to the different variance specifications used in the models; the model in Ref. [22]

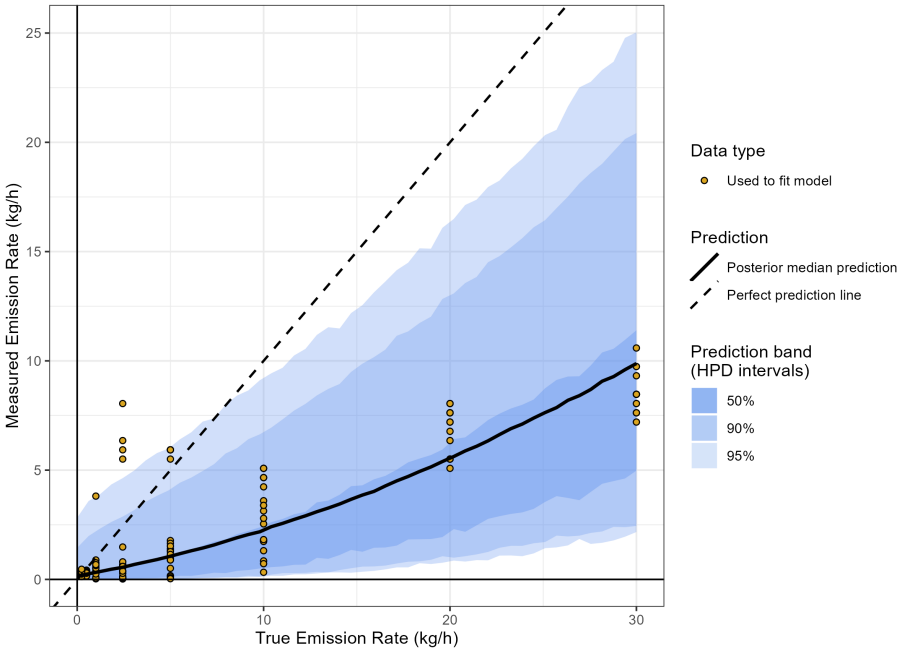

Figure 3: Uncertainty quantification model results for QOGI Operator A. The model was fit to data from the first field trial. No external data were available.

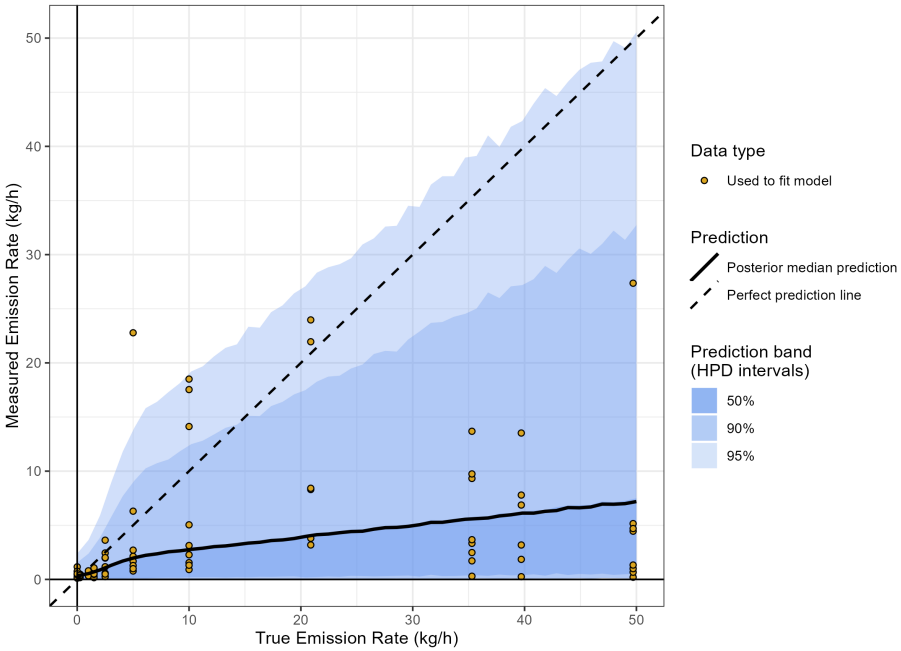

Figure 4: Uncertainty quantification model results for QOGI Operator B. The model was fit to data from the second field trial. No external data were available.

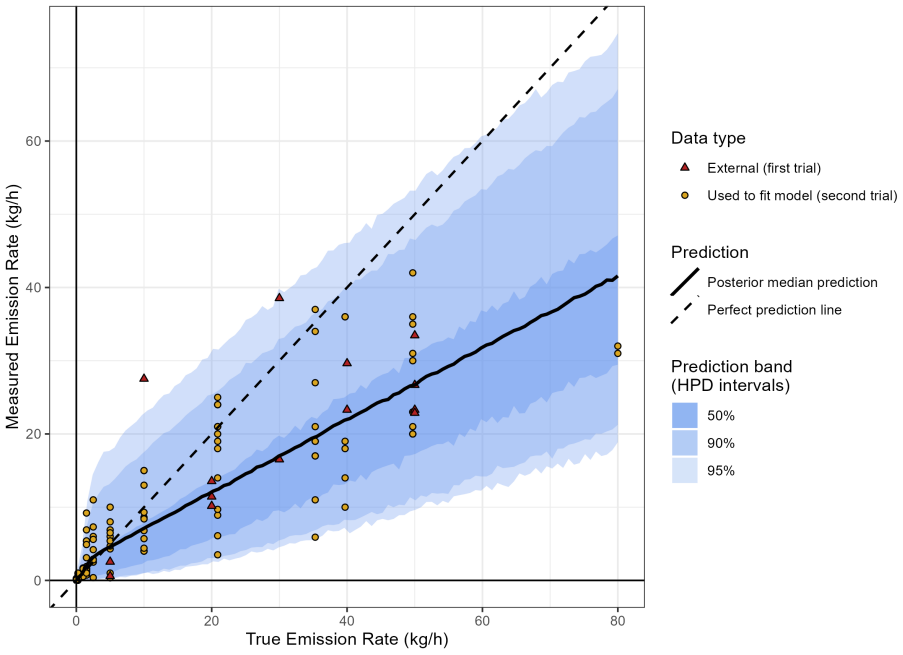

Figure 5: Uncertainty quantification model results from QOGI Operator C. The model was fit to data from the second field trial. Controlled release data from the first field trial was used as external data.

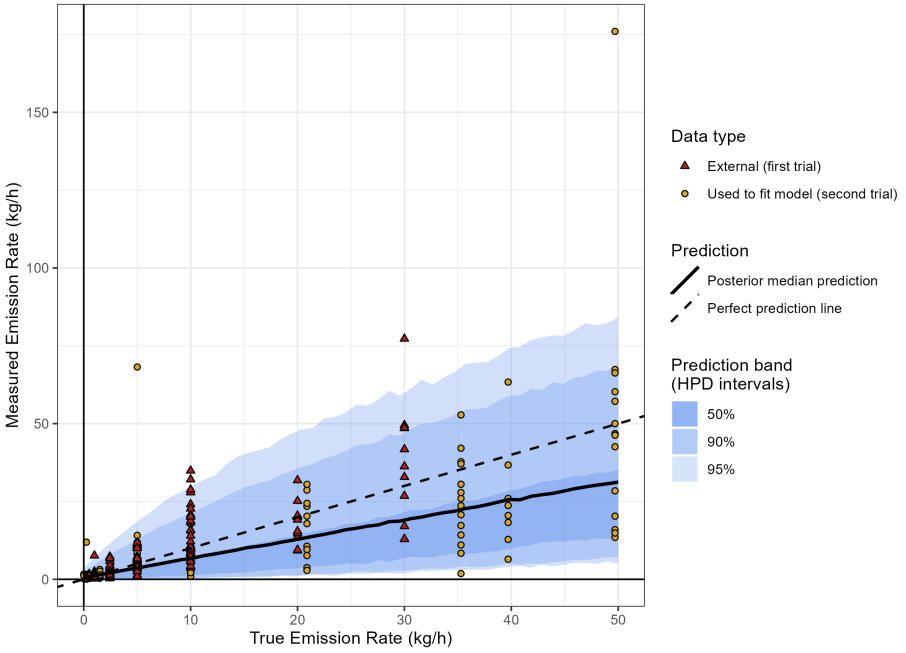

Figure 6: Uncertainty quantification model results for truck-based TDLAS data. The model was fit to data from the second field trial. Controlled release data from the first field trial was used as external data.

uses a constant variance whereas the selected model from our proposed Bayesian approach allows the variance to change with Q_i. The median predictions (solid black line) are very similar between the two models.

4.1.3. Airborne NIR HS Imaging

The prediction bands from the selected model for the airborne NIR HS imaging technology are shown in Fig. 9. The technology tends to overestimates emissions. The model was fit to data from the second field trial, while data from the first field trial were used as external data. There appear to be systemic differences between the data from the two field trials, with the data from the first field trial underestimating emissions more often, and data from the second field trial overestimating emissions more often. However, most of the external data points still fall within the 95% prediction bands, which is a positive indication of the applicability of the model. During the field trials, the operator remarked that the conditions were considered marginal due to excessive cloud cover and would not be typical of those under which commercial measurements were conducted, while those of the second field trial were nearly ideal. This could explain the lack of implausibly large estimates calculated during the first trial, as compared to the second field trial, where there two estimated measurements that

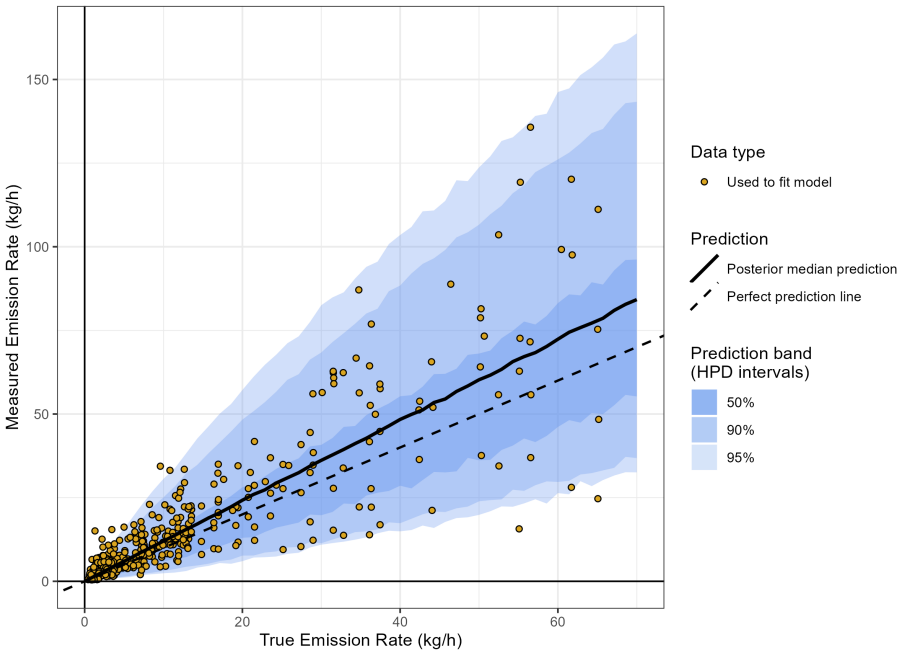

Figure 7: Uncertainty quantification model results for airborne TDLAS data provided by [22].

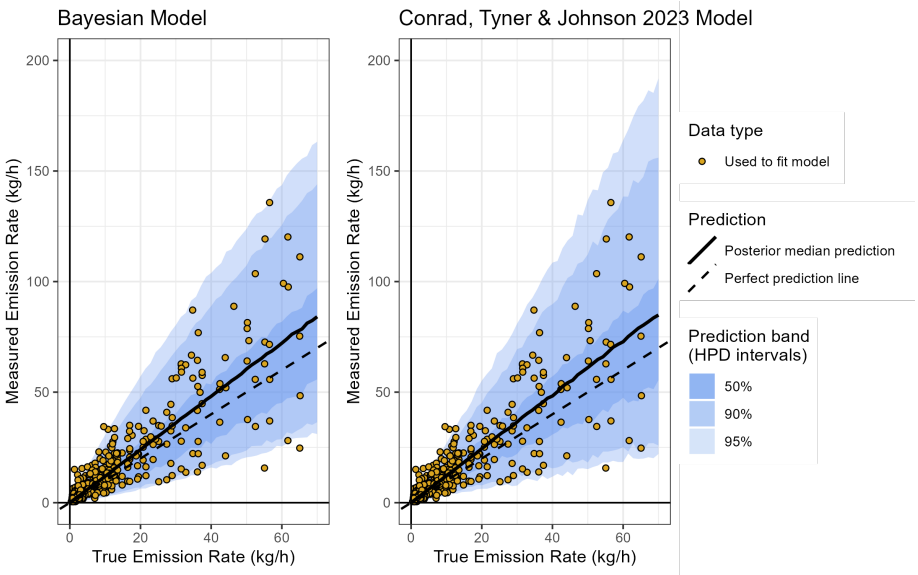

Figure 8: Side-by-side comparison of Bayesian model proposed in table 4 and that presented in [22] for airborne TDLAS.

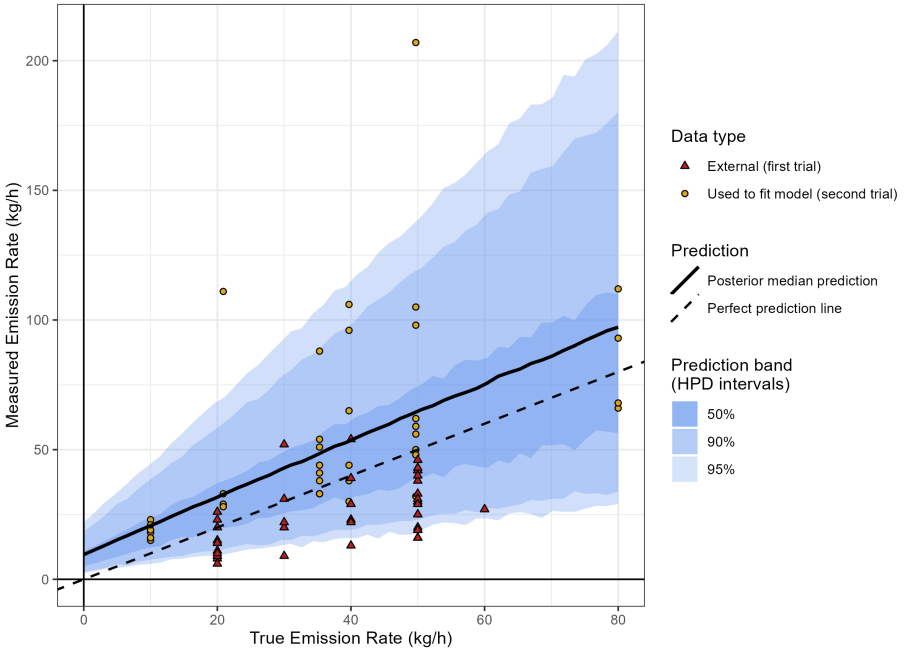

Figure 9: Uncertainty quantification model results for airborne NIR HS data. The model was fit to data from the second field trial. Controlled release data from the first field trial was used as external data.

are so large they fall outside the 95% prediction band.

4.2. Application: Quantifying Uncertainty in New Measurements

4.2.1. Algorithm

Finally, we describe a possible downstream application of the distributions derived in section 4.1. Suppose we wish to calculate a credible interval for the true emission rate based on methane measurements made in the field where the true emission rate is not known. Let Q^{new} represent the unknown true emission rate associated with the new measurement and M^{new} be the measurement made in the field. We wish to know the distribution of Q^{new} given M^{new} and our uncertainty model derived from controlled release data. The distribution of interest is $p(Q^{new} \mid M^{new}, \underline{M})$. Using Bayes equation, Eq. (3), we can say that

$$\begin{aligned} p(Q^{new} \mid M^{new}, \underline{M}) &\propto p(M^{new} \mid Q^{new}, \underline{M}) p(Q^{new} \mid \underline{M}) \\ &= p(M^{new} \mid Q^{new}, \underline{M}) p(Q^{new}), \end{aligned} \tag{5}$$

where the true emission rate is modeled as independent of the measurements from the controlled release trials. Computational techniques can be used to

obtain $p(M^{new} \mid Q^{new}, \underline{M})$. We can say that

$$p(M^{new} \mid Q^{new}, \underline{M}) = \int p(M^{new} \mid Q^{new}, \underline{\theta}, \underline{M}) p(\underline{\theta} \mid Q^{new}, \underline{M}) d\underline{\theta}$$
$$= \int p(M^{new} \mid Q^{new}, \underline{\theta}) p(\underline{\theta} \mid \underline{M}) d\underline{\theta}. \quad (6)$$

The distribution $p(M^{new} \mid Q^{new}, \underline{\theta})$ is the likelihood given by the uncertainty model from 3.1.1 where $Q_i = Q^{new}$. $p(\underline{\theta} \mid \underline{M})$ is the posterior distribution of $\underline{\theta}$ given the controlled release data. This integral can be approximated as follows:

1. Sample $\underline{\theta}^j$ $j = 1, \ldots, J$ times from $p(\underline{\theta} \mid \underline{M})$
2. For $j = 1, \ldots, J$, calculate $p(M^{new} \mid Q^{new}, \underline{\theta}^j)$
3. The integral in (6) $\approx \frac{\sum_{j=1}^{J} p(M^{new} \mid Q^{new}, \underline{\theta}^j)}{J}$

This process can also be performed for repeated measurements of the same source, that is, where $\underline{M}^{new}$ is a vector, $\underline{M}^{new} = (M_1^{new}, \ldots, M_n^{new})$. Under the assumption that the uncertainty in each measurement is independent of previous measurements, conditional on Q^{new}, at step two, $p(\underline{M}^{new} \mid Q^{new}, \underline{\theta}^j) = \prod_{i=1}^{n} p(M_i^{new} \mid Q^{new}, \underline{\theta}^j)$.

The distribution $p(Q^{new})$ in Eq. (5) is the prior distribution of Q^{new}, where we assume the new measurement is independent of the controlled release data. This distribution can be informed using relevant pre-existing data, such as survey data on leak rates in the region where the measurement was made. This facilitates the natural synthesis of external data with controlled release data and the new observed measurement.

The posterior distribution of interest $p(Q^{new} \mid M^{new}, \underline{M})$ can be sampled as follows:

1. Draw a sample of size L from the prior distribution $p(Q^{new})$: $S = \{Q_1^{new}, \ldots, Q_L^{new}\}$
2. For $l = 1, \ldots, L$, approximate $p(M^{new} \mid Q_l^{new}, \underline{M})$ using the previously described algorithm
3. Calculate weights

$$w_l = \frac{p(M^{new} \mid Q_l^{new}, \underline{M})}{\sum_{k=1}^{L} p(M^{new} \mid Q_k^{new}, \underline{M})} \quad (7)$$

4. Resample from S with sampling probabilities w_l for $l = 1, \ldots, L$ to get a sample of size K from the posterior distribution $p(Q^{new} \mid M^{new}, \underline{M})$.

4.2.2. Set-up

The algorithm described in the previous section is demonstrated using the model selected in Sec. 4.1 for QOGI Operator C. We consider two data scenarios: one where a single measurement of a source is made, and one where five independent measurements of the same source are made. We also investigate two different prior distributions, $p(Q^{new})$, to show how information flows from the prior to the posterior distribution.

We simulate the process of performing measurements in the field (e.g., as part of a LDAR survey program) as follows: First, we choose a hypothetical true value for the source we will measure, which we set to $Q_i = 25$ kg/h. Then, we simulate "measurements" by drawing from the posterior predictive distribution defined in Sec. 3.1.2, $p(\tilde{M}_i \mid M_i, Q_i = 25 \text{ kg/hr})$. This is done five times to produce five independent "measurements" of the same source. Let $M_1^{new}, \ldots, M_5^{new}$ represent the five simulated measurement values: 20.3 kg/h, 11.7 kg/h, 8.9 kg/h, 16.2 kg/h, and 16.4 kg/h.

To understand how prior information impacts the estimates, we use two different prior distributions. The first prior is a uniform distribution where we only impose the upper limit of 200 kg/h, while the second is a log-normal distribution, as this distribution has been suggested to model leak rate distributions [42].

To summarise, we investigate four different scenarios:

1. Only M_1^{new} is used, and $p(Q^{new})$ is a uniform distribution from 0 to 200 kg/h.
2. All of $M_1^{new}, \ldots, M_5^{new}$ are used, and $p(Q^{new})$ is a uniform distribution from 0 to 200 kg/h.
3. Only M_1^{new} is used, and $p(Q^{new})$ is a log-normal distribution with shape parameter 1, location parameter equal to zero, and scale parameter equal to 2.6 [39]
4. All of $M_1^{new}, \ldots, M_5^{new}$ are used, and $p(Q^{new})$ is a log-normal distribution with shape parameter 1, location parameter equal to zero, and scale parameter equal to 2.6 [39]

4.2.3. Results

The different information expressed in the two priors can be visualized by comparing the histograms in Fig. 10. The uniform prior expresses that all values between 0 and 200 kg/h are equally likely. This may be a naive choice, because surveys have shown that extremely high emitters are much less likely than lower emitters [43, 42]. The log-normal prior expresses that there is about a 50% chance that the emission rate is less than 13.6 kg/h and 75% chance that the emission rate is less than 26.2 kg/h, with very large values being rare.

Also, posterior distributions are not centred around the measured values, which were all less than the "true" value of 25 kg/h. This is a reflection of the results shown in Fig. 5, where it is clear that the technology systematically underestimates the true emission rate. The algorithm presented in this section allows the information captured in the model derived from Sec. 3.1 to be inverted and produce estimates that are equal to the true emission rate, on average.

4.3. Discussion

As highlighted by Figs.3 - 9, relying solely on measurements without considering uncertainty can lead to significant misinterpretations of the emission rate. The methods presented in Sec 3.1 provide a way to summarise both the variability and systematic bias of a technology. They are situated in the Bayesian

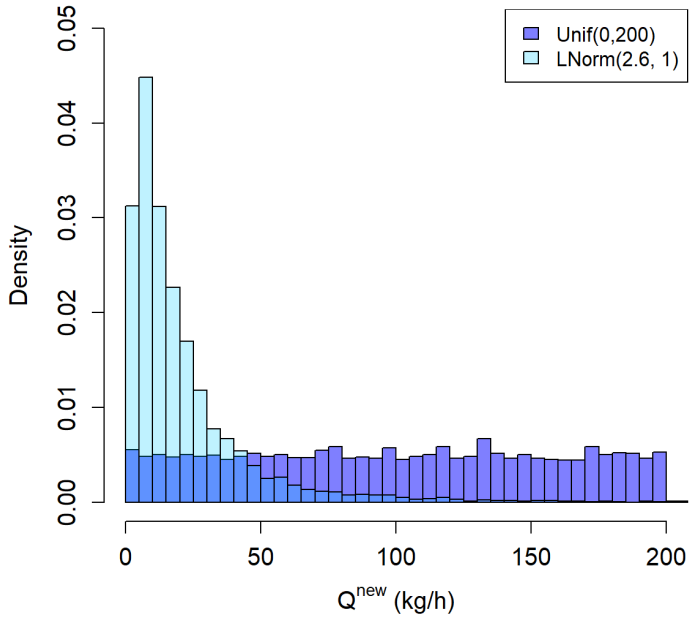

Figure 10: Comparison of histograms for the two different prior distributions $p(M^{new})$ investigated in the analysis.

statistical framework which facilitates probabilistic inference, the derivation of credible intervals, and downstream approaches, as exemplified in Sec. 4.2.

Specifically, the algorithm presented in Sec. 4.2 provides a distribution of the true emission rate given all available information, including a measurement or set of measurements, controlled release data, and external prior knowledge, e.g., what is a believable leak rate for a given scenario? This prior knowledge strongly informs the posterior when the measurement data is limited, but its influence diminishes as more measurements becomes available. This is beneficial because it formalises an informal process: in the absence of data, we must rely more on previous knowledge, whereas when more data are available, we rely less on our previous knowledge. The results also show that increasing the numbers of measurements reduces uncertainty in the posterior; that is, as we collect more data, we can be more certain about the true value of the emission rate. The models presented here could be used to determine how many measurements should be performed for a certain technology and emission rate to ensure that the credibility intervals are within a certain value. This algorithm in tandem with the models described in Sec. 3.1 could be used in the future to help plan or assess the effectiveness of LDAR programs, e.g., by identifying the optimal

Table 5: Summary of 90% credible intervals (CrI) for the four different data scenarios, and their lengths.

Prior	Number of Values	90% CrI	Length of CrI
Unif(0,200)	1	(12.3, 71.0)	58.7
Unif(0,200)	5	(14.9, 37.3)	22.3
LogNormal(2.6, 1)	1	(3.1, 47.9)	44.8
LogNormal(2.6, 1)	5	(13.9, 35.5)	21.6

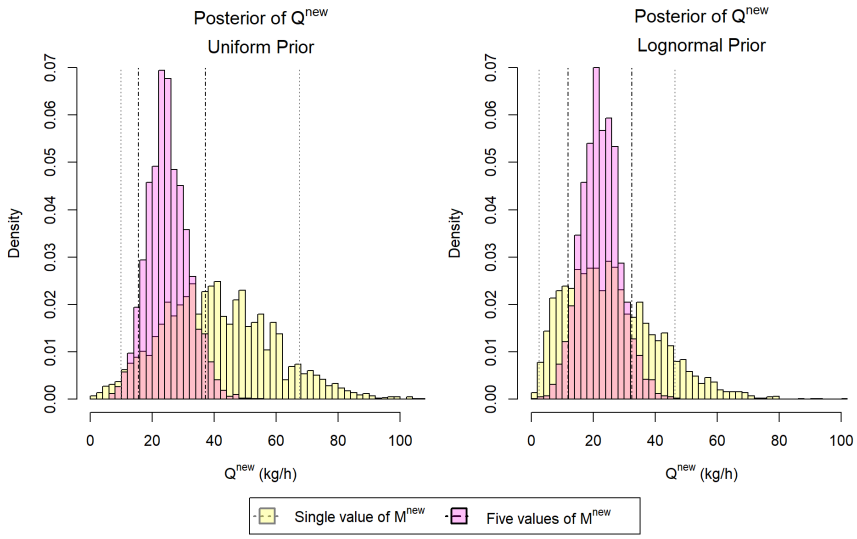

Figure 11: Posterior distributions for the true value of the emission rate given the observed measurement(s), prior distribution, and controlled release data.

combination of technologies that achieve a certain credibility interval.

Although extensive meteorological data were collected during the field trials (as detailed in the SI), we refrain from incorporating them in the statistical models. This is because the goal of the models is to summarise the performance of a technology over a variety of conditions. To this end, the field trials were conducted at different times-of-year and each trial over multiple days, so that the results could be used to assess the performance of the technologies over a variety of conditions. This also enabled the use of external validation - checking to see if data collected under one set of conditions can be used to predict data collected under a different set of conditions. However, it may be possible to improve the predictive performance of the models by incorporating meteorological data into the likelihood, discussed further in the Conclusion section.

5. Conclusions

A wide range of quantification technologies have been deployed to measure methane emissions from the upstream oil and gas sector, including ground-based infrared cameras, airborne hyperspectral imaging, and truck and airborne laser absorption spectroscopy. However, measurements from these technologies can only be interpreted properly in the context of uncertainty, which arise from measurement noise, uncertain model parameters, and, especially, the approximations and simplifications that must be made to render the models tractable (model error).

This paper presented a formalism for developing uncertainty estimates from controlled release data within the Bayesian framework. The outcome of this analysis are posterior probability density functions that comprehensively define what is known about an emission rate, based on the measurement data, controlled release data, and prior information. This approach is entirely technology-agnostic, does not require knowledge of the underlying physical model, and may be adapted to a wide range of scenarios.

The probability density functions may be summarized as credibility intervals (e.g., the range of emission rates that correspond to a given probability) and may be used for other purposes, such as the input to probabilistic simulations to assess the effectiveness of alternative fugitive emissions management plans (alt-FEMPS) or the in calculating the uncertainty attached to inventory estimates. This study demonstrates clearly the importance of multiple measurements during any particular emission survey study. Two implications of this result are immediately of significance for regulation and policy. Currently the structure and schedule of regulations tends to specify only the annual frequencies of site and equipment emission monitoring surveys (cite the Alberta or BC regulations if you wish). A key result of this study is that the uncertainty of a given survey is dependent on the number of measurements made. Due to the relatively large uncertainties observed for different technologies in this study, it is likely that multiple measurements would be required to achieve a desired emission rate uncertainty. Therefore, regulators would be advised to specify the minimum number of observations at any emission source in addition to the annual emissions survey frequencies. Alternatively, and perhaps more appropriately due to the relationship between uncertainty and true emission rate, a desired uncertainty range per emission source should be specified and the number of measurements required to achieve this uncertainty should be made.

An important aspect of methane detection and quantification technologies not covered by the methodology so far is the detection probability, that is, how likely it is that a technology will detect a given emission source under a set of conditions. Work has been done to characterise probability of detection and uncertainty separately, e.g., [22]. An advantage of the Bayesian approach taken in this paper is that it lends itself well to model extension. Modelling of detection probability could be done simultaneously to uncertainty modelling by using a hierarchical Bayesian model. Extending the model to simultaneously consider detection probability and measurement uncertainty is an important avenue of

future work.

Another area of further exploration lies in the incorporation of meteorological data into the likelihood. Given a large amount of controlled release data, it may be possible to narrow the posterior probability by incorporating the meteorological data into the likelihood. For example, the likelihood could be modified so that the median function is

$$\phi_i = \alpha_0 + \alpha_1 Q_i + \beta_1(\text{wind speed}_i).$$

In other words, the meteorological data becomes an additional observable. Another caveat to the inclusion of covariates is that more controlled release data would be needed to accurately estimate the increased number of model parameters. However, it has the potential to improve the predictive performance of the models. Incorporating meteorological data into the model could be investigated further.

6. Acknowledgements

This research was sponsored by the Alberta Upstream Petroleum Research Fund (AUPRF), the Clean Resources Information Network (CRIN), and the Natural Sciences and Engineering Research Council (ALLRP 571135-2021). The authors are grateful for the support of Carbon Management Canada, Petroleum Technology Alliance Canada, and Arolytics, Inc.

References

[1] K. Calvin, D. Dasgupta, G. Krinner, A. Mukherji, P. W. Thorne, C. Trisos, J. Romero, P. Aldunce, K. Barrett, G. Blanco, W. W. Cheung, S. Connors, F. Denton, A. Diongue-Niang, D. Dodman, M. Garschagen, O. Geden, B. Hayward, C. Jones, F. Jotzo, T. Krug, R. Lasco, Y.-Y. Lee, V. Masson-Delmotte, M. Meinshausen, K. Mintenbeck, A. Mokssit, F. E. Otto, M. Pathak, A. Pirani, E. Poloczanska, H.-O. Pörtner, A. Revi, D. C. Roberts, J. Roy, A. C. Ruane, J. Skea, P. R. Shukla, R. Slade, A. Slangen, Y. Sokona, A. A. Sörensson, M. Tignor, D. Van Vuuren, Y.-M. Wei, H. Winkler, P. Zhai, Z. Zommers, J.-C. Hourcade, F. X. Johnson, S. Pachauri, N. P. Simpson, C. Singh, A. Thomas, E. Totin, P. Arias, M. Bustamante, I. Elgizouli, G. Flato, M. Howden, C. Méndez-Vallejo, J. J. Pereira, R. Pichs-Madruga, S. K. Rose, Y. Saheb, R. Sánchez Rodríguez, D. Ürge Vorsatz, C. Xiao, N. Yassaa, A. Alegría, K. Armour, B. Bednar-Friedl, K. Blok, G. Cissé, F. Dentener, S. Eriksen, E. Fischer, G. Garner, C. Guivarch, M. Haasnoot, G. Hansen, M. Hauser, E. Hawkins, T. Hermans, R. Kopp, N. Leprince-Ringuet, J. Lewis, D. Ley, C. Ludden, L. Niamir, Z. Nicholls, S. Some, S. Szopa, B. Trewin, K.-I. Van Der Wijst, G. Winter, M. Witting, A. Birt, M. Ha, J. Romero, J. Kim, E. F. Haites, Y. Jung, R. Stavins, A. Birt, M. Ha, D. J. A. Orendain, L. Ignon, S. Park, Y. Park, A. Reisinger, D. Cammaramo, A. Fischlin, J. S. Fuglestvedt, G. Hansen,

C. Ludden, V. Masson-Delmotte, J. R. Matthews, K. Mintenbeck, A. Pirani, E. Poloczanska, N. Leprince-Ringuet, C. Péan, IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland., Tech. rep., Intergovernmental Panel on Climate Change (IPCC), edition: First (Jul. 2023). doi:10.59327/IPCC/AR6-9789291691647.
URL https://www.ipcc.ch/report/ar6/syr/

[2] A. Wilson, New Optical Gas-Imaging Technology for Quantifying Fugitive-Emission Rates, Journal of Petroleum Technology 68 (08) (2016) 78–79. doi:10.2118/0816-0078-JPT.
URL https://doi.org/10.2118/0816-0078-JPT

[3] A. L. Mitchell, D. S. Tkacik, J. R. Roscioli, S. C. Herndon, T. I. Yacovitch, D. M. Martinez, T. L. Vaughn, L. L. Williams, M. R. Sullivan, C. Floerchinger, M. Omara, R. Subramanian, D. Zimmerle, A. J. Marchese, A. L. Robinson, Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results, Environmental Science & Technology 49 (5) (2015) 3219–3227, publisher: American Chemical Society. doi:10.1021/es5052809.
URL https://doi.org/10.1021/es5052809

[4] R. Kang, P. Liatsis, D. C. Kyritsis, Emission Quantification via Passive Infrared Optical Gas Imaging: A Review, Energies 15 (9) (2022) 3304, number: 9 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en15093304.
URL https://www.mdpi.com/1996-1073/15/9/3304

[5] R. S. Heltzel, D. R. Johnson, M. T. Zaki, A. K. Gebreslase, O. I. Abdul-Aziz, Machine learning techniques to increase the performance of indirect methane quantification from a single, stationary sensor, Heliyon 8 (12) (2022) e11962. doi:10.1016/j.heliyon.2022.e11962.
URL https://www.sciencedirect.com/science/article/pii/S2405844022032509

[6] H. L. Brantley, E. D. Thoma, W. C. Squier, B. B. Guven, D. Lyon, Assessment of Methane Emissions from Oil and Gas Production Pads using Mobile Measurements, Environmental Science & Technology 48 (24) (2014) 14508–14515, publisher: American Chemical Society. doi:10.1021/es503070q.
URL https://doi.org/10.1021/es503070q

[7] E. Atherton, D. Risk, C. Fougère, M. Lavoie, A. Marshall, J. Werring, J. P. Williams, C. Minions, Mobile measurement of methane emissions from natural gas developments in northeastern British Columbia, Canada, Atmospheric Chemistry and Physics 17 (20) (2017) 12405–12420, publisher:

Copernicus GmbH. doi:10.5194/acp-17-12405-2017.
URL https://acp.copernicus.org/articles/17/12405/2017/

[8] D. R. Caulton, Q. Li, E. Bou-Zeid, J. P. Fitts, L. M. Golston, D. Pan, J. Lu, H. M. Lane, B. Buchholz, X. Guo, J. McSpiritt, L. Wendt, M. A. Zondlo, Quantifying uncertainties from mobile-laboratory-derived emissions of well pads using inverse Gaussian methods, Atmospheric Chemistry and Physics 18 (20) (2018) 15145–15168, publisher: Copernicus GmbH. doi:10.5194/acp-18-15145-2018.
URL https://acp.copernicus.org/articles/18/15145/2018/

[9] B. M. Conrad, D. R. Tyner, H. Z. Li, D. Xie, M. R. Johnson, A measurement-based upstream oil and gas methane inventory for Alberta, Canada reveals higher emissions and different sources than official estimates, Communications Earth & Environment 4 (1) (2023) 1–10, number: 1 Publisher: Nature Publishing Group. doi:10.1038/s43247-023-01081-0.
URL https://www.nature.com/articles/s43247-023-01081-0

[10] E. D. Sherwin, Y. Chen, A. P. Ravikumar, A. R. Brandt, Single-blind test of airplane-based hyperspectral methane detection via controlled releases, Elementa: Science of the Anthropocene 9 (1) (2021) 00063. doi:10.1525/elementa.2021.00063.
URL https://doi.org/10.1525/elementa.2021.00063

[11] D. J. Varon, J. McKeever, D. Jervis, J. D. Maasakkers, S. Pandey, S. Houweling, I. Aben, T. Scarpelli, D. J. Jacob, Satellite Discovery of Anomalously Large Methane Point Sources From Oil/Gas Production, Geophysical Research Letters 46 (22) (2019) 13507–13516, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL083798. doi:10.1029/2019GL083798.
URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2019GL083798

[12] J. P. Canfin, Pascal, Report on the proposal for a regulation of the European Parliament and of the Council on methane emissions reduction in the energy sector and amending Regulation (EU) 2019/942 | A9-0162/2023 | European Parliament (Apr. 2023).
URL https://www.europarl.europa.eu/doceo/document/A-9-2023-0162_EN.html

[13] OGMP 2.0 – The Oil & Gas Methane Partnership 2.0.
URL https://ogmpartnership.com/

[14] T. A. Fox, C. H. Hugenholtz, T. E. Barchyn, T. R. Gough, M. Gao, M. Staples, Can new mobile technologies enable fugitive methane reductions from the oil and gas industry?, Environmental Research Letters 16 (6) (2021) 064077, publisher: IOP Publishing. doi:10.1088/1748-9326/ac0565.
URL https://dx.doi.org/10.1088/1748-9326/ac0565

[15] E. Solazzo, M. Crippa, D. Guizzardi, M. Muntean, M. Choulga, G. Janssens-Maenhout, Uncertainties in the Emissions Database for Global

Atmospheric Research (EDGAR) emission inventory of greenhouse gases, Atmospheric Chemistry and Physics 21 (7) (2021) 5655–5683, publisher: Copernicus GmbH. doi:10.5194/acp-21-5655-2021.
URL https://acp.copernicus.org/articles/21/5655/2021/

[16] K. MacKay, M. Lavoie, E. Bourlon, E. Atherton, E. O'Connell, J. Baillie, C. Fougère, D. Risk, Methane emissions from upstream oil and gas production in Canada are underestimated, Scientific Reports 11 (1) (2021) 8041, number: 1 Publisher: Nature Publishing Group. doi:10.1038/s41598-021-87610-3.
URL https://www.nature.com/articles/s41598-021-87610-3

[17] A. Montazeri, X. Zhou, J. D. Albertson, On the Viability of Video Imaging in Leak Rate Quantification: A Theoretical Error Analysis, Sensors 21 (17) (2021) 5683, number: 17 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/s21175683.
URL https://www.mdpi.com/1424-8220/21/17/5683

[18] M. O. L. Cambaliza, P. B. Shepson, D. R. Caulton, B. Stirm, D. Samarov, K. R. Gurney, J. Turnbull, K. J. Davis, A. Possolo, A. Karion, C. Sweeney, B. Moser, A. Hendricks, T. Lauvaux, K. Mays, J. Whetstone, J. Huang, I. Razlivanov, N. L. Miles, S. J. Richardson, Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions, Atmospheric Chemistry and Physics 14 (17) (2014) 9029–9050, publisher: Copernicus GmbH. doi:10.5194/acp-14-9029-2014.
URL https://acp.copernicus.org/articles/14/9029/2014/

[19] D. Singh, B. Barlow, C. Hugenholtz, W. Funk, C. Robinson, A. P. Ravikumar, Field Performance of New Methane Detection Technologies: Results from the Alberta Methane Field Challenge, publisher: EarthArXiv (Jun. 2021).
URL https://eartharxiv.org/repository/view/1860/

[20] M. R. Johnson, D. R. Tyner, A. J. Szekeres, Blinded evaluation of airborne methane source detection using Bridger Photonics LiDAR, Remote Sensing of Environment 259 (2021) 112418. doi:10.1016/j.rse.2021.112418.
URL https://www.sciencedirect.com/science/article/pii/S003442572100136X

[21] C. Bell, J. Rutherford, A. Brandt, E. Sherwin, T. Vaughn, D. Zimmerle, Single-blind determination of methane detection limits and quantification accuracy using aircraft-based LiDAR, Elementa: Science of the Anthropocene 10 (1) (2022) 00080. doi:10.1525/elementa.2022.00080.
URL https://doi.org/10.1525/elementa.2022.00080

[22] B. M. Conrad, D. R. Tyner, M. R. Johnson, Robust probabilities of detection and quantification uncertainty for aerial methane detection: Examples for three airborne technologies, Remote Sensing of Environment 288 (2023) 113499. doi:10.1016/j.rse.2023.113499.
URL https://www.sciencedirect.com/science/article/pii/S0034425723000500

[23] C. E. Kemp, A. P. Ravikumar, A. R. Brandt, Comparing Natural Gas Leakage Detection Technologies Using an Open-Source "Virtual Gas Field" Simulator, Environmental Science & Technology 50 (8) (2016) 4546–4553, publisher: American Chemical Society. doi:10.1021/acs.est.5b06068.
URL https://doi.org/10.1021/acs.est.5b06068

[24] T. A. Fox, M. Gao, T. E. Barchyn, Y. L. Jamin, C. H. Hugenholtz, An agent-based model for estimating emissions reduction equivalence among leak detection and repair programs, Journal of Cleaner Production 282 (2021) 125237. doi:10.1016/j.jclepro.2020.125237.
URL https://www.sciencedirect.com/science/article/pii/S0959652620352811

[25] M. Lavoie, D. Risk, E. O'Connell, E. Atherton, J. Gorski, J. Johnson, Evaluating the benefits of alternative leak detection programs, publisher: EarthArXiv (May 2021).
URL https://eartharxiv.org/repository/view/2355/

[26] M. R. Johnson, B. M. Conrad, D. R. Tyner, Creating measurement-based oil and gas sector methane inventories using source-resolved aerial surveys, Communications Earth & Environment 4 (1) (2023) 1–9, number: 1 Publisher: Nature Publishing Group. doi:10.1038/s43247-023-00769-7.
URL https://www.nature.com/articles/s43247-023-00769-7

[27] Y. Zeng, J. Morris, Detection limits of optical gas imagers as a function of temperature differential and distance, Journal of the Air & Waste Management Association 69 (3) (2019) 351–361, publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10962247.2018.1540366. doi:10.1080/10962247.2018.1540366.
URL https://doi.org/10.1080/10962247.2018.1540366

[28] A. P. Ravikumar, J. Wang, M. McGuire, C. S. Bell, D. Zimmerle, A. R. Brandt, "Good versus Good Enough?" Empirical Tests of Methane Leak Detection Sensitivity of a Commercial Infrared Camera, Environmental Science & Technology 52 (4) (2018) 2368–2374, publisher: American Chemical Society. doi:10.1021/acs.est.7b04945.
URL https://doi.org/10.1021/acs.est.7b04945

[29] D. Zimmerle, T. Vaughn, C. Bell, K. Bennett, P. Deshmukh, E. Thoma, Detection Limits of Optical Gas Imaging for Natural Gas Leak Detection in Realistic Controlled Conditions, Environmental Science & Technology 54 (18) (2020) 11506–11514, publisher: American Chemical Society. doi:10.1021/acs.est.0c01285.
URL https://doi.org/10.1021/acs.est.0c01285

[30] J. M. Supplee, E. A. Whittaker, W. Lenth, Theoretical description of frequency modulation and wavelength modulation spectroscopy, Applied Optics 33 (27) (1994) 6294–6302, publisher: Optica Publishing Group. doi:10.1364/AO.33.006294.
URL https://opg.optica.org/ao/abstract.cfm?uri=ao-33-27-6294

[31] T. K. Flesch, J. D. Wilson, E. Yee, Backward-Time Lagrangian Stochastic Dispersion Models and Their Application to Estimate Gaseous Emissions, Journal of Applied Meteorology and Climatology 34 (6) (1995) 1320–1332, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology. doi:10.1175/1520-0450(1995)034¡1320:BTLSDM¿2.0.CO;2.
URL https://journals.ametsoc.org/view/journals/apme/34/6/1520-0450_1995_034_1320_btlsdm_2

[32] T. K. Flesch, J. D. Wilson, L. A. Harper, B. P. Crenna, R. R. Sharpe, Deducing Ground-to-Air Emissions from Observed Trace Gas Concentrations: A Field Trial, Journal of Applied Meteorology and Climatology 43 (3) (2004) 487–502, publisher: American Meteorological Society Section: Journal of Applied Meteorology and Climatology. doi:10.1175/1520-0450(2004)043¡0487:DGEFOT¿2.0.CO;2.
URL https://journals.ametsoc.org/view/journals/apme/43/3/1520-0450_2004_043_0487_dgefot_2

[33] A. E. Esparza, J.-F. Gauthier, Monitoring Leaks at Oil and Gas Facilities Using the Same Sensor on Aircraft and Satellite Platforms, OnePetro, 2021. doi:10.15530/urtec-2021-5666.
URL https://dx.doi.org/10.15530/urtec-2021-5666

[34] A. E. Esparza, G. Rowan, A. Newhook, H. J. Deglint, B. Garrison, B. Orth-Lashley, M. Girard, W. Shaw, Analysis of a tiered top-down approach using satellite and aircraft platforms to monitor oil and gas facilities in the Permian basin, Renewable and Sustainable Energy Reviews 178 (2023) 113265. doi:10.1016/j.rser.2023.113265.
URL https://www.sciencedirect.com/science/article/pii/S1364032123001211

[35] D. J. Varon, D. J. Jacob, J. McKeever, D. Jervis, B. O. A. Durak, Y. Xia, Y. Huang, Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes, Atmospheric Measurement Techniques 11 (10) (2018) 5673–5686, publisher: Copernicus GmbH. doi:10.5194/amt-11-5673-2018.
URL https://amt.copernicus.org/articles/11/5673/2018/

[36] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin, Bayesian Data Analysis, Third Edition, CRC Press, 2013, google-Books-ID: ZXL6AQAAQBAJ.

[37] M. Plummer, JAGS version 4.3.0 user manual (Jun. 2023).
URL http://www.stat.yale.edu/ jtc5/238/materials/jags_4.3.0_manual_with_distributions.pd

[38] M. J. Denwood, runjags: An R Package Providing Interface Utilities, Model Templates, Parallel Computing Methods and Additional Distributions for MCMC Models in JAGS, Journal of Statistical Software 71 (2016) 1–25. doi:10.18637/jss.v071.i09.
URL https://doi.org/10.18637/jss.v071.i09

[39] NIST, 1.3.6.6.9. Lognormal Distribution (Jun. 2023).
URL https://www.itl.nist.gov/div898/handbook/eda/section3/eda3669.htm

[40] A. Gelman, Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper), Bayesian Analysis 1 (3) (2006) 515–534, publisher: International Society for Bayesian Analysis. doi:10.1214/06-BA117A.
URL https://projecteuclid.org/journals/bayesian-analysis/volume-1/issue-3/Prior-distrib

[41] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, A. Van Der Linde, Bayesian measures of model complexity and fit, J R Stat Soc Series B Stat Methodology 64 (4) (2002) 583–639.

[42] D. Zavala-Araiza, S. C. Herndon, J. R. Roscioli, T. I. Yacovitch, M. R. Johnson, D. R. Tyner, M. Omara, B. Knighton, Methane emissions from oil and gas production sites in Alberta, Canada, Elementa: Science of the Anthropocene 6 (2018) 27. doi:10.1525/elementa.284.
URL https://doi.org/10.1525/elementa.284

[43] A. R. Brandt, G. A. Heath, D. Cooley, Methane Leaks from Natural Gas Systems Follow Extreme Distributions, Environmental Science & Technology 50 (22) (2016) 12512–12520, publisher: American Chemical Society. doi:10.1021/acs.est.6b04303.
URL https://doi.org/10.1021/acs.est.6b04303