ACS ES&T Air

Open Access
This article is licensed under CC-BY-NC-ND 4.0
pubs.acs.org/estair

Article

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

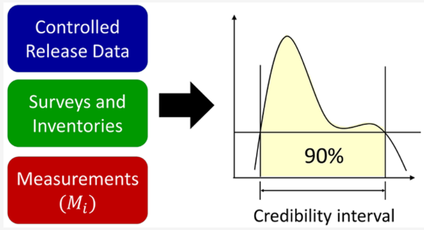
Augustine Wigle,* Audrey Béliveau, Daniel Blackmore, Paule Lapeyre, Kirk Osadetz, Christiane Lemieux, and Kyle J. Daun

Cite This: *ACS EST Air* 2024, 1, 1000–1014

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: An accurate understanding of uncertainty is needed to properly interpret methane emission estimates from upstream oil and gas sources in a variety of contexts, from component-level measurements to yearly jurisdiction-wide inventories. To characterize measurement uncertainty, we examine controlled release (CR) data from five different technology providers including quantitative gas imaging (QOGI), tunable diode laser-absorption spectroscopy (TDLAS); and airborne near-infrared hyperspectral (NIR HS) imaging. We introduce a novel empirical method to develop probability distributions of measurements given a true emission rate using the CR data. The approach includes flexible likelihoods which capture complex relationships in the data. An algorithm which provides the distribution of the true emission rate given a measurement is also developed, which synthesizes the measurement with the CR data and external information about the possible true emission rate. The results show that flexible models that accommodate complex nonlinear behavior are needed to adequately model measurement error. We also show that measurement error can vary under different conditions. We demonstrate that measurement uncertainty can be reduced by performing repeated measurements. A limitation of the study is that the collected CR data is collected under controlled conditions that may differ from those in industrial settings. As new CR data become available, the models presented in this paper can be refit to consider more diverse scenarios. The methodology can be extended to explicitly model different conditions to improve performance.



KEYWORDS: *Methane, Oil and gas, Measurement uncertainty, Measurement-based inventories, LDAR, FEMP, Bayesian*

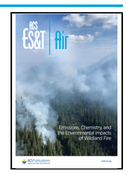
INTRODUCTION

Methane is responsible for approximately 30% of the rise in global temperatures since the industrial revolution and is the second most important climate forcing mechanism, only after carbon dioxide.[1,2] Approximately 60% of methane emissions come from anthropogenic sources,[1] and deep and rapid reductions in these sources are crucial in order to avoid the worst outcomes of climate change.[2] Emissions from the oil and gas sector (approximately 40 % of anthropogenic sources globally[1]) are more amenable for reductions compared to some other sources, through opportunities created by emerging technology and improved operational practices. Doing this requires instrumentation that can detect and quantify methane emissions to assess regulatory compliance, develop effective and science-based policy, design and deploy cost-effective emission mitigation strategies, and for climate modeling.

Methane emission measurements 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,[2−4] 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 high quantification uncertainty can impair the cost-effectiveness of LDAR programs.[5] Furthermore, methane emissions measurements are used to develop broader jurisdiction-wide and global inventories,[6−8] 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. For example, the well-known "gap" between inventories compiled from "top-down" (aircraft or satellite-based modalities) and "bottom-up" techniques (hand-held or in-situ technologies) is partially explained by the uncertainties attached to individual measurements.[7,9] Given the high global warming potential of methane and the large volumes involved, variations in inventory estimates have a far reaching impact on climate modeling.

A wide range of technologies have been deployed to quantify emissions from the oil and gas industry, including quantitative optical gas imaging (QOGI) using mid-wavelength infrared

Received: February 10, 2024
Revised: June 24, 2024
Accepted: June 25, 2024
Published: July 22, 2024



© 2024 The Authors. Published by American Chemical Society

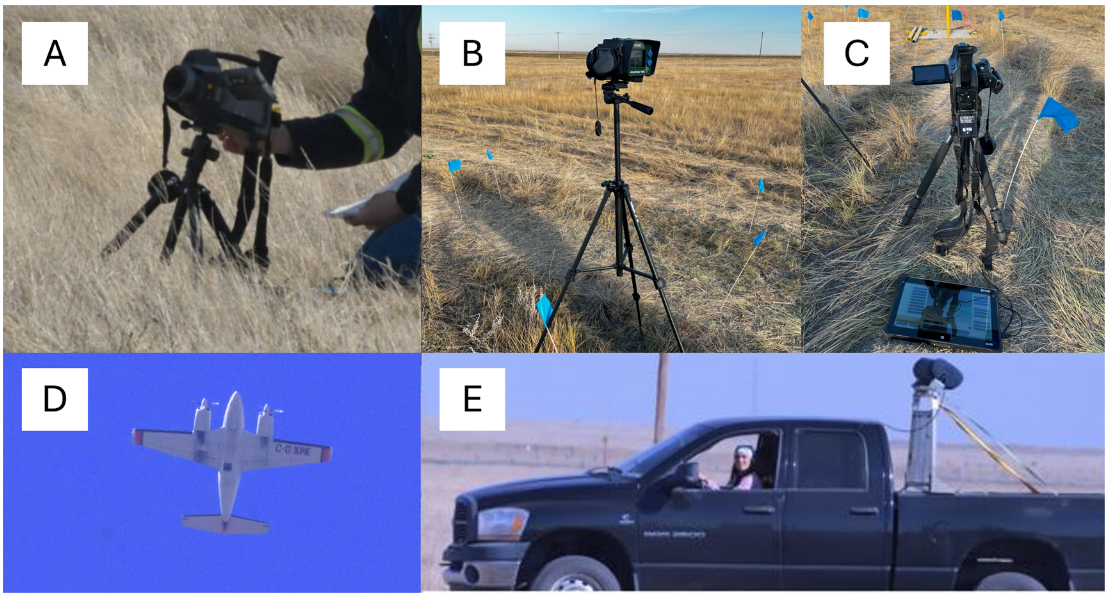


Figure 1. Technologies studied in the measurement campaigns. A: QOGI camera used by operator A, B: QOGI camera used by operator B, C: QOGI camera used by operator C, D: airborne NIR HSI, and E: truck-based TDLAS driven by Paule Lapeyre. Photos B and C from Michael Nagorski. Other photos from Kyle J. Daun.

(MWIR) cameras;[10−12] stationary[13] and mobile[14−16] methane concentration sensors; and airborne[8,17] and satellite-based[18] 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 × 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/h).

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.[19] 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). Caulton et al.[16] 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.[20] 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 can provide insights into the uncertainty of methane emission estimates obtained from specific 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 completely unknown due to proprietary aspects of the technology. Second, the results of a physics-based uncertainty analysis are specific to a given technology and would therefore need to be derived separately for every system 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.[16] 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.

Data-driven 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. Data-driven approaches have two key benefits over physics-driven approaches: (1) they are fitted to field data, 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.

Data-driven approaches require controlled-release data. To this end, many single- or double-blinded measurement campaigns have been conducted with the goal of assessing the performance of methane emissions quantification technology.[17,21−25] However, data-driven approaches employed on these measurements have mainly been limited to linear regression approaches,[17,21−24] with the exception of Conrad et al.,[25] who provide an approach to derive the distribution of the true emission rate given a measurement from an airborne methane detection and quantification technology. A limitation of these approaches is that the models have a constant relationship between uncertainty and the true emission rate (either strictly additive or multiplicative). Also, existing approaches often do not account for the possibility that a technology may detect methane when there is none being released. Finally, there is a lack of reporting on how the methods should be applied in different scenarios, such as in performing field measurements in LDAR programs or as part of measurement-based surveys.

Uncertainty in quantifying methane emissions using detection and quantification technology can be decomposed into uncertainty due to missed detection of sources ("detection uncertainty") and uncertainty in the measurements when methane is detected ("measurement uncertainty"). In this work, we introduce a flexible framework to elucidate measurement uncertainty that can be applied to any technology modality and illustrate its use. The models are derived from single-blind controlled release (CR) data from two field campaigns carried out using four methane detection and quantification technologies, as well as CR data reported by Conrad et al.[25] The framework allows for the derivation of two important probability distributions: (1) the distribution of non-zero measurements given the true emission rate and (2) the distribution of the true emission rate given a successful 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.[26−28] The second distribution is an important input to simulation methods used to derive measurement-based inventory estimates as proposed by Johnson et al.[29] Our approach to deriving the second distribution also incorporates context-specific prior information into the analysis, such as knowledge of the emission rate distribution typical for a type of facility or component. The results of the analysis are data-driven and the design of the measurement campaigns allows for the assessment of the potential real-world effectiveness of the measurement uncertainty results.

MATERIALS AND METHODS

Methane Quantification Technologies. We demonstrate the analysis procedure using CR data from four methane quantification technologies, three of which were evaluated in the measurement campaigns: 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.[25] Examples of the technologies investigated in the campaigns are shown in Figure 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 band. Intensity entering the camera aperture is imaged through the cold-filter and onto a focal plane array (FPA) that produces a pixel intensity. The camera data is then analyzed in near real-time by software on a peripheral tablet. The measurement model is composed of a spectroscopic submodel 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.[22,30,31] Identifying favorable measurement scenarios draws considerably on operator experience.[32] Three QOGI systems were deployed by three operators of varying experience, as summarized in Table 1.

Table 1. Description of QOGI Operators and Equipment Characteristics

Operator	Experience	System	Campaign
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 and 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 measurement campaign. 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.

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)[33] and then converted to a path-average concentration (ppm). The Boreal GasFinder3-VB system consists of a TDLAS operating at 1653 cm^{-1} in wavelength modulated spectroscopy mode. The laser/detector and retroreflector define three measurement paths within a 1.3 m long perforated measurement cell mounted behind the cab of a pickup truck. 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[34,35] to obtain a release estimate for each plume transect.

Airborne NIR HS Imaging. 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,[36] mounted inside an aircraft.[37] The aircraft overflew the releases at approximately 1500 m above ground level at 240 km/h. 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[38] using wind data from an online weather model to find the emission rate.

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.[23]

Measurement Campaign Design and Execution. The QOGI, truck-mounted TDLAS, and airborne NIR HS imaging systems were evaluated through two CR measurement 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. Providers conducted measurements and analyzed data in the same way they would deploy in "real world" scenarios. In a small minority of the cases, the technology providers provided a "void" or "null" measurements due to sensor failures, or, in the case of the airborne NIR HS measurements, excessive cloud cover. The number of these instances are provided in the SI 1. In almost all of these cases the technology provider included an explanation for why the inferred rate was not included.

The conditions of the CRs have important implications for the model results. Measurement error depends on confounding "real world" factors such as complex aerodynamics from nearby structures and reflecting surfaces in the case of radiometric measurements that are not reflected by the ideal conditions of the CR measurement campaign. The statistical model is fit to data collected under a certain set of conditions (range of true emission rates, meteorological conditions, etc.). Model predictions should only be considered reliable under conditions similar to those that prevailed during the CRs.[39] To maximize the observed meteorological conditions in our CR data, we conducted two campaigns at different times of year. We consider a range of release rates that correspond to realistic scenarios from 0 to 80 kg/h. Finally, we performed releases from a variety of industrially relevant scenarios, including 1.7 m, 3.4 m, and 4.8 tall stacks, a 14-m tall unlit flare, and a storage tank.

Due to the location of the measurement campaigns and modality of several technologies, it was not possible to carry out fully double-blinded CRs where the technology providers were unaware of the existence, location, and release rate of CRs, which is considerably more challenging compared to the single-blind releases in this study. Due to this limitation, our study focuses on quantifying measurement uncertainty rather than detection error.

Information on releases and meteorological conditions are provided in SI 1. A summary of the number of controlled releases completed for each technology provider and campaign is given in Table 2. Additional data was taken from the CR studies of the Bridger GML system reported by Conrad et al.[25] to demonstrate the applicability of the methods. Anonymized release and emissions estimates data from both trials are available at https://github.com/augustinewigle/methaneUQ. 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.

Deriving Measurement Error Models Using Controlled Release Data. We propose a statistical model that

Table 2. Summary of Technologies, Providers, and Available Data from the Measurement Campaigns and External Sources[a]

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

[a]N_1 and N_2 refer to the number of observations collected for a given technology during the first and second campaigns, respectively.

answers the following question: For a given true emission rate, what range of measurements could be expected, given the observed CR data, considering model error and measurement noise? As discussed in the Measurement Campaign Design and Execution section, the models described in this section are conditional on the data used to fit the models. That is, their prediction accuracy can only be guaranteed for measurements made under similar conditions. Additionally, the model is fit only to data where methane was detected, and therefore does not account for missed detections. The implications of this are discussed in the Results and Discussion section.

Novel Flexible Model. Let Q_i be the true emission rate corresponding to the ith observation in the measurement campaign, and M_i be the corresponding emission rate estimated by the technology for the ith observation, $i = 1, ..., n$ where n is the total number of observations for the given technology. Here, we describe a likelihood that models the relationship between Q_i and M_i, given that methane was detected.

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, such as the case in Figure 2. 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 that fulfills these requirements is given by

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

where

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

and

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

This likelihood is normal on the log scale, which corresponds to a log-normal likelihood on the measurement scale; a justification for this treatment is given below. The median of M_i is modeled by a continuous piece-wise function of Q_i:

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

which is quadratic for values of Q_i below a threshold γ and linear above γ. To ensure that the function is continuous at $Q_i = \gamma$, we impose the restriction that $\beta_0 = \alpha_2\gamma^2 - \beta_1\gamma$. The specification of ϕ_i can be modified to give the best prediction results and fit to the data, depending on the detection

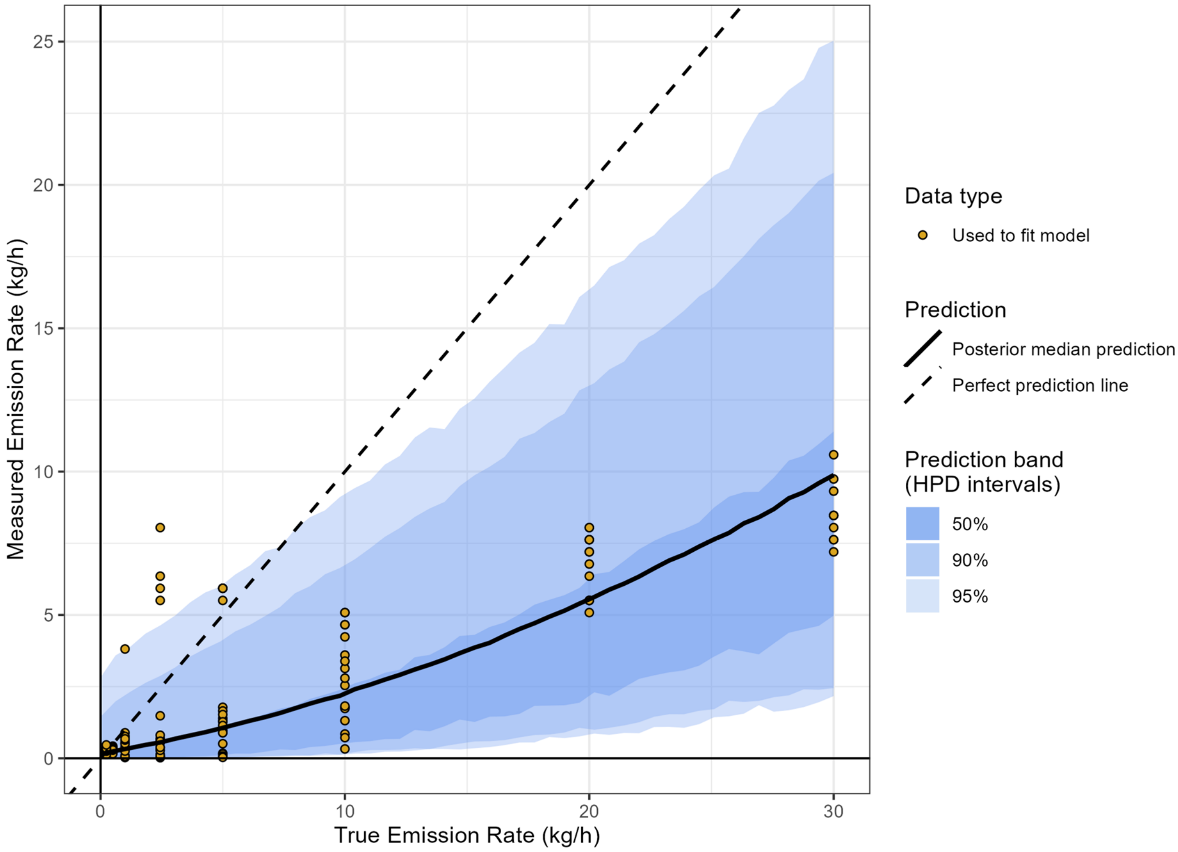


Figure 2. Uncertainty quantification model results for QOGI Operator A. The dotted line represents a perfect technology where the measured emission rate equals the true emission rate Q_i. The solid line represents the median predicted by the model for a given Q_i. The blue prediction bands represent the credible region for model predictions given Q_i; that is, a new measurement lies in the interval (X, Y) with a probability of Z. The model was fit to data from the first campaign. No external data were available.

technology. For example, the threshold parameter γ, β_0, and β_1 could be removed, which would give a quadratic relationship characterized by α_0, α_1, and α_2 over the whole range of Q_i.

The model can be rewritten to facilitate interpretation by exponentiating both sides of eq 1:

$$M_i = \phi_i \times e^{\epsilon_i} \tag{3}$$

where ϕ_i is the median measurement for a true emission rate of Q_i and e^{ϵ_i} is equal to one on average. That is, the function ϕ_i adjusts for systematic bias and e^{ϵ_i} is a multiplicative error term that accounts for remaining variation.

The likelihood in eq 1 is an extension of the scheme proposed by Conrad et al.[25] 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 $i = 1, ..., n$. 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 piecewise linear and quadratic relationships between the median measurement and Q_i and accommodate false positives rather than using a strictly linear model which does not allow false positives. Second, the inclusion of the threshold parameter γ 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 is that Conrad et al.[25] 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 that are motivated by the data. Finally, we take a fully Bayesian approach to estimation and inference as discussed in the Bayesian Analysis and Prior Distributions section, whereas Conrad et al.[25] use 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 3. 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. 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 while also ensuring that the variance is continuous along the range of Q_i. With this variance form, when $Q_i = 0$, the variance on the log scale is equal to τ^{-1}, which decreases as Q_i increases. The parameter η controls how quickly the variance decreases with Q_i. The approach to choosing an appropriate likelihood, including the variance specification, is described in the Model Selection section. Table 3 summarizes the parameters in the likelihood, their

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

units, scenarios in which they may be included or removed and interpretations of their roles in the model.

Table 3. Summary of the Units for Each Parameter in the Likelihood, Important to Keep in Mind When Specifying Prior Distributions[a]

Parameter	Units	Exclusion criteria	Interpretation
α_0	$[M_i]$	*	Median measurement when $Q_i = 0$
α_1	Unitless	None	Coefficient for Q_i in the quadratic relationship between ϕ_i and Q_i (when $Q_i < \gamma$ if applicable)
α_2	$[M_i]^{-1}$	†	Coefficient for Q_i^2 in the quadratic relationship between ϕ_i and Q_i (when $Q_i < \gamma$ if applicable)
β_0	$[M_i]$	*, †	Change in intercept of linear relationship when $Q_i > 0$
β_1	Unitless	†	Change in slope of linear relationship when $Q_i > 0$
γ	$[M_i]$	†	Allows relationship between Q_i and ϕ_i to change when $Q_i > \gamma$
τ	Unitless	None	Variance parameter
η	$[M_i]$	†	Allows variance to decrease as Q_i increases—larger value means milder reduction in variance

[a]Exclusion criteria: * = may be removed if technology does not have false positives in CR data, † = may be removed if its removal leads to a simpler model which has adequate fit and predictive performance.

As a concrete example to show how the form of the likelihood can connect to physical models, consider the problem of recovering the emission rate from a downwind concentration measurement at a precisely known location and one instant in time, as is done for the truck-mounted TLDAS system. In principle, the true concentration c^{true} is related to the true emission rate as[40]

$$c^{true} = Q \times D^{true}(U^{true}) \tag{4}$$

where D^{true} is the true advection model that depends on the wind speed, U^{true}. The true advection model is complicated and stochastic due to the turbulent atmospheric boundary layer that advects the methane from its source to the downwind detection location. Instead, the methane emission rate is estimated by substituting measured concentration values c_i and measured wind speeds U_i into the inverted form of a simplified advection model D (e.g., the Gaussian plume model), which is a simplification of D^{true}. The methane emission rate estimate is then obtained by

$$M_i = \frac{1}{D(U_i)} c_i$$

The uncertainty in the estimate is due to measurement error in c_i and U_i, and model error, that is, the model D does not perfectly reflect D^{true}. We will assume that error in $D(U_i)$ is independent of error in c_i. Expressing the inverted advection model with measured wind speed as the product of the true model $D^{true}(U^{true})$ inverted times a random variable δ_i which captures the variation from the truth, and treating the measurement error in the concentration as negligible as would be expected for high release rates, we can say that

$$M_i = Q\delta_i \tag{5}$$

That is, the relationship between the measurement and the truth depends on the probability distribution of δ_i, $p(\delta_i)$, which may be unknown. Taking logs on both sides of eq 5, we obtain

$$\log(M_i) = \log(Q) + \log(\delta_i) \tag{6}$$

which has the same form as eq 1 with $\epsilon_i = \log(\delta_i)$. This means, if it can be assumed that the δ_i's are identically distributed across measurements, then the variance of ϵ_i is constant, which is equivalent to setting

$$\sigma_i^2 = \tau^{-1}$$

in eq 2. Further, the median estimated emission rate is equal to the true emission rate times the median of $p(\delta_i)$, where a median of one would represent a technology whose median measurement is equal to the truth. This is consistent with the treatment used by Conrad et al.[25]

However, as we will show, the data from the CR experiments often do not exhibit a constant variance on the log scale and thus require more advanced variance modelling. Further, measurement error in the concentration may be non-negligible. Similarly to $p(\delta_i)$, the probability distribution of the concentration measurements $p(c_i)$ is likely unknown. Since $p(c_i)$ could be dependent on the true emission rate, the median and variance of M_i may change with Q_i. Thus, we require a flexible model that can accommodate non-constant variances and medians. As demonstrated in this example, inferring the underlying likelihood structure directly from the physical model is complex even for relatively straightforward systems and involves unknown probability distributions. Thus, the flexibility in our likelihood specification function allows us to find a model that fits the data well without requiring knowledge of the distribution of δ_i or c_i, and avoids the need to make the simplifying assumption that there is no error in the concentration measurements.

Bayesian Analysis and Prior Distributions. 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. A description of Bayesian analysis is provided in SI 2.

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 summarize the units of each parameter in Table 3.

The parameters of the likelihood to be estimated are α_0, α_1, and τ, along with optional parameters α_2, γ, β_1, and/or η. Let the vector of all statistical model parameters be represented by $\underline{\theta}$. All parameters are assumed to be independent, that is, the prior distribution can be written as

$$\begin{aligned} p(\underline{\theta}) &= p(\alpha_0, \alpha_1, \tau, \alpha_2, \gamma, \beta_1, \eta) \\ &= p(\alpha_0)p(\alpha_1)p(\tau)p(\alpha_2)p(\gamma)p(\beta_1)p(\eta) \end{aligned} \tag{7}$$

so we can specify individual prior distributions for each parameter.

The prior distributions are specified by considering the role of each parameter in the model and its units. Since α_0 is the

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

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 the prior is 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 overestimates by a factor of k is the same as the probability that it underestimates by a factor of $1/k$. This property is desirable for the prior because information about under- or overestimation 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[41]) because it fulfills this requirement. For example, if α_1 follows the standard log-normal distribution the probability that $\Pr(\alpha_1 < 1/2) = \Pr(\alpha_1 > 2) = 0.244$.

Coefficient α_2 is associated with Q_i^2 when $Q_i \leq \gamma$. Like α_0 and α_1, it must be non-negative. Therefore, we use a half-normal distribution with scale parameter equal to one. This gives a variance of 0.60 kg/h^{-2}. In units of kg/h, this is equivalent to a variance of 1.29.

Parameter β_1 (unitless, non-negative) 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. We must constrain the parameter γ to the range of Q_i in the CR data; otherwise it cannot be estimated. We use a uniform prior distribution on $(0, Q_{max})$ where Q_{max} is the largest value of Q_i observed in the CR 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 ref 42.

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 and Discussion section, the model results were insensitive to the prior specification.

Model Selection. As discussed in previous sections, 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 CR data better but may also be prone to over-fitting, leading to poor predictive performance. Therefore, we use a combination of Deviance Information Criteria, 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 (DIC) combines goodness-of-fit to the training data and model complexity to provide an overall assessment of the model.[39] It is analogous to the AIC, a frequentist model selection tool used by Conrad et al.[25] 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.[43]

We are also guided by plotting prediction bands derived from the posterior predictive distribution described in SI 2 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 indicates that the variance is not modelled well. The generalizability of the model may also be assessed by comparing the prediction bands to additional data that was excluded from the model fit ("external data"). If the model predictions are consistent with the external validation data, this is an encouraging sign that either the CR data was collected in diverse conditions, or the bias and variance of the measurements do not depend on factors which differ between the training and external data. In either of these scenarios, the results are generalizable to 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, $M_i - \hat{M}_i$ provides further 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 show structured and/or large residuals, this indicated that the median parameterization may be need to be modified. 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.

Quantifying Measurement Uncertainty in New Measurements. A key application of the models is to calculate a credible interval for the true emission rate based on methane measurements made in the field. Let M^{new} be the new non-zero measurement made in the field and Q^{new} represent the unknown true emission rate associated with M^{new}. We wish to know the distribution of Q^{new} given M^{new} and our

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

uncertainty model derived from CR data. The distribution of interest is $p(Q^{new} \mid M^{new}, \underline{M})$. Using Bayes equation (eq 1 in SI 2), we can say that

$$\begin{aligned} p(Q^{new}|M^{new}, \underline{M}) &\propto p(M^{new}|Q^{new}, \underline{M})p(Q^{new}|\underline{M}) \\ &= p(M^{new}|Q^{new}, \underline{M})p(Q^{new}) \end{aligned} \tag{8}$$

where the new measurement and true emission rate are modeled as independent of the CR data. That is, we assume the distribution of true emission rates investigated in the CR trials does not provide any information about the true emission rate observed in the field. The distribution $p(Q^{new})$ can be thought of as a prior distribution representing our beliefs about the emission rate distribution we expect to see. The distribution $p(M^{new} \mid Q^{new}, \underline{M})$ is the likelihood of observing M^{new} given Q^{new} and $\underline{M}$. If we assume that the new measurements obey the same uncertainty model as the CR data, then this distribution is the likelihood selected in the previous section for the given technology, marginalised over the posterior distribution of the statistical model parameters $\underline{\theta}$.

In SI 2, we describe how a weighted bootstrap algorithm[39,44] can be used to produce the probability distribution of the true emission rate given the new measurement(s), CR data, and prior knowledge about the distribution of emissions shown in eq 8. Briefly, we specify a prior distribution $p(Q^{new})$ for the true measurement, for example, based on previous emissions survey data. A sample of size L is drawn from the prior distribution to give $\{Q_1^{new}, ..., Q_L^{new}\}$. The value of the likelihood $p(M^{new} \mid Q_l^{new}, \underline{M})$ for $l = 1, ..., L$ is then determined computationally. We then re-sample from the initial prior sample K times with weights proportional to the likelihood value. The final sample represents a sample from the desired posterior distribution. The key inputs are the new measurement(s), the specification of the prior, and the form of the likelihood and statistical model parameters determined by the previous section. The algorithm depends on the assumption that the new measurements follow the same uncertainty model as the CR data.

RESULTS AND DISCUSSION

Uncertainty Results. In this section, we present the selected likelihoods for five different methane quantification technologies/operators. The chosen models are summarized in Table 4. The results are conditional on the CR data used to fit the models, therefore we assess the generalizability of the models to an arbitrary release scenario with different conditions by comparing to external data, if available. We also compare the model predictions to a line representing perfect prediction.

QOGI Technologies. Prediction bands and posterior median predictions are shown for QOGI Operators A, B, and C in Figures 2, 3, and 4, 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, and a linear function at larger release rates. QOGI Operator A is best modelled using a quadratic function for the median. The releases measured by Operator A have a more 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 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 observed during the field campaign. This lack-of-familiarity manifests as an additional factor that influences (broadens and biases) the likelihood.

QOGI Operator C was present for both campaigns. Only 14 measurements were made for this technology at the first campaign, which we use as external data. These data points fall within the 95% prediction band, suggesting that the model is generalizable. The other technology providers were only present for one campaign and thus no external data are available for these providers.

TDLAS. Results from the selected models for truck and aerial TDLAS systems are shown in Figure 5 and the left panel of Figure 6. The truck-based TDLAS tends to underestimate emissions, while the aerial-based TDLAS overestimates on average.

For truck-based TDLAS, the model was fit using data from the second campaign, while data from the first campaign were 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 the two campaigns. The generalizability would be improved if more CR data were collected and included in the model fitting.

Figure 6 compares the predictions resulting from the Bayesian uncertainty model derived for the airborne TDLAS and the one presented by Conrad et al.[25] 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 Conrad et al.[25] uses a constant variance whereas the present model allows the variance to change with Q_i. The median predictions are very similar between the two models.

Airborne NIR HS Imaging. The prediction bands from the selected model for the airborne NIR HS imaging technology are shown in Figure 7. The technology tends to overestimate emissions. The model was fit to data from the second campaign, while data from the first campaign were used as external data. The technology tended to underestimate the true releases during the first campaign and overestimate the true releases during the second campaign. Conditions during the first campaign were considered marginal due to excessive cloud cover and atypical of those under which commercial measurements were conducted, while those of the second

Table 4. Summary of Selected Models for Each Methane Quantification Technology Provider

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

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

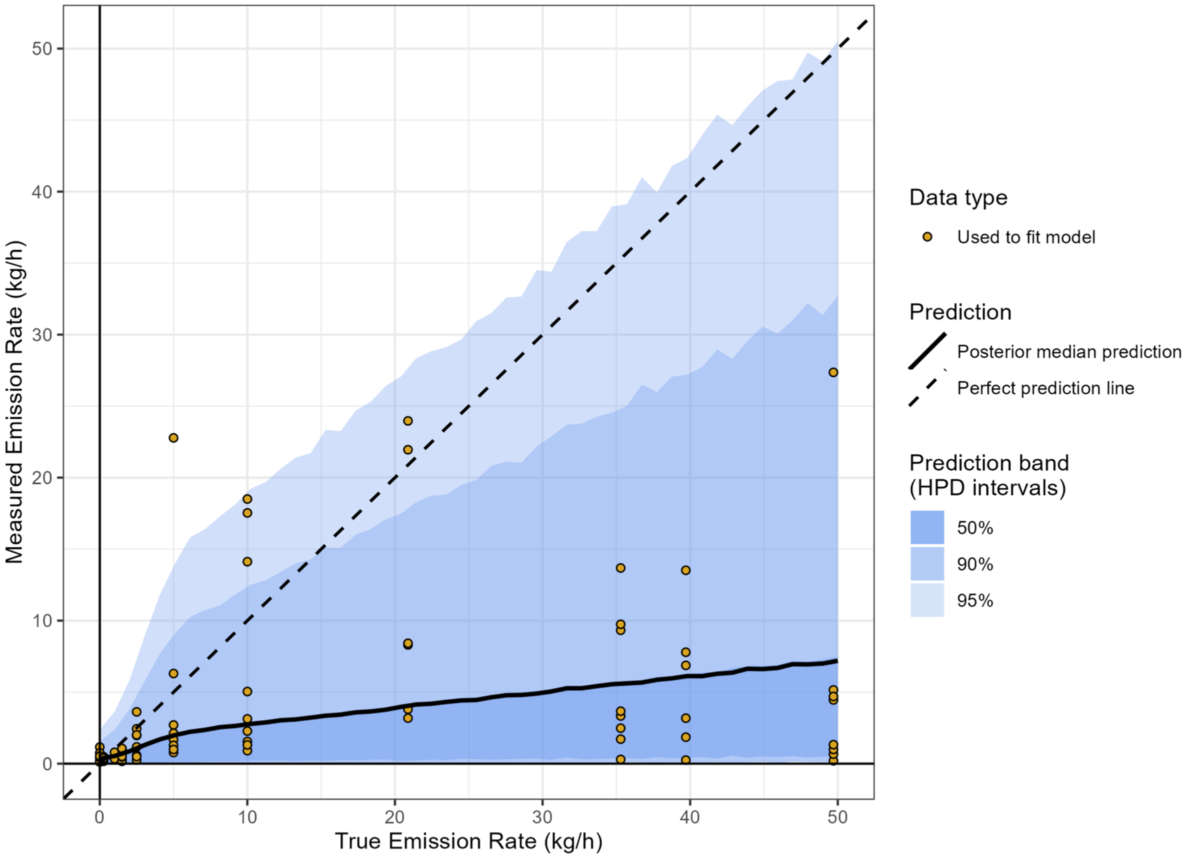


Figure 3. Uncertainty quantification model results for QOGI Operator B. The model was fit to data from the second campaign. No external data were available.

campaign were nearly ideal. Given the different conditions between the campaigns, it is not surprising that the model derived from data taken from the second campaign consistently overestimates the observed data from the first campaign. Therefore, the model is not generalizable.

Application: Quantifying Uncertainty in New Measurements. We now demonstrate how the models developed in the previous section may be used to estimate uncertainty in a new measurement using the QOGI Operator C model as an example. The weighted bootstrap algorithm described in SI 2 was followed with $L = 5000$, $J = 10{,}000$, and $K = 4000$. When conducting QOGI measurements, a key operational consideration are the number of independent measurements that should be conducted by the operator, given the well-known variability of this technology. Accordingly, we consider two measurement 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 as follows: First, we choose a hypothetical true value for the source we will measure, $Q_i = 25$ kg/h. Then, we simulate measurements by drawing from the posterior predictive distribution defined in SI 2, $p(\tilde{M}_i \mid M_i, Q_i = 25 \text{ kg/h})$. To demonstrate how prior information impacts the estimates, we consider two different prior distributions, reflecting different degrees of prior information that may be available to the operator. We investigate four different scenarios:

1. Only one measurement, $M_1^{new} = 19.7$ kg/h, is drawn, and $p(Q^{new})$ is a uniform distribution from 0 to 200 kg/h.
2. Five measurements are drawn, $M_1^{new} = 19.7$ kg/h, $M_2^{new} = 11.6$ kg/h, $M_3^{new} = 8.4$ kg/h, $M_4^{new} = 18.1$ kg/h, and $M_5^{new} = 17.0$ kg/h, and $p(Q^{new})$ is the same uniform distribution as Scenario 1
3. Only one measurement (M_1^{new}) is used, and $p(Q^{new})$ is a log-normal distribution, which is often used to model leak rate distributions.[45] We set the shape parameter to 1, location parameter equal to zero, and scale parameter equal to 2.6.[41]
4. All of M_1^{new}, ..., M_5^{new} are used, and $p(Q^{new})$ is the same log-normal distribution as Scenario 3

The different information expressed in the two priors is shown in the histograms in Figure 8.

The posterior distributions shown in Figure 9 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 Figure 4, where it is clear that the technology systematically underestimates the true emission rate. Also apparent is the important role played by the prior information: The narrower prior distribution provides more information and narrows the posterior PDF. An important caveat to this is that if the chosen prior is a poor representation of the area under study, then the resulting posterior will be less accurate. For this reason, it is crucial that the prior PDF, in terms of both its distribution type and width, accurately represents the true state of prior knowledge. Tools like maximum entropy priors may be deployed to minimize the information content of

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

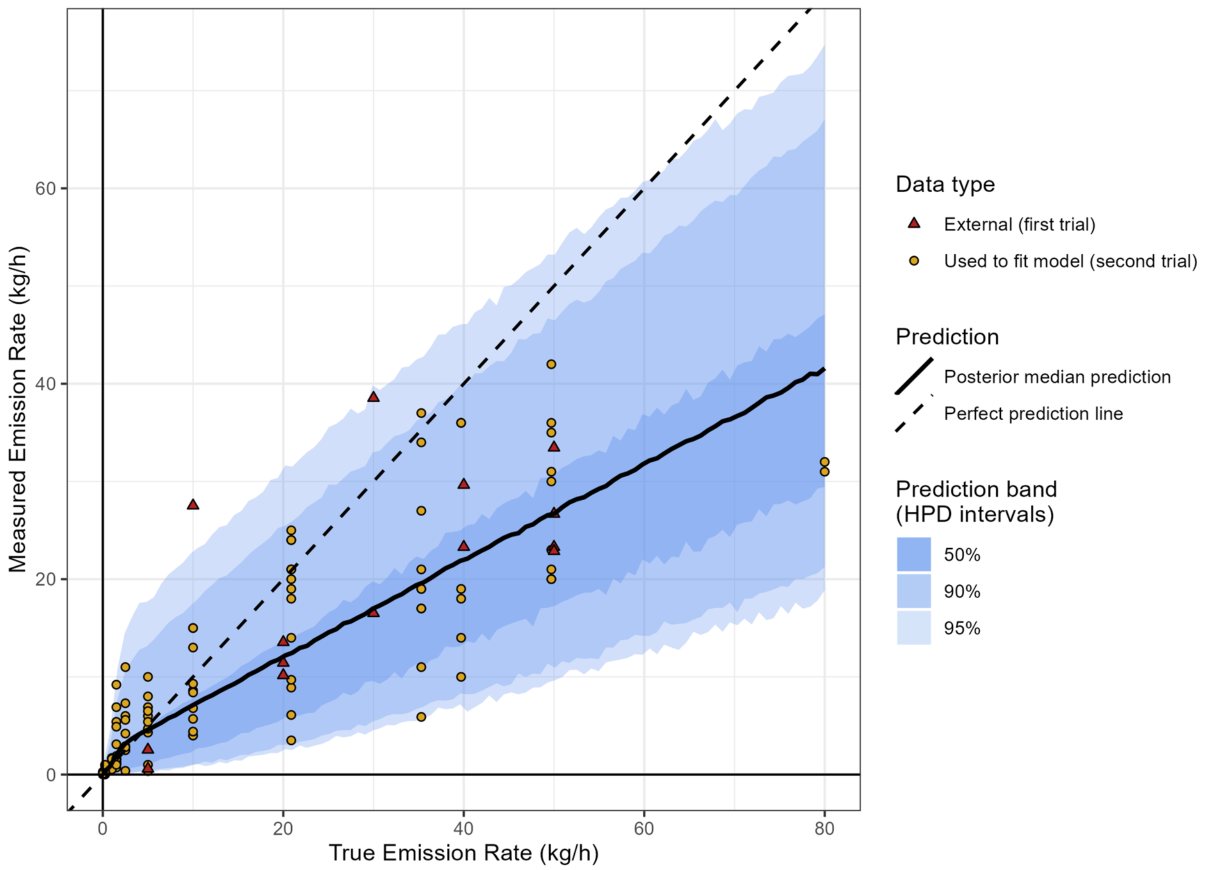


Figure 4. Uncertainty quantification model results from QOGI Operator C. The model was fit to data from the second campaign. CR data from the first campaign was used as external data.

the priors subject to constraints imposed by testable information.

An important consideration when applying this algorithm is that the error model will only reflect the true measurement uncertainty if the new measurement was obtained under conditions similar to those under which the CR data were obtained. In particular, we caution against fitting error models to CR data collected only in ideal operating conditions, because the resulting CIs will be too narrow. In summary, the algorithm provides a CI derived from new measurement(s) which is centred around the true emission rate and whose width reflects measurement uncertainty. The results are subject to the sensibility of the prior and the agreement between the CR conditions and the new measurements.

Discussion. Methane emission estimates can only be interpreted properly in the context of uncertainty. This paper presented a formalism for developing estimates of measurement uncertainty from CR data within the Bayesian framework. The outcome of this analysis are posterior probability distributions that comprehensively define what is known about an emission rate, based on the measurement data, CR 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 posterior distributions may be summarized as CIs (e.g., the range of emission rates that correspond to a given probability) and used for other purposes, such as inputs to probabilistic simulations to assess the effectiveness of alternative fugitive emissions management plans (alt-FEMPs) or in calculating the uncertainty attached to inventory estimates.

As highlighted by Figures 2–7, relying solely on measurements without considering their uncertainty can lead to significant misinterpretations of the underlying emission rate. The methods presented in this work provide a way to summarize both the variability and systematic bias of a technology. They are situated in the Bayesian statistical framework which facilitates probabilistic inference, the derivation of credible intervals, and downstream approaches.

Specifically, the weighted bootstrap algorithm provides a distribution of the true emission rate given all available information, including a measurement or set of measurements, CR 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 become available. This is beneficial because it formalizes an informal process: in the absence of data, we must rely more on previous knowledge, and when more data are available, we rely less on our previous knowledge. The results also show that increasing the number of measurements reduces the width of 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 CIs have a given width. This algorithm in tandem with the models derived from CR data could be used in the future to help plan or assess the effectiveness of LDAR programs, e.g., by identifying the

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

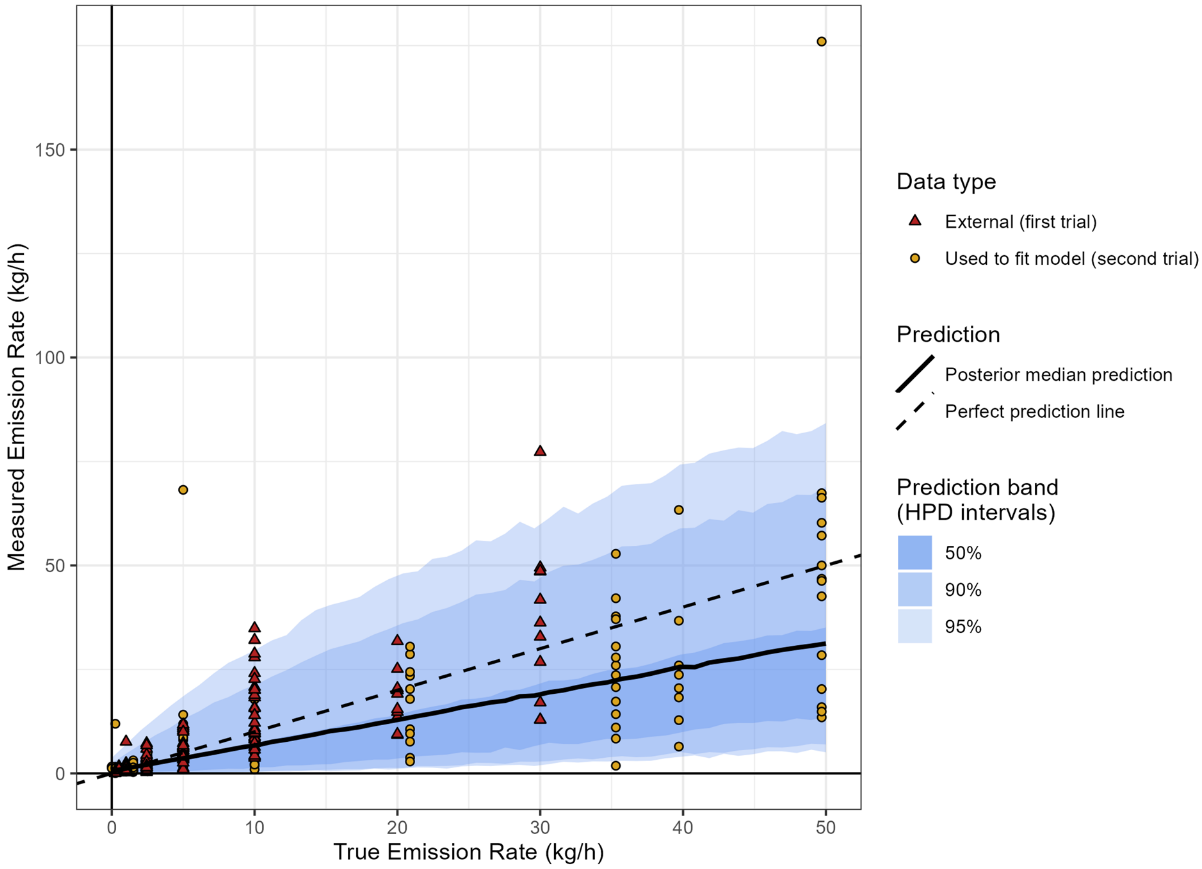


Figure 5. Uncertainty quantification model results for truck-based TDLAS data. The model was fit to data from the second campaign. CR data from the first campaign was used as external data.

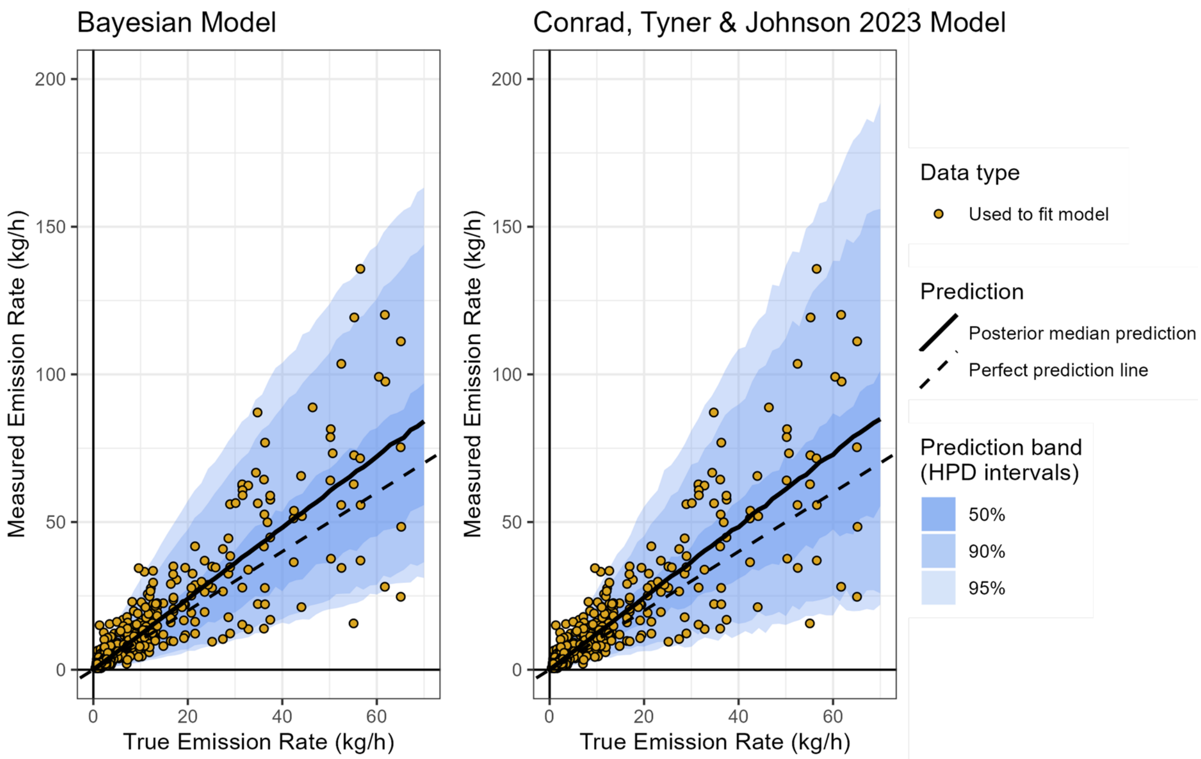


Figure 6. Side-by-side comparison of Bayesian model proposed in Table 4 and that presented in Conrad et al.[25] for airborne TDLAS.

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

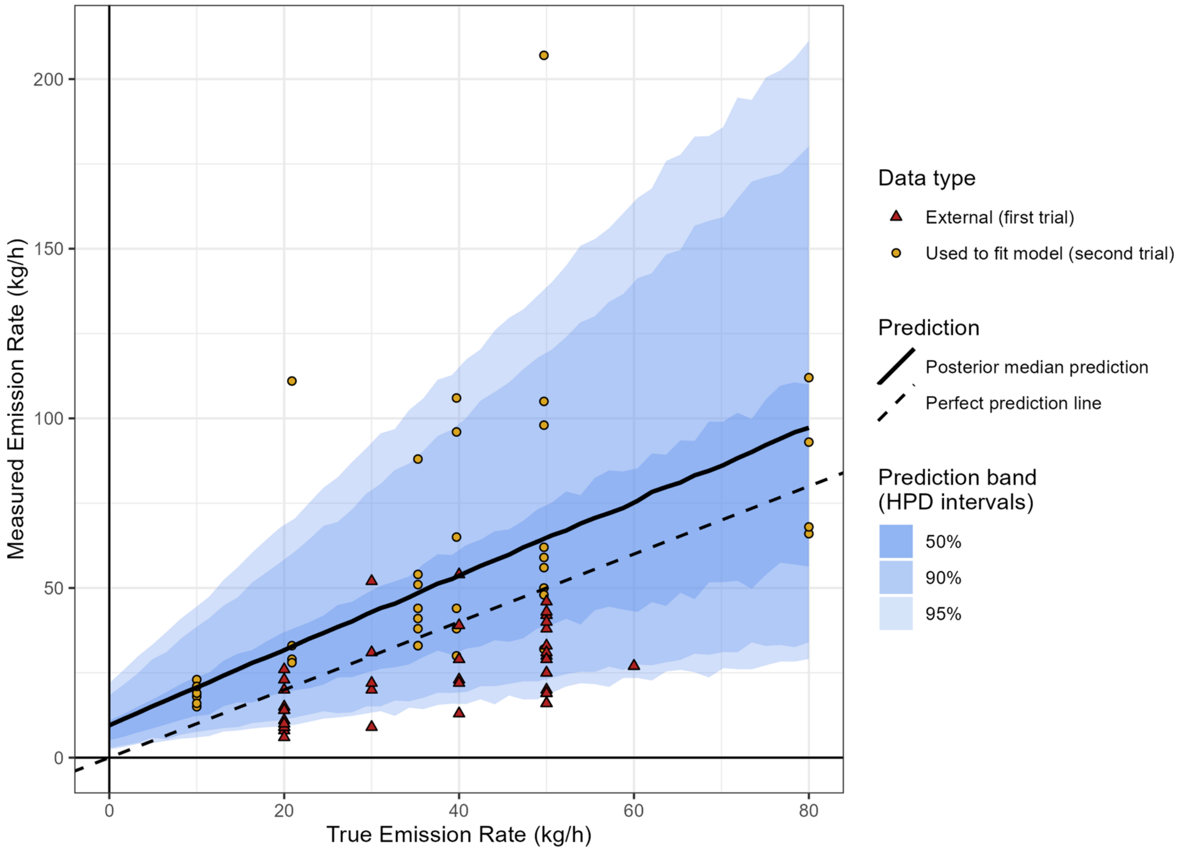


Figure 7. Uncertainty quantification model results for airborne NIR HS data. The model was fit to data from the second campaign. CR data from the first campaign was used as external data.

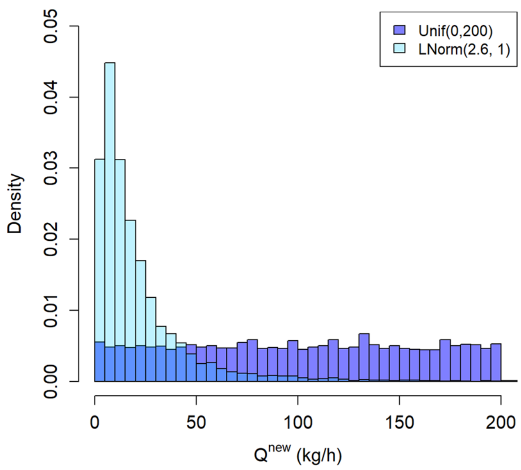


Figure 8. Comparison of histograms for the two different prior distributions $p(M^{new})$ investigated in the analysis.

optimal combination of technologies that achieve a certain precision.

While not the focus of this work, this study also demonstrates the importance of multiple measurements during any particular emission survey study. Table 5 shows that the uncertainty in estimating a true emission rate is reduced when we have five measurements compared to only one measurement. The 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. A key result of this study is that the uncertainty of a given survey depends on the number of measurements made. 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.

Although extensive meteorological data were collected during the campaigns (as detailed in SI 1), we refrain from incorporating them in the statistical models. This is because the goal of the models is to summarize the performance of a technology over a variety of conditions. To this end, the campaigns were conducted at different times of year, involved different release structures, and each trial was carried out over multiple days, so that the results could be used to assess the performance of the technologies over a variety of conditions. However, it may be possible to improve the predictive performance of the models by incorporating meteorological data into the likelihood.

While we scheduled the measurement campaigns at two times of the year over multiple days to capture variations in environmental conditions, the limited time allowed for the measurements means that not all weather conditions are encompassed in the CR data. The measurement campaign conditions also departed from those that prevail at typical upstream oil and gas sites, e.g., the presence of nearby structures that could complicate the advection models. The

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

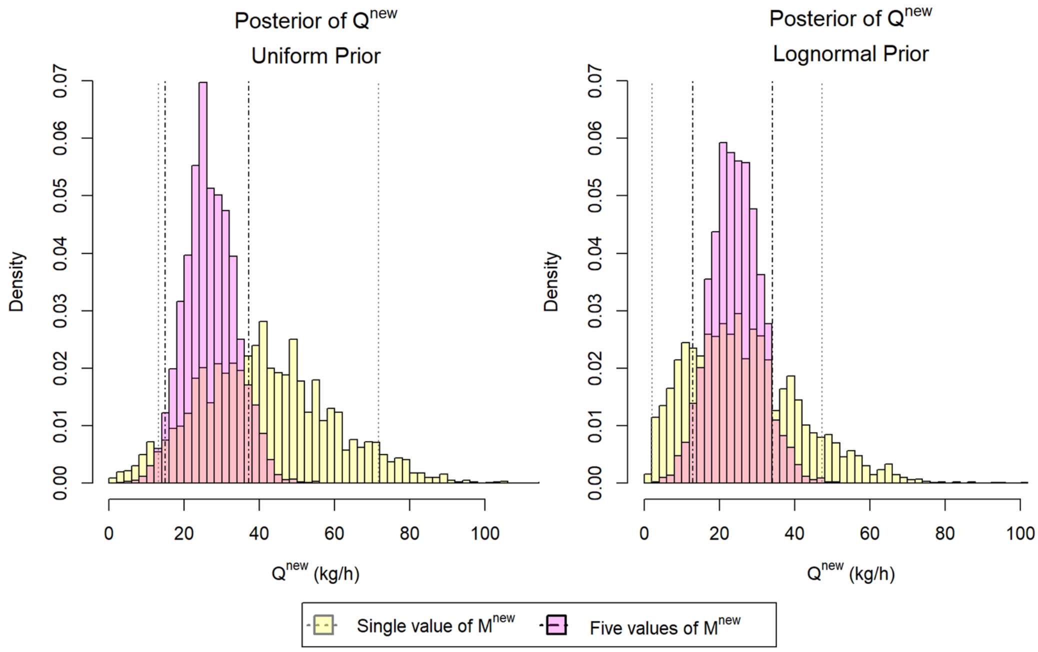


Figure 9. Posterior distributions for the true value of the emission rate given the observed measurement(s), prior distribution, and CR data.

Table 5. Summary of 90% Credible Intervals (CI) for the Four Different Data Scenarios and Their Lengths

Prior	Number of Values	90% CI	Length of CI
Unif(0, 200)	1	(13.1, 71.7)	58.6
Unif(0, 200)	5	(15.0, 37.2)	22.2
LogNormal(2.6, 1)	1	(2.1, 47.3)	45.2
LogNormal(2.6, 1)	5	(13.0, 34.1)	21.1

external validation indeed shows that for some technologies (truck-based TDLAS, airborne NIR HSI), the measurement error in the first campaign data behaves differently than predicted by the model fitted to the second campaign data. A similar result was reported recently by Day et al.,[46] who found that the detection probability and measurement accuracy of fixed sensor systems were significantly lower when installed at upstream oil and gas facilities compared to dedicated testing facilities. This finding highlights the important role played by latent environmental factors in quantification uncertainty. This is evidence that more diverse CR data are needed for broadly applicable measurement uncertainty models.

Overall quantification uncertainty comes from detection probability and measurement uncertainty. The methods in this paper address measurement uncertainty only. These components may be considered separately, e.g., by Conrad et al.[25] They found that the detection probability of Bridger's GML is related to different factors than the measurement uncertainty, emphasizing that the two sources of uncertainty are distinct. An advantage of the Bayesian approach taken in this paper is that it lends itself well to model extension. A possible avenue of future work is the modelling of detection probability and measurement uncertainty simultaneously using a hierarchical Bayesian model.

■ ASSOCIATED CONTENT

Data Availability Statement

Example code for fitting the models and anonymized CR data for both campaigns are available at https://github.com/augustinewigle/methaneUQ.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsestair.4c00030.

Information on the set-up, meteorological conditions, and release rates for the two measurement campaigns (PDF)

Overview of Bayesian analysis, pseudocode to aid in the implementation of the models, and the algorithm for quantifying uncertainty in a new measurement (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Augustine Wigle − *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada;* orcid.org/0000-0003-0225-1949; Email: amhwigle@uwaterloo.ca

Authors

Audrey Béliveau − *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

Daniel Blackmore − *Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

Paule Lapeyre − *Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

Kirk Osadetz − *Carbon Management Canada, Bow City, Alberta T0J 2M0, Canada*

Christiane Lemieux − *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014

Kyle J. Daun – *Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada;* orcid.org/0000-0001-9611-7462

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsestair.4c00030

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

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. AW gratefully acknowledges the support of Natural Sciences and Engineering Research Council through a Canada Graduate Scholarship.

REFERENCES

(1) International Energy Agency. *Global Methane Tracker 2023 – Analysis*. https://www.iea.org/reports/global-methane-tracker-2023 (accessed 2024-02-10).

(2) Calvin, K.; Dasgupta, D.; Krinner, G.; Mukherji, A.; Thorne, P. W.; Trisos, C.; Romero, J.; Aldunce, P.; Barrett, K.; Blanco, G.; Cheung, W. W.; Connors, S.; Denton, F.; Diongue-Niang, A.; Dodman, D.; Garschagen, M.; Geden, O.; Hayward, B.; Jones, C.; Jotzo, F.; Krug, T.; Lasco, R.; Lee, Y.-Y.; Masson-Delmotte, V.; Meinshausen, M.; Mintenbeck, K.; Mokssit, A.; Otto, F. E.; Pathak, M.; Pirani, A.; Poloczanska, E.; Pörtner, H.-O.; Revi, A.; Roberts, D. C.; Roy, J.; Ruane, A. C.; Skea, J.; Shukla, P. R.; Slade, R.; Slangen, A.; Sokona, Y.; Sörensson, A. A.; Tignor, M.; Van Vuuren, D.; Wei, Y.-M.; Winkler, H.; Zhai, P.; Zommers, Z.; Hourcade, J.-C.; Johnson, F. X.; Pachauri, S.; Simpson, N. P.; Singh, C.; Thomas, A.; Totin, E.; Arias, P.; Bustamante, M.; Elgizouli, I.; Flato, G.; Howden, M.; Méndez-Vallejo, C.; Pereira, J. J.; Pichs-Madruga, R.; Rose, S. K.; Saheb, Y.; Sánchez Rodríguez, R.; Ürge Vorsatz, D.; Xiao, C.; Yassaa, N.; Alegría, A.; Armour, K.; Bednar-Friedl, B.; Blok, K.; Cissé, G.; Dentener, F.; Eriksen, S.; Fischer, E.; Garner, G.; Guivarch, C.; Haasnoot, M.; Hansen, G.; Hauser, M.; Hawkins, E.; Hermans, T.; Kopp, R.; Leprince-Ringuet, N.; Lewis, J.; Ley, D.; Ludden, C.; Niamir, L.; Nicholls, Z.; Some, S.; Szopa, S.; Trewin, B.; Van Der Wijst, K.-I.; Winter, G.; Witting, M.; Birt, A.; Ha, M.; Romero, J.; Kim, J.; Haites, E. F.; Jung, Y.; Stavins, R.; Birt, A.; Ha, M.; Orendain, D. J. A.; Ignon, L.; Park, S.; Park, Y.; Reisinger, A.; Cammaramo, D.; Fischlin, A.; Fuglestvedt, J. S.; Hansen, G.; Ludden, C.; Masson-Delmotte, V.; Matthews, J. R.; Mintenbeck, K.; Pirani, A.; Poloczanska, E.; Leprince-Ringuet, N.; Péan, C. *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*; Intergovernmental Panel on Climate Change (IPCC): 2023.

(3) Canfin, P.; Jutta, P. *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.* https://www.europarl.europa.eu/doceo/document/A-9-2023-0162_EN.html (accessed 2023-11-23).

(4) UNEP. *OGMP 2.0 The Oil & Gas Methane Partnership 2.0.* https://ogmpartnership.com/ (accessed 2023-11-23).

(5) Fox, T. A.; Hugenholtz, C. H.; Barchyn, T. E.; Gough, T. R.; Gao, M.; Staples, M. Can new mobile technologies enable fugitive methane reductions from the oil and gas industry? *Environmental Research Letters* **2021**, *16*, 064077.

(6) Solazzo, E.; Crippa, M.; Guizzardi, D.; Muntean, M.; Choulga, M.; Janssens-Maenhout, G. Uncertainties in the Emissions Database for Global Atmospheric Research (EDGAR) emission inventory of greenhouse gases. *Atmospheric Chemistry and Physics* **2021**, *21*, 5655–5683.

(7) MacKay, K.; Lavoie, M.; Bourlon, E.; Atherton, E.; O'Connell, E.; Baillie, J.; Fougère, C.; Risk, D. Methane emissions from upstream oil and gas production in Canada are underestimated. *Sci. Rep.* **2021**, *11*, 8041.

(8) Conrad, B. M.; Tyner, D. R.; Li, H. Z.; Xie, D.; Johnson, M. R. A measurement-based upstream oil and gas methane inventory for Alberta, Canada reveals higher emissions and different sources than official estimates. *Communications Earth & Environment* **2023**, *4*, 416.

(9) Tyner, D. R.; Johnson, M. R. Where the Methane Is–Insights from Novel Airborne LiDAR Measurements Combined with Ground Survey Data. *Environ. Sci. Technol.* **2021**, *55*, 9773–9783.

(10) Wilson, A. New Optical Gas-Imaging Technology for Quantifying Fugitive-Emission Rates. *Journal of Petroleum Technology* **2016**, *68*, 78–79.

(11) Mitchell, A. L.; Tkacik, D. S.; Roscioli, J. R.; Herndon, S. C.; Yacovitch, T. I.; Martinez, D. M.; Vaughn, T. L.; Williams, L. L.; Sullivan, M. R.; Floerchinger, C.; Omara, M.; Subramanian, R.; Zimmerle, D.; Marchese, A. J.; Robinson, A. L. Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results. *Environ. Sci. Technol.* **2015**, *49*, 3219–3227.

(12) Kang, R.; Liatsis, P.; Kyritsis, D. C. Emission Quantification via Passive Infrared Optical Gas Imaging: A Review. *Energies* **2022**, *15*, 3304.

(13) Heltzel, R. S.; Johnson, D. R.; Zaki, M. T.; Gebreslase, A. K.; Abdul-Aziz, O. I. Machine learning techniques to increase the performance of indirect methane quantification from a single, stationary sensor. *Heliyon* **2022**, *8*, e11962.

(14) Brantley, H. L.; Thoma, E. D.; Squier, W. C.; Guven, B. B.; Lyon, D. Assessment of Methane Emissions from Oil and Gas Production Pads using Mobile Measurements. *Environ. Sci. Technol.* **2014**, *48*, 14508–14515.

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

(16) Caulton, D. R.; Li, Q.; Bou-Zeid, E.; Fitts, J. P.; Golston, L. M.; Pan, D.; Lu, J.; Lane, H. M.; Buchholz, B.; Guo, X.; McSpiritt, J.; Wendt, L.; Zondlo, M. A. Quantifying uncertainties from mobile-laboratory-derived emissions of well pads using inverse Gaussian methods. *Atmospheric Chemistry and Physics* **2018**, *18*, 15145–15168.

(17) Sherwin, E. D.; Chen, Y.; Ravikumar, A. P.; Brandt, A. R. Single-blind test of airplane-based hyperspectral methane detection via controlled releases. *Elementa: Science of the Anthropocene* **2021**, *9*, 00063.

(18) Varon, D. J.; McKeever, J.; Jervis, D.; Maasakkers, J. D.; Pandey, S.; Houweling, S.; Aben, I.; Scarpelli, T.; Jacob, D. J. Satellite Discovery of Anomalously Large Methane Point Sources From Oil/Gas Production. *Geophysical Research Letters* **2019**, *46*, 13507–13516.

(19) Montazeri, A.; Zhou, X.; Albertson, J. D. On the Viability of Video Imaging in Leak Rate Quantification: A Theoretical Error Analysis. *Sensors* **2021**, *21*, 5683.

(20) Cambaliza, M. O. L.; Shepson, P. B.; Caulton, D. R.; Stirm, B.; Samarov, D.; Gurney, K. R.; Turnbull, J.; Davis, K. J.; Possolo, A.; Karion, A.; Sweeney, C.; Moser, B.; Hendricks, A.; Lauvaux, T.; Mays, K.; Whetstone, J.; Huang, J.; Razlivanov, I.; Miles, N. L.; Richardson, S. J. Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions. *Atmospheric Chemistry and Physics* **2014**, *14*, 9029–9050.

(21) Edie, R.; Robertson, A. M.; Field, R. A.; Soltis, J.; Snare, D. A.; Zimmerle, D.; Bell, C. S.; Vaughn, T. L.; Murphy, S. M. Constraining the accuracy of flux estimates using OTM 33A. *Atmospheric Measurement Techniques* **2020**, *13*, 341–353.

(22) Singh, D.; Barlow, B.; Hugenholtz, C.; Funk, W.; Robinson, C.; Ravikumar, A. P. Field Performance of New Methane Detection

Technologies: Results from the Alberta Methane Field Challenge. *EarthArXiv*, 2021-06-04. https://eartharxiv.org/repository/view/1860/ (accessed 2023-11-24).
(23) Johnson, M. R.; Tyner, D. R.; Szekeres, A. J. Blinded evaluation of airborne methane source detection using Bridger Photonics LiDAR. *Remote Sensing of Environment* **2021**, *259*, 112418.
(24) Bell, C.; Rutherford, J.; Brandt, A.; Sherwin, E.; Vaughn, T.; Zimmerle, D. Single-blind determination of methane detection limits and quantification accuracy using aircraft-based LiDAR. *Elementa: Science of the Anthropocene* **2022**, *10*, 00080.
(25) Conrad, B. M.; Tyner, D. R.; Johnson, M. R. Robust probabilities of detection and quantification uncertainty for aerial methane detection: Examples for three airborne technologies. *Remote Sensing of Environment* **2023**, *288*, 113499.
(26) Kemp, C. E.; Ravikumar, A. P.; Brandt, A. R. Comparing Natural Gas Leakage Detection Technologies Using an Open-Source "Virtual Gas Field" Simulator. *Environ. Sci. Technol.* **2016**, *50*, 4546−4553.
(27) Fox, T. A.; Gao, M.; Barchyn, T. E.; Jamin, Y. L.; Hugenholtz, C. H. An agent-based model for estimating emissions reduction equivalence among leak detection and repair programs. *Journal of Cleaner Production* **2021**, *282*, 125237.
(28) Lavoie, M.; Risk, D.; O'Connell, E.; Atherton, E.; Gorski, J.; Johnson, J. Evaluating the benefits of alternative leak detection programs. *EarthArXiv*, 2021-05-18. https://eartharxiv.org/repository/view/2355/ (accessed 2023-11-24).
(29) Johnson, M. R.; Conrad, B. M.; Tyner, D. R. Creating measurement-based oil and gas sector methane inventories using source-resolved aerial surveys. *Communications Earth & Environment* **2023**, *4*, 139.
(30) Zeng, Y.; Morris, J. Detection limits of optical gas imagers as a function of temperature differential and distance. *J. Air Waste Manage. Assoc.* **2019**, *69*, 351−361.
(31) Ravikumar, A. P.; Wang, J.; McGuire, M.; Bell, C. S.; Zimmerle, D.; Brandt, A. R. Good versus Good Enough?" Empirical Tests of Methane Leak Detection Sensitivity of a Commercial Infrared Camera. *Environ. Sci. Technol.* **2018**, *52*, 2368−2374.
(32) Zimmerle, D.; Vaughn, T.; Bell, C.; Bennett, K.; Deshmukh, P.; Thoma, E. Detection Limits of Optical Gas Imaging for Natural Gas Leak Detection in Realistic Controlled Conditions. *Environ. Sci. Technol.* **2020**, *54*, 11506−11514.
(33) Supplee, J. M.; Whittaker, E. A.; Lenth, W. Theoretical description of frequency modulation and wavelength modulation spectroscopy. *Applied Optics* **1994**, *33*, 6294−6302.
(34) Flesch, T. K.; Wilson, J. D.; Yee, E. Backward-Time Lagrangian Stochastic Dispersion Models and Their Application to Estimate Gaseous Emissions. *Journal of Applied Meteorology and Climatology* **1995**, *34*, 1320−1332.
(35) Flesch, T. K.; Wilson, J. D.; Harper, L. A.; Crenna, B. P.; Sharpe, R. R. Deducing Ground-to-Air Emissions from Observed Trace Gas Concentrations: A Field Trial. *Journal of Applied Meteorology and Climatology* **2004**, *43*, 487−502.
(36) Esparza, A. E.; Gauthier, J.-F. Monitoring Leaks at Oil and Gas Facilities Using the Same Sensor on Aircraft and Satellite Platforms; *Proceedings of the 9th Unconventional Resources Technology Conference*; 2021; DOI: 10.15530/urtec-2021-5666 (accessed 2023-11-29).
(37) Esparza, A. E.; Rowan, G.; Newhook, A.; Deglint, H. J.; Garrison, B.; Orth-Lashley, B.; Girard, M.; Shaw, W. 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* **2023**, *178*, 113265.
(38) Varon, D. J.; Jacob, D. J.; McKeever, J.; Jervis, D.; Durak, B. O. A.; Xia, Y.; Huang, Y. Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmospheric Measurement Techniques* **2018**, *11*, 5673−5686.
(39) Gelman, A.; Carlin, J. B.; Stern, H. S.; Dunson, D. B.; Vehtari, A.; Rubin, D. B. *Bayesian Data Analysis*, 3rd ed.; CRC Press: 2013.
(40) Zhou, X.; Montazeri, A.; Albertson, J. D. Mobile sensing of point-source gas emissions using Bayesian inference: An empirical examination of the likelihood function. *Atmospheric Environment* **2019**, *218*, 116981.
(41) NIST. NIST Engineering Statistics Handbook 1.3.6.6.9 Lognormal Distribution. *NIST Engineering Statistics Handbook*, 2023. https://www.itl.nist.gov/div898/handbook/eda/section3/eda3669.htm (accessed 2023-06-28).
(42) Gelman, A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis* **2006**, *1*, 515−534.
(43) Spiegelhalter, D. J.; Best, N. G.; Carlin, B. P.; Van Der Linde, A. Bayesian measures of model complexity and fit. *J. R Stat Soc. Series B Stat Methodology* **2002**, *64*, 583−639.
(44) Smith, A. F. M.; Gelfand, A. E. Bayesian Statistics without Tears: A Sampling-Resampling Perspective. *American Statistician* **1992**, *46*, 84−88.
(45) Zavala-Araiza, D.; Herndon, S. C.; Roscioli, J. R.; Yacovitch, T. I.; Johnson, M. R.; Tyner, D. R.; Omara, M.; Knighton, B. Methane emissions from oil and gas production sites in Alberta, Canada. *Elementa: Science of the Anthropocene* **2018**, *6*, 27.
(46) Day, R. E.; Emerson, E.; Bell, C.; Zimmerle, D. Point Sensor Networks Struggle to Detect and Quantify Short Controlled Releases at Oil and Gas Sites. *Sensors* **2024**, *24*, 2419.

https://doi.org/10.1021/acsestair.4c00030
ACS EST Air 2024, 1, 1000−1014