Quantitative comparison of methods used to estimate methane emissions from small point sources

Stuart N. Riddick[1], Riley Ancona[1], Clay Bell[1], Aidan Duggan[1], Tim Vaughn[1], Kristine Bennett[1] and Dan Zimmerle[1]

[1] The Energy Institute, Colorado State University, Fort Collins, CO, 80524, USA

5

Correspondence to: Stuart.Riddick@colostate.edu

Abstract Recent interest in quantifying trace gas emissions from point sources, such as measuring methane (CH_4) emissions from oil and gas wells, has resulted in several methods being used to estimate emissions from sources with emission rates below 200g CH_4 hour^{-1}. The choice of measurement approach depends on how close observers can get to the source, the instruments available

10 and the meteorological/micrometeorological conditions. As such, static chambers, dynamic chambers, HiFlow measurements, Gaussian plume modelling and backward Lagrangian stochastic (bLs) models have all been used, but there is no clear understanding of the accuracy or precision of each method. To address this, we copy the experimental design for each of the measurement methods to make single field measurements of a known source, to simulate single measurement field protocol, and then make repeat measurements to generate an understanding of the accuracy and precision of each method. Here, for comparison,

15 we present estimates for the percentage difference between the measured emission and the known emission, A, and the average percentage difference for three repeat measurements, A_r, for emissions of 200 g CH_4 h^{-1}. Our results show that, even though the dynamic chamber repeatedly underestimates the emission, it is the most accurate for a single measurement and the accuracy improves with subsequent measurements ($A = -11\%$, $A_r = -10\%$). The single HiFlow emission estimate was also an underestimate, however, poor instrument precision resulted in reduced accuracy of emission estimate to becomes less accurate after repeat

20 measurements ($A = -16\%$, $A_r = -18\%$). Of the far field methods, the bLs method underestimated emissions both for single and repeat measurements ($A = -11\%$, $A_r = -7\%$) while the GP method significantly overestimated the emissions ($A = 33\%$, $A_r = 29\%$) despite using the same meteorological and concentration data as input. Additionally, our results show that the accuracy and precision of the emission estimate increases as the flow rate of the source is increased for all methods. To our knowledge this is the first time that methods for measuring CH_4 emissions from point sources less than 200 g CH_4 h^{-1} have been quantitively assessed

25 against a known reference source and each other.

1 Introduction

Recent interest in quantifying trace gas emissions from point sources, such as measuring methane (CH_4) emissions from oil and gas wells (Omara et al., 2016; Pekney et al., 2018; Riddick et al., 2019a; Vaughn et al., 2018; Alvarez et al., 2018; Zavala-Araiza et al., 2021), has resulted in several methods being used to measure emissions from smaller point sources, i.e. less than 200g CH_4

30 hour^{-1}. Measurement approaches depend on how close observers can get to the source, due to access restrictions, instrumentation availability and the meteorological/micrometeorological conditions at the measurement site. Broadly put, these methods can be classed as direct measurements, i.e. touching/enclosing the source, and downwind measurements where access is not possible. Direct methods include: static chambers (Livingston and Hutchinson, 1995; Kang et al., 2014), dynamic flux chambers (Riddick et al., 2019a, 2020; Aneja et al., 2006) and Hi-Flow sampling (Pekney et al., 2018; Allen et al., 2013; Brantley et al., 2015). While

35 downwind methods include: Gaussian plume models (Baillie et al., 2019; Caulton et al., 2014; Riddick et al., 2017; Edie et al.,

2020; Bell et al., 2017) and Lagrangian dispersion models (Riddick et al., 2019a, 2017; Denmead, 2008; Flesch et al., 1995). Despite the interest in developing methods, emissions calculated using these methods have not been comprehensively compared using a known emission source, even though the results of these studies are being used to guide policy.

Static chambers have been used to estimate CH_4 emissions from abandoned wells that are very small sources (0.6 mg hr^{-1}) to those
40   that are much larger (90 g hr^{-1}) (Kang et al., 2014). This method is relatively simple, where a container of a known volume (V, m^3) is place over the emission source and the change in concentration (C, g m^{-3}) inside the container over time (t, s) can be used to calculate the emission (Q, g s^{-1}; Equation 1). Kang et al. (2014) reported using a gas chromatograph after the measurement campaign to assay CH_4 concentrations taken from the chamber at time intervals. Therefore, the static chamber method requires no power, apart from batteries to run a fan in the chamber and is very portable. Major shortcomings of this method are the uncertainty
45   caused by incomplete mixing of methane in the chamber, estimated at a factor of 2 (Kang et al., 2014), and large emission sources can result in the concentration inside the chamber exceeding the CH_4 lower explosive limit (LEL).

$$Q = \frac{dC}{dt} \cdot V \qquad \text{(Equation 1)}$$

To address LEL issues inside the chamber, a dynamic flux method has also been used to measure CH_4 leakage from abandoned and active oil and gas wells (Riddick et al., 2019a). Like the static chamber, the dynamic chamber comprises of a container enclosing the source and a propeller is used to circulate the air. Additionally, a flow of air is passed through the chamber, which
50   reduces the likelihood of exceeding LEL inside the chamber. Unlike the static chamber, the CH_4 concentration becomes stable after a period of time depending on the source emission rate. The CH_4 flux (Q, g s^{-1}) is calculated (Equation 2) from the CH_4 concentration at steady state (C_{eq}, g m^{-3}), the background CH_4 concentration (C_b, g m^{-3}) in the air used to flush the chamber, the height of the chamber (h, m), the flow of air through the chamber (q, m^3 s^{-1}), the area of the chamber (a, m^2) and the volume of the
55   chamber (V, m^3) (Aneja et al., 2006; Riddick et al., 2019a). As well as improving the safety, the dynamic chamber reduces the theoretical uncertainty in emission rate to $\pm$ 7% (Riddick et al., 2019a), however, the added power requirement of a pump means the dynamic chamber is less portable than the static chamber. Methane emissions have been calculated using this method between 4 µg CH_4 hr^{-1} and 100 g CH_4 hr^{-1} (Riddick et al., 2019a).

$$Q = \frac{(c_{eq} - c_b) \cdot h \cdot q \cdot a}{V} \qquad \text{(Equation 2)}$$

60   Another way of addressing the issue of enclosing methane at concentrations approaching LEL is to use a Hi-Flow sampler. As the name suggests, a Hi-Flow sampler draws high volumes of air – typical rates are 300 $l\ min^{-1}$ – into a measurement chamber, where the concentration of CH_4 in the air is measured and then calculates an emission rate. The current commercial HiFlow sampler uses two sensors, a catalytic oxidation sensor to measure CH_4 concentrations between 0 and 5% and a thermal conductivity sensor between 5% and 100% CH_4. These concentration measurements are then used to calculate emissions between 1 and 140 g
65   CH_4 hr^{-1} (Equation 3), at an accuracy of $\pm$ 10%. A recent independent study found the Bacharach HiFlow Sampler to be a suitable instrument for measuring CH_4 emissions if operated correctly within its limitations, but found the uncertainty largest when measuring at the transition of the sensors, i.e. ~5% methane (Connolly et al., 2019).

$$Q = F \cdot (X_s - X_b) \qquad \text{(Equation 3)}$$

In some circumstances, access and safety restrictions mean that direct measurements are impossible, and an observer must use a
70   far-field method to measure the emissions remotely. The most widely used of these far-field approaches is the Gaussian plume (GP) model. First used in the 1940s, a GP model describes the concentration of a gas as a function of distance downwind from a point source (Seinfeld and Pandis, 2016). When a gas is emitted from the source, it is entrained in the prevailing ambient air flow and disperses laterally and vertically with time, forming a dispersed concentration cone. The concentration of the gas (X, µg m^{-3}), at any point x meters downwind of the source, y meters laterally from the center line of the plume and z meters above ground level

Atmospheric
Measurement
Techniques
Discussions

75   can be calculated (Equation 4) using the emission rate (Q, g s^{-1}), the height of the source (h_s, m) and the Pasquill-Gifford stability

class (PGSC) as a measure of air stability. The standard deviation of the lateral (σ_y, m) and vertical (σ_z, m) mixing ratio distributions

are calculated from the PGSC of the air (Pasquill, 1962; Busse and Zimmerman, 1973; US EPA, 1995). The GP model assumes

that the vertical eddy diffusivity and wind speed are constant and there is total reflection of CH$_4$ at the surface. The GP is the

simplest of the far-field methods and assumes that the emissions are well-defined plumes injected above the near-surface turbulent

80   layer from point sources and not affected by aerodynamic obstructions that cause mechanical turbulence at the surface. However,

in most situations there are aerodynamic obstacles and plumes are rarely perfectly Gaussian in shape. Another shortcoming of the

GP model is the parameterization of the PGSC, which are discrete values and incorrectly assigning them can lead to significant

uncertainty. Generally, speaking the GP is rarely used for emissions less than 100 g CH4 h^{-1}. However, an example of using a GP

model is its use estimating CH$_4$ emissions from oil production platforms in the North Sea, where emissions ranged from 10 to 80

85   kg CH$_4$ hr^{-1} with an uncertainty of ± 45% (Riddick et al., 2019b)

$$X(x,y,z) = \frac{Q}{2\pi u \sigma_y \sigma_z} e^{-\frac{y^2}{(2\sigma_y)^2}} \left(e^{-\frac{(z-h_s)^2}{(2\sigma_z)^2}} + e^{-\frac{(z+h_s)^2}{(2\sigma_z)^2}} \right) \qquad \text{(Equation 4)}$$

As an alternative to the GP model, Lagrangian dispersion models can be used to calculate the emission of a source. In a backward

Lagrangian stochastic (bLS) model, the measurement position, gas concentration, meteorology and micrometeorology are known

inputs and the model works iteratively backwards to simulate the motion of the air parcel, this is then used to infer the rate of

90   emission from the source (Flesch et al., 1995). For given meteorological conditions, the model calculates the ratio of downwind

concentration to emission, $(C/Q)_{sim}$, depending on the size and location of the source. The emission rate (Q, g m^{-2} s^{-1}) is then

inferred from the measured gas concentration (X_m, g m^{-3}) and the background gas concentration (X_b, g m^{-3}) (Equation 5). The bLS

models can be used to calculate the emissions from point or area sources in a range of micrometeorological conditions. However,

a major shortcoming of the model is its inability to adequately model emissions from sources with complex topography or near

95   large objects, such as buildings. This can be mitigated by measuring far away from the source over a relatively flat fetch, but an

accurate measurement of the micrometeorology is required. As an example, CH$_4$ emissions from individual point sources on oil

and gas infrastructure have been estimated using a bLS model between 4 µg CH$_4$ hr^{-1} and 3 kg CH$_4$ hr^{-1} with an uncertainty of ±

38% (Riddick et al., 2019a)

$$Q = \frac{X_m - X_b}{\left(\frac{C}{Q}\right)_{sim}} \qquad \text{(Equation 5)}$$

100 In general, as access becomes more restricted, emission rates larger, or safety concerns increase, the methods to estimate the

emission rate of a source become more complex. We limit the scope of this study to estimating emissions from a single point

source that we would realistically be able to approach and measure, i.e. less than 200 g CH$_4$ h^{-1}. The study compares each method's

accuracy against known emission rates. Explicitly, our objectives are: 1) Copy the experimental design for each of the measurement

methods; 2) Conduct single measurements as a researcher would do in the field by taking measurements to generate a single

105 emission estimate from a point source and compare this to known emission rate; 3) Conduct repeat measurements to generate an

understanding of the accuracy and precision of the methods; and 4) Make recommendations on the suitability of each method for

measuring emissions from relatively small point sources. To our knowledge this is the first time that methods for measuring CH$_4$

emissions from point sources less than 200 g CH$_4$ h^{-1} have been quantitively assessed against a known reference source and each

other.

2 Methods

Each of the methods, static chambers, dynamic chambers, Hi-Flow, bLs and GP, are tested at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) site in Fort Collins, CO, USA. METEC can reproduce the range of CH_4 emissions typically seen from individual point sources at oil and gas operations, i.e. between 20 g CH_4 hr^{-1} and 40 kg CH_4 hr^{-1}, from realistic locations on O&G equipment. For the purposes of this study, where we are comparing the ability of each method to estimate the emission from a point source, we will constrain the known emission rates to those that can be measured safely, i.e. up to 200 g CH_4 hr^{-1}. To accomplish this, CH_4 emission rates will be set from a point source 20 cm above the ground at 40, 100 and 200 g CH_4 hr^{-1}.

Two instruments are used to report CH_4 mixing ratios: the Picarro (ww.picarro.com) GasScouter G4301 mobile gas concentration analyser and the Agilent (www.aglient.com) 7890B Flame-Ionization Detector Gas Chromatograph (GC-FID). The Picarro GasScouter reports CO_2, H_2O and CH_4 mixing ratios every 3 s, with a precision (300s, 1σ) for CH_4 of 300 ppb over an operating range of 0 to 800 ppm. The Agilent 7890B gas chromatograph-flame ionization detector (GC-FID), as used here, has a detection limit of 1.5 ppb and linear dynamic range from 1 ppm to 100% CH_4. The instrument was calibrated every 10 samples using a 5,000 ppm gas standard. The GC-FID was checked for linearity before and after each set of measurements using zero-air, 5,000 ppm, 2.5% and 100% CH_4.

2.1 Static Chamber

Following method descriptions presented in Pihalatie et al. (2013) and Collier et al. (2014), the static chamber is made by enclosing air within a fixed volume over the emission source (Figure 1A). A fan is secured inside the chamber and used to circulate the air. When the chamber is sealed with the ground an air sample is drawn. During the experiment at least three further air samples are taken at regular intervals (Pihlatie et al., 2013; Collier et al., 2014), the time interval was pre-calculated depending on the emission rate to ensure the increase in concentration was linear. The emission is then calculated from the linear increase in concentration over time.

Three sizes of static chambers were used in this experiment: 0.12 m^3, 0.5 m^3 and 1.2 m^3. The largest chamber reportedly used by Kang et al. (2014) was 2 m^3, however, we found it was impossible to maintain a ground seal with a chamber of this size in anything other than zero wind (0 m s^{-1}) wind conditions. In each case, the chamber was placed over a point source 20 cm above the ground emitting gas at approximately 40, 100 and 200 g CH_4 hr^{-1}. During the experiment, four sample of 25 ml of air were drawn from the chamber using a 50 ml gas syringe at equal time intervals. The air samples were injected into glass vials containing 30 ml of nitrogen and then stored in a fridge before the CH_4 concentrations were measured using the GC-FID. All samples were measured within two weeks of collection.

The minimum time between air sampling was set at one minute to ensure that the correct vial could be found and the sample outlet purged of gas. When sampling times were less, the experiment became too rushed and errors occurred. Additionally, as a health and safety precaution, a handheld CH_4 sensor, HXG-2D (Sensit Technologies, USA, www.gasleaksensors.com), was placed in the chamber and if the CH_4 concentration exceeded the lower explosive limit before three samples were taken the test was abandoned.
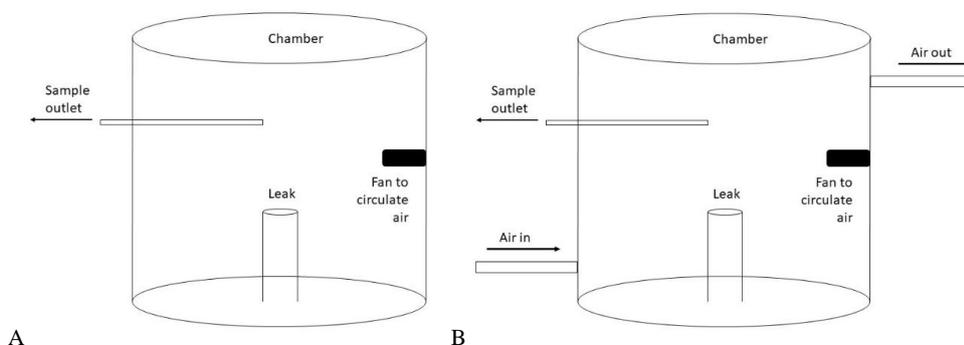
Figure 1 Schematics of the A. Static chamber and B. Dynamic flux chamber.

145 **2.2 Dynamic Chamber**

A single chamber 0.12 m^3 was used for testing the dynamic chamber method. The plastic chamber, open at one end, was placed over known leaks of approximately 40, 100 and 200 g CH$_4$ hr^{-1} and air was passed through the chamber at a constant rate of 67 l min^{-1}, following the method of Riddick et al. (2019). The chamber was left until the CH$_4$ concentration inside had become constant, as measured by the HXG-2D sensor. When steady state was reached, three sample of 25 ml of air were drawn from the chamber

150 using a 50 ml gas syringe injected into glass vials containing 30 ml of nitrogen. As with the samples from the static chamber, the vials were stored in a fridge and measured withing two weeks of collection. Following the methods of Aneja et al. (2006) and Riddick et al. (2019; 2020), the emission is calculated from the steady state gas concentration using Equation 2.

2.3 Hi-Flow

As the Hi-Flow sampler method is relatively simple, no data is required other than the direct measurements made by the instrument.

155 Following the methods of Pekney et al. (2018), the hose end of the Bacharach Hi-Flow sampler (Heath Consultants Inc., www.heathus.com) was placed over the point source and the instrument was turned on. This was repeated three times and the average emission calculated. The Hi-Flow sampler used in this study was calibrated monthly as recommended by the manufacturer. This experiment is repeated three times.

2.4 Gaussian Plume

160 The GP model uses downwind measurement coupled with meteorology to estimate the emission rate of a source using equation 4. Explicitly, the data used are wind speed (u, m s^{-1}), wind direction (WD, °), temperature (T, °C), CH$_4$ concentration downwind of the source (X, µg m^{-3}), location and height of the CH$_4$ detector, background CH$_4$ concentration (X_b, µg m^{-3}) and the Pasquill-Gifford stability class (PGSC).

Methane emissions are calculated using CH$_4$ concentrations measured 5 m downwind and background CH$_4$ concentrations 5 m

165 upwind of the source by the Picarro GasScouter. Wind speed and wind direction were measured every 10 s using a Kestrel 5500 weather meter (www.kestrelmeters.com) on a mast 2 m above the ground. To reduce any impact of mechanical turbulence while maintaining real changes to CH$_4$ emission caused by changing environmental or atmospheric factors, both CH$_4$ concentrations and meteorological data are averaged over 15 min (Laubach et al., 2008; Flesch et al., 2009). The PGSC was calculated from the meteorological data using the method of Seinfeld and Pandis (2006). The lookup table, Table S1, is presented in Supplementary

170 Material Section 1.

Atmospheric
Measurement
Techniques
Discussions

2.5 bLS dispersion model point measurements

WindTrax (www.thunderbeachscientific.com), a commercial software program, uses a bLS dispersion model to calculate the rate of gas emission from a point, area or line source. In this application, the inversion function of the WindTrax inverse dispersion model version 2.0 was used (Flesch et al., 1995). Data used as input are wind speed (u, m s^{-1}), wind direction (WD, °), temperature (T, °C), downwind CH$_4$ concentration (X, µg m^{-3}), location and height of the CH$_4$ detector, background CH$_4$ concentration (X_b, µg m^{-3}), the roughness length (z_0, m) and the Pasquill-Gifford stability class. The ideal terrain for WindTrax modelling is an obstruction-free surface (Sommer et al., 2005; Laubach et al., 2008) with the maximum distance between the source and the detector of 1 km (Flesch et al., 2005, 2009). Data for downwind average CH$_4$ concentration, background CH$_4$ concentration, meteorological and micrometeorological data used in WindTrax will be the same as described in Section 2.4.

2.6 Measures of accuracy and precision

To gain a better understanding of method accuracy and precision, experiments described in sections 2.1 to 2.5 are repeated twice more. In each individual experiment the difference between the known emission rate and the calculated emission rate will be presented as a percentage (Equation 6), where A is the accuracy, Q_c is the calculated emission and Q_k is the known emission. The average accuracy of the three experiments (A_r, %) will be presented as a measure of the accuracy and the standard deviation ($A_{S.D.}$) of the individual uncertainties will be used as a comparative measure of the precision.

$$A = \frac{(Q_c - Q_k)}{Q_k} x\ 100 \qquad\qquad \text{(Equation 6)}$$

3 Results

3.1 Method narrative – anecdotal description of methods

Of the methods tested in this study (summary in Table 1), the static chamber is logistically the simplest. The researcher fixes a container around an emission source and extracts air samples at known time intervals. These vials can be stored for up to a month before analysis on a gas chromatograph. As such, the samples can be analyzed by a third party and the researcher only requires access to the flux chamber, LEL sensor, and the required gas sampling equipment. We found the main shortcomings of the static chamber method are: 1. It was difficult to take samples fast enough during the linear change in concentration; and 2. The method is inherently dangerous. We were aware throughout the experiment that the chamber will become explosive and pre-calculated the time between sample measurement based on the emission rate. During the 200 g CH$_4$ h^{-1} experiment, the lower explosive limit of CH$_4$ was reached after three minutes of the chamber being sealed. We would like to highlight that collecting unprocessed natural gas that may contain aromatic hydrocarbons and hydrogen sulphide is ill-advised and we do not recommend its use. As such, we have not presented our measurement data and strongly encourage the use of an alternative method.

Atmospheric
Measurement
Techniques
Open Access
Discussions
EGU

200 **Table 1 Condensed description of logistical needs and results of each experiment.** *Access* **describes if physical access to the emission source is required,** *Inst* **describes if a dedicated instrument is required, and cost is the approximate cost of the lowest price instrument capable of the measurements.** *Met* **describes if meteorological data is required.** T_{meas} **and** $T_{analysis}$ **are the times it takes to conduct and analyse one measurement, respectively.** *A* **is the accuracy of one measurement of a 200 g CH₄ h⁻¹ source (as defined above in Section 2.6),** A_r **is the average accuracy when repeating the measurement of a 200 g CH₄ h⁻¹ source three times,** $A_{S.D.}$ **is the standard deviation of the accuracy of the three repeated experiments and** *U* **is the** 205 **theoretical uncertainty as presented in previous studies.**

Method	Access	Inst	Cost (k$)	Met	T_{meas} (mins)	$T_{analysis}$ (mins)	A (%)	A_r (%)	$A_{S.D.}$	U (%)
Static chamber	Y	N	-	N						+100, - 50*
Dynamic chamber	Y	N	-	N	15	5	-11	-10	5.9	± 7#
HiFlow	Y	Y	5	N	5	-	-16	-18	8.2	± 10†
Gaussian Plume	N	Y	32	Y	15	60	33	29	12.5	± 18‡
bLs model	N	Y	32	Y	15	90	-11	-7	14.1	± 12§

* Kang et al. (2014), # Riddick et al. (2019), † Pekney et al. (2015), ‡ Riddick et al. (2020), § Riddick et al. (2016)

The dynamic chamber is logistically one step more advanced than the static chamber and requires a pump to draw air through the chamber at a known rate, and, ideally, a flow meter to measure the air flow. This reduces the potential for CH_4 concentration inside the chamber becoming explosive. This means the main advantages of the static chamber are preserved, i.e. cost and ease of 210 analysis, but mitigates the health and safety concerns. Again, the major shortcoming of the dynamic chamber method is that it requires direct access to the emission source.

The HiFlow is an off-the-shelf method/instrument, and as an integrated solution, is easier than the dynamic chamber. Once calibrated, the HiFlow bag is loosely cinched around the emission source and turned on. The instrument displays the methane emission, in l min⁻¹, within a minute at a precision of one significant figure. The data are stored in the instrument and can be 215 downloaded later. The advantages of the HiFlow are the ease of use and amount of time needed to measure a source, typically five minutes per emission source. The main shortcomings are that the researcher needs to have a HiFlow instrument, direct access to the source, calibration gas, and a means of charging batteries and/or powering the instrument.

Measurement data required for the GP and bLs methods were the same. After CH_4 is emitted from a source it quickly disperses and to measure the concentration downwind access to a sub-ppm CH_4 analyzer is required. In 2020, the least-expensive, suitable 220 instrument on the market costs around $32,000. In addition to near-ambient CH_4 concentration measurements, meteorological data are required to populate the models. Despite the cost and time required to make the measurements, the practical advantages of these methods are that access is not required, emissions can be calculated from remote sources and that conditions are generally safe. However, ensuring that the measurement location is in the plume for long enough to detect an enhancement large enough for the instrument to measure accurately can be challenging. In light winds the plume can move laterally and the sensor becomes 225 offset.

3.2 Accuracy of single measurements

To simulate typical measurement methods and gain an understanding of how good a 'snapshot' measurement can be, a single measurement was taken from a source of known emission rates: ~40, 100, 200 g CH₄ h⁻¹. The emissions were generated for each source and the percentage difference between measured and known emissions also calculated (*A*, %). Emissions calculated using

230 the dynamic chamber and HiFlow methods were -11 % and -16 %, respectively (Figure 2A and Supplementary Material Section 2 Table S2). From these measurements, using single measurement data in a Gaussian Plume model generally results in an overestimate of CH_4 emission, between 33 and 104% higher than the actual emission rate. However, $|A|$ decreases as the emission rate increases; $A = -33\%$ for emissionrates of 200 g CH_4 h^{-1}. For the other three methods, dynamic chamber, HiFlow and bLs, the emission generated are an underestimate, lying within 21% of the known emission rate, and not changing with increased
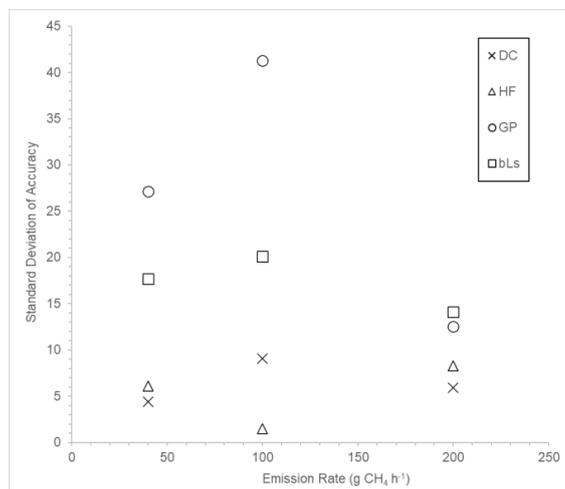
235 emission.
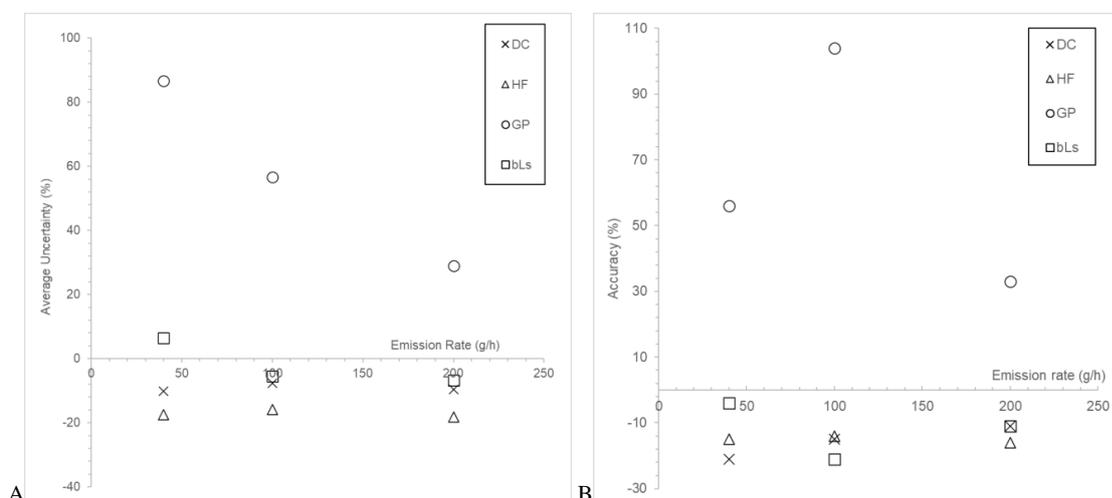
Figure 2 A) Accuracy (% difference from known emission rate) of emission estimates from a single measurement using each of the measurement methodologies at different known emission rates (~ 40, 100 and 200 g CH_4 h^{-1}). B) Accuracy of emission estimates from the average of three repeated measurements. Abbreviations as follows: DC – Dynamic chamber,
240 **HF – HiFlow, GP – Gaussian Plume, bLs – Backwards Lagrangian stochastic method.**

3.2 Accuracy and precision of repeat measurements

Repeating the experiments three times and generating an average value A_r, generally became closer to zero as the flow rate of the source increased. Our results show that the most accurate method for generating emissions after repeat measurements from a 200 g CH_4 h^{-1} source was the bLs method (-7%), then the dynamic chamber (-10%) and then the HiFlow (-18%) (Table 1). The least
245 accurate method after repeat measurements was the GP model (29%). For most methods, the dynamic chamber, GP and bLs, the accuracy of the emission estimate improved, i.e. $|A_r| < |A|$, while for the HiFlow the accuracy became slightly worse, i.e. $|A_r| > |A|$. Repeating the experiments improved the accuracy of the emission estimate by 4% for the GP model. Data are all presented in Supplementary Material Section 3.

For the 40 g CH_4 h^{-1} source, repeating the experiments generally improved the accuracy of the emission estimate except for the GP
250 model which became 20% less accurate (Figure 2B). Like the accuracy, the precision of the methods became better, i.e. the S.D. of the individual uncertainties became smaller, as the emission rate of the source increased (Figure 3). Methods that made measurements while being attached to the source – chamber and HiFlow methods – were more precise than those that measured remotely – bLs and GP methods.

255   **Figure 3 The standard deviation of the uncertainties of repeated measurements against the emission rate of the experiment.**

4 Discussion and conclusions

4.1 Precision and accuracy

This study investigates the accuracy and precision of five methods that have recently been used to estimate smaller, < 200 g CH_4 h^{-1}, CH_4 emissions from oil and gas infrastructure and include, dynamic chamber, the Bacharach HiFlow sensor, Gaussian plume

260   modelling and backward Lagrangian stochastic models. Here, we generate CH_4 emission estimates from a known CH_4 source emitting approximately 40, 100 and 200 g CH_4 h^{-1}. Experiments simulating published methods are carried out once to generate a single visit estimate and are then repeated twice more to better understand how repeat experiments can improve the accuracy and precision of the emission estimate.

Even though the dynamic chamber repeatedly underestimates the emission, it is the most accurate for a single measurement, which

265   improves with subsequent measurements ($A = -11\%$, $A_r = -10\%$). The HiFlow also underestimates the single emission and becomes less accurate with repetition ($A = -16\%$, $A_r = -18\%$). We suggest the dynamic chamber and HiFlow may be more accurate as the flow of air results in increased mixing within the chamber and the sampled air is a better representation of a homogenous gas mixture.

For the far field methods, the bLs method underestimated emissions both for single and repeat measurements ($A = -11\%$, $A_r = -7\%$) while the GP method significantly overestimated the emissions ($A = 33\%$, $A_r = 29\%$) despite using the same meteorological

270   and concentration data as input. These findings are consistent with another study (Bonifacio et al., 2013), however, this is the first study that has compared both to a known emission rate.

4.2 Which method to use – a decision making paradigm

Regardless of accuracy and precision, this study shows that all methods can be used to estimate emissions from a source less than

275   200 g CH_4 h^{-1} to an accuracy of at least 40%. It is reasonable to assume that this level of uncertainty is acceptable in some studies where the research is only aiming to determine relative sizes of emission, e.g. Riddick et al. (2019), while other studies require time-resolved emission estimates to compare against modelled output, e.g. Riddick et al. 2017.

It is, however, concerning that many of the methods show a bias in measurement results. In most studies, it is assumed that averaging large numbers of measurements reduces aggregate error *in the mean* for the measured locations – for example, for a

280   study that surveyed with a downwind method at many well pads, it is generally assumed that, while reported results may be inaccurate for any one facility, the aggregate emission rate across all facilities is an unbiased estimate of total emissions. For the downwind methods and conditions in this study, this is not a correct assumption for most of the methods, and, in particular, will result in substantial error for studies relying on the GP method.

It is also important to note that the study performed here did not simulate or account for issues which increase error in field

285   conditions. For example, when using downwind methods (GP or bLs), the scientist may not know the exact location of the emission point and may be further downwind of the emission location. These knowledge errors may result in uncertainties, or bias in excess of what is presented here; our study should be viewed as a best case bound on the accuracy of the methods.

To provide our findings with some form of context we present our data as a flow diagram to help researchers decide on a suitable method based on accuracy, precision and effort (Figure 4). Here, we give suggestions of which methods could be used given

290   practical and logistical circumstances, i.e. site access and instrument availability. If precision is important in an experiment, our flow diagram would direct a researcher to either the dynamic chamber or the bLs model, while time limitation would suggest the HiFlow. We also hope that the findings of our study could help industry better understand the veracity of researcher's findings and to help put findings in context. As such, we present data that can be used quickly to provide meaning, and balance, to studies that present unexpected findings.

295
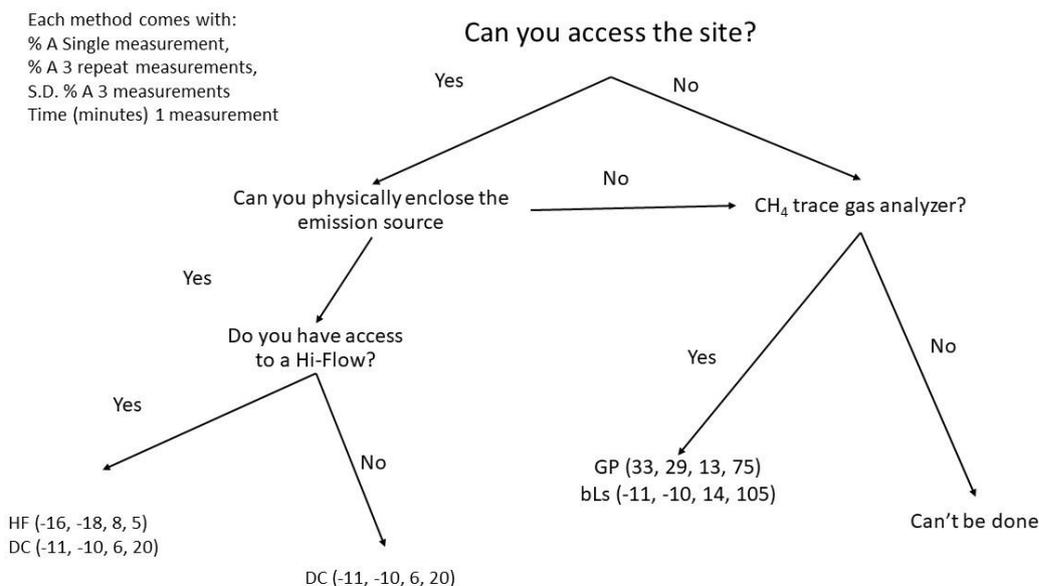
Figure 4 Data presented as a flow diagram to help researchers decide on a suitable method based on accuracy, precision and effort

Author Contributions

300   *Stuart N. Riddick:* Conceptualization, Investigation, Methodology, Supervision, Writing – original draft preparation, review and editing

Riley Ancona: Investigation, Data Curation and Analysis, Writing: original draft preparation

Clay Bell: Writing – Analysis, review and editing

Aidan Duggan: Investigation

305   *Tim Vaughn*: Investigation, Methodology

Kristine Bennett: Writing – Analysis, review and editing

Dan Zimmerle: Writing – Analysis, review and editing

Stuart N. Riddick: Conceptualization, Investigation, Methodology, Supervision, Writing – original draft preparation, review and editing

Competing Interests

The authors declare that they have no conflict of interest.

Disclosure Statement

315   The authors declare that no financial interest or benefit that has arisen from the direct applications of this research.

References

Allen, D. T., Torres, V. M., Thomas, J., Sullivan, D. W., Harrison, M., Hendler, A., Herndon, S. C., Kolb, C. E., Fraser, M. P., Hill, A. D., Lamb, B. K., Miskimins, J., Sawyer, R. F., and Seinfeld, J. H.: Measurements of methane emissions at natural gas production sites in the United States, Proceedings of the National Academy of Sciences, 110, 17768–17773,

320   https://doi.org/10.1073/pnas.1304880110, 2013.

Alvarez, R. A., Zavala-Araiza, D., Lyon, D. R., Allen, D. T., Barkley, Z. R., Brandt, A. R., Davis, K. J., Herndon, S. C., Jacob, D. J., Karion, A., Kort, E. A., Lamb, B. K., Lauvaux, T., Maasakkers, J. D., Marchese, A. J., Omara, M., Pacala, S. W., Peischl, J., Robinson, A. L., Shepson, P. B., Sweeney, C., Townsend-Small, A., Wofsy, S. C., and Hamburg, S. P.: Assessment of methane emissions from the U.S. oil and gas supply chain, eaar7204, https://doi.org/10.1126/science.aar7204, 2018.

325   Aneja, V. P., Blunden, J., Claiborn, C. S., and Rogers, H. H.: Dynamic Chamber System to Measure Gaseous Compounds Emissions and Atmospheric-Biospheric Interactions, in: Environmental Simulation Chambers: Application to Atmospheric Chemical Processes, 97–109, 2006.

Baillie, J., Risk, D., Atherton, E., O'Connell, E., Fougère, C., Bourlon, E., and MacKay, K.: Methane emissions from conventional and unconventional oil and gas production sites in southeastern Saskatchewan, Canada, Environ. Res. Commun., 1,
330   011003, https://doi.org/10.1088/2515-7620/ab01f2, 2019.

Bell, C. S., Vaughn, T. L., Zimmerle, D., Herndon, S. C., Yacovitch, T. I., Heath, G. A., Pétron, G., Edie, R., Field, R. A., Murphy, S. M., Robertson, A. M., and Soltis, J.: Comparison of methane emission estimates from multiple measurement techniques at natural gas production pads, 5, 79, https://doi.org/10.1525/elementa.266, 2017.

Bonifacio, H. F., Maghirang, R. G., Razote, E. B., Trabue, S. L., and Prueger, J. H.: Comparison of AERMOD and WindTrax
335   dispersion models in determining PM 10 emission rates from a beef cattle feedlot, Journal of the Air & Waste Management Association, 63, 545–556, https://doi.org/10.1080/10962247.2013.768311, 2013.

Brantley, H. L., Thoma, E. D., and Eisele, A. P.: Assessment of volatile organic compound and hazardous air pollutant emissions from oil and natural gas well pads using mobile remote and on-site direct measurements, Journal of the Air & Waste Management Association, 65, 1072–1082, https://doi.org/10.1080/10962247.2015.1056888, 2015.

340   Caulton, D. R., Shepson, P. B., Santoro, R. L., Sparks, J. P., Howarth, R. W., Ingraffea, A. R., Cambaliza, M. O. L., Sweeney, C., Karion, A., Davis, K. J., Stirm, B. H., Montzka, S. A., and Miller, B. R.: Toward a better understanding and quantification of methane emissions from shale gas development, 111, 6237–6242, https://doi.org/10.1073/pnas.1316546111, 2014.

Collier, S. M., Ruark, M. D., Oates, L. G., Jokela, W. E., and Dell, C. J.: Measurement of Greenhouse Gas Flux from Agricultural Soils Using Static Chambers, JoVE, 52110, https://doi.org/10.3791/52110, 2014.

345   Connolly, J. I., Robinson, R. A., and Gardiner, T. D.: Assessment of the Bacharach Hi Flow® Sampler characteristics and potential failure modes when measuring methane emissions, Measurement, 145, 226–233, https://doi.org/10.1016/j.measurement.2019.05.055, 2019.

Denmead, O. T.: Approaches to measuring fluxes of methane and nitrous oxide between landscapes and the atmosphere, Plant Soil, 309, 5–24, https://doi.org/10.1007/s11104-008-9599-z, 2008.

350   Edie, R., Robertson, A. M., Field, R. A., Soltis, J., Snare, D. A., Zimmerle, D., Bell, C. S., Vaughn, T. L., and Murphy, S. M.: Constraining the accuracy of flux estimates using OTM 33A, Atmos. Meas. Tech., 13, 341–353, https://doi.org/10.5194/amt-13-341-2020, 2020.

Flesch, T. K., Wilson, J. D., and Yee, E.: Backward-Time Lagrangian Stochastic Dispersion Models and Their Application to Estimate Gaseous Emissions, J. Appl. Meteor., 34, 1320–1332, https://doi.org/10.1175/1520-
355   0450(1995)034<1320:BTLSDM>2.0.CO;2, 1995.

Kang, M., Kanno, C. M., Reid, M. C., Zhang, X., Mauzerall, D. L., Celia, M. A., Chen, Y., and Onstott, T. C.: Direct measurements of methane emissions from abandoned oil and gas wells in Pennsylvania, 111, 18173–18177, https://doi.org/10.1073/pnas.1408315111, 2014.

Livingston, G. P. and Hutchinson, G. L.: Enclosure-based measurement of trace gas exchange: applications and sources of error.,
360   in: In: Matson, P.A. and Harris, R.C., Eds., Biogenic trace gases: measuring emissions from soil and water., Blackwell Science Ltd., Oxford, UK, 14–51, 1995.

Omara, M., Sullivan, M. R., Li, X., Subramanian, R., Robinson, A. L., and Presto, A. A.: Methane Emissions from Conventional and Unconventional Natural Gas Production Sites in the Marcellus Shale Basin, 50, 2099–2107, https://doi.org/10.1021/acs.est.5b05503, 2016.

365   Pekney, N. J., Diehl, J. R., Ruehl, D., Sams, J., Veloski, G., Patel, A., Schmidt, C., and Card, T.: Measurement of methane emissions from abandoned oil and gas wells in Hillman State Park, Pennsylvania, Carbon Management, 9, 165–175, https://doi.org/10.1080/17583004.2018.1443642, 2018.

Pihlatie, M. K., Christiansen, J. R., Aaltonen, H., Korhonen, J. F. J., Nordbo, A., Rasilo, T., Benanti, G., Giebels, M., Helmy, M.,

Sheehy, J., Jones, S., Juszczak, R., Klefoth, R., Lobo-do-Vale, R., Rosa, A. P., Schreiber, P., Serça, D., Vicca, S., Wolf, B., and

370   Pumpanen, J.: Comparison of static chambers to measure CH4 emissions from soils, Agricultural and Forest Meteorology, 171–

172, 124–136, https://doi.org/10.1016/j.agrformet.2012.11.008, 2013.

Riddick, S. N., Connors, S., Robinson, A. D., Manning, A. J., Jones, P. S. D., Lowry, D., Nisbet, E., Skelton, R. L., Allen, G.,

Pitt, J., and Harris, N. R. P.: Estimating the size of a methane emission point source at different scales: from local to landscape,

17, 7839–7851, https://doi.org/10.5194/acp-17-7839-2017, 2017.

375   Riddick, S. N., Mauzerall, D. L., Celia, M. A., Kang, M., Bressler, K., Chu, C., and Gum, C. D.: Measuring methane emissions

from abandoned and active oil and gas wells in West Virginia, 651, 1849–1856, https://doi.org/10.1016/j.scitotenv.2018.10.082,

2019a.

Riddick, S. N., Mauzerall, D. L., Celia, M., Harris, N. R. P., Allen, G., Pitt, J., Staunton-Sykes, J., Forster, G. L., Kang, M.,

Lowry, D., Nisbet, E. G., and Manning, A. J.: Methane emissions from oil and gas platforms in the North Sea, 19, 9787–9796,

380   https://doi.org/10.5194/acp-19-9787-2019, 2019b.

Riddick, S. N., Mauzerall, D. L., Celia, M. A., Kang, M., and Bandilla, K.: Variability observed over time in methane emissions

from abandoned oil and gas wells, International Journal of Greenhouse Gas Control, 100, 103116,

https://doi.org/10.1016/j.ijggc.2020.103116, 2020.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics: from air pollution to climate change, Third edition., John

385   Wiley & Sons, Inc, Hoboken, New Jersey, 1120 pp., 2016.

Vaughn, T. L., Bell, C. S., Pickering, C. K., Schwietzke, S., Heath, G. A., Pétron, G., Zimmerle, D. J., Schnell, R. C., and

Nummedal, D.: Temporal variability largely explains top-down/bottom-up difference in methane emission estimates from a

natural gas production region, 115, 11712–11717, https://doi.org/10.1073/pnas.1805687115, 2018.

Zavala-Araiza, D., Omara, M., Gautam, R., Smith, M. L., Pandey, S., Aben, I., Almanza-Veloz, V., Conley, S., Houweling, S.,

390   Kort, E. A., Maasakkers, J. D., Molina, L. T., Pusuluri, A., Scarpelli, T., Schwietzke, S., Shen, L., Zavala, M., and Hamburg, S.

P.: A tale of two regions: methane emissions from oil and gas production in offshore/onshore Mexico, Environ. Res. Lett., 16,

024019, https://doi.org/10.1088/1748-9326/abceeb, 2021.