

Reviewer 1

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

The Powerhouse Energy Campus
Colorado State University
430 North College Avenue
Fort Collins, CO 80524

E-mail: Stuart.Riddick@colostate.edu

26th April 2022

Dear Dr Chen,

We thank reviewer 1 for their comments. As suggested, we have amended the manuscript to address the reviewers' comments and have indicated changes to the manuscript in red text.

Please find our detailed responses below.

Yours sincerely,

Stuart Riddick

Reviewer 1 General comment 1:

One concern with this paper is that it's unclear who this work is applicable to. For example, for greenhouse gas emission inventories, there is always the emission factor approach, in contrast to what's shown in Figure 4. The authors mention oil and gas wells but without providing much context into the type of oil and gas wells, which can explain how the flow rates studied were chosen.

Response to reviewer:

The aim of this work to produce the first study to compare emissions quantified by a range of methods against known methane release rates. The target audience is anyone who has thought of, or who has used one of the methods included. As mentioned in the original introduction, these methods were chosen as they have been used extensively without a clear understanding of the precision and accuracy of quantification. Furthermore, as the methods have been used to guide policy decisions, it is important to regulators to understand how well these methods perform under controlled conditions. Therefore, we suggest this work is of interest to those currently quantifying methane emissions, those that plan to quantify methane emission and regulators that have justify decisions on policy. We also add the caveat that we only endorse measurement methodologies can be done safely wearing PPE as regulated at METEC (steel toe boot, FR overalls, hard hat, safety glasses and 4 gas monitor).

Changes to the manuscript:

In response to the reviewer's comment we have edited the opening paragraph to make the target audience more obvious.

At L27:

“Methane (CH₄) gas is a powerful greenhouse gas with a greenhouse warming potential 84 times larger than carbon dioxide over 100 years. Quantification of CH₄ emissions from abandoned wells has recently become an area of interest as studies suggest over 200 Gg CH₄ yr⁻¹ is emitted from 2.2 million abandoned wells in the US alone (US EPA, 2021). Quantifying and then plugging these wells makes them an attractive target for achieving goals set out in the Paris Agreement (Nisbet et al., 2020). Additionally, private companies are beginning initiatives to generate revenue through carbon credits gained by plugging wells and accurate quantification is essential for realizing the capital.

As there are millions of abandoned wells globally, there is a growing need to measure as many wells as quickly as possible to identify the most emissive wells. Typically, an emission from an abandoned well can be considered as an above-ground point source that is relatively small in emission size, up to 180 g CH₄ hour⁻¹ (Riddick et al., 2019a; Pekney et al., 2018; Townsend-Small et al., 2016; Boothroyd et al., 2016). Other emission sources, such as emissions pipeline leakage, are fundamentally different in behavior, where gas travels through the soil and forms an area emission at the surface, these sources require different methods for estimating the emission, e.g. mass balance or eddy covariance. Area emissions could form if a plugged well leaks from corrosion of the borehole casing, but this will not be discussed in this study.

Several methods are being used to measure emissions from these smaller point sources, i.e. less than 180 g CH₄ hour⁻¹. The chosen measurement approach depends on how close an observer can get to the source, instrumentation availability and the meteorological/micrometeorological conditions at the measurement site. Measurement methods can be classed as direct, i.e. touching/enclosing the source, and downwind measurements where access is not possible. Direct methods include static chambers (Livingston and Hutchinson, 1995), dynamic flux chambers (Riddick et al., 2019a, 2020b; Aneja et al., 2006) and Hi Flow sampling (Pekney et al., 2018; Allen et al., 2013; Brantley et al., 2015). While downwind methods include Gaussian-based plume models (Baillie et al., 2019; Caulton et al., 2014; Riddick et al., 2019b, 2020a; Edie et al., 2020; Bell et al., 2017) and Lagrangian dispersion models (Riddick et al., 2019b, 2017; Denmead, 2008; Flesch et al., 1995). Emissions calculated using the majority of these methods have not been comprehensively compared using controlled emission source rates.

Other quantification methods are generally unsuitable for measuring emissions from abandoned wells. Infra-red cameras, such as FLIR cameras, cannot be used to quantify emission and have difficulty detecting plume smaller emissions (Zimmerle et al., 2020). Mass balance approaches are unlikely to detect the small and narrow plume from the abandoned well. Tracer release is technically demanding, takes a long time to make a single measurement and requires road access for measurement. Remote sensing has typical detection limits of 10+ kg CH₄ h⁻¹ for aircraft (Duren et al., 2019), 100+ kg CH₄ h⁻¹ for satellites (Cooper et al., 2022) and unsuitable for these types of emission source. As such, these other quantification methods will not be investigated in this study.

In general, as access becomes more restricted, emission rates larger, or safety concerns increase (such as the co-emission of harmful gases), the method used to estimate the CH₄ emission rate of a source must be carefully considered. From experience and the response of a 4-gas monitor, working close enough to measure emissions greater than 200 g CH₄ h⁻¹ for many of these methods (especially the chambers and Hi Flow) can be unsafe, therefore this study is limited to quantifying CH₄ emissions between the lowest flow METEC can produce (40 g CH₄ h⁻¹) and the highest flow we feel comfortable measuring with these methods (200 g CH₄ h⁻¹). Putting these emission ranges into real-world context, the maximum emission from unplugged and abandoned wells was measured at 177 g CH₄ h⁻¹ in West Virginia (Riddick et al., 2019a), 175 g CH₄ h⁻¹ in Pennsylvania (Pekney et al., 2018), 146 g CH₄ h⁻¹ across the US (Townsend-Small et al., 2016) and 35 g CH₄ h⁻¹ in the UK (Boothroyd et al., 2016). As most of the methods presented here require access to the source, we considered 200 g CH₄ h⁻¹ to be a sensible limit to the emission rate and is larger than the emissions observed by many previous studies. Therefore, the scope of this study is limited to estimating CH₄ emissions from a single point source that we would realistically be able to approach and measure, i.e. between 40 and 200 g CH₄ h⁻¹.”

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 General comment 2:

Also, the authors caution against doing measurements at sites with hydrogen sulfide and aromatic hydrocarbons, which would bias samples for regional or national inventories. From a policy perspective, wells that emit hydrogen sulfide or aromatic hydrocarbons are often prioritized for mitigation, and it would be unfortunate if we can't quantify the methane emissions being reduced through these efforts.

Response to reviewer:

We do not caution against measurement of H₂S and aromatic hydrocarbons, but caution against using a static chamber to measure methane emissions. The stated aim of this study is to present controlled release data for methods that can be used to measure methane emissions from a point source between 40 and 200 g CH₄ h⁻¹. This paper only refers to the measurement of methane and does not comment on the measurement of other gases.

Change to the manuscript

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 General comment 3:

The methods can give very different uncertainties depending on how the experiment is conducted. For example, there are many ways to implement the Gaussian Plume method, including how and where methane is analyzed and the micrometeorology measured. The same goes for the static and dynamic chambers. Overall, the authors need to add more detail on the methodology and experimental conditions (dates/times, exact equipment and supplies, etc.).

Response to reviewer:

We have added text to the manuscript (as suggested by the reviewer below) and dates/times of the experiments to the SI to clarify some of the details of the methodologies. However, the assertion that method uncertainty changes over time is conjecture. This may be true about downwind measurements, where atmospheric stability makes calculating emissions more complex resulting in propagated uncertainty, but chamber measurement uncertainties are unlikely to change over time. To address these concerns we have added text to the discussion.

Change to the manuscript

At L317:

“Both the dynamic chamber ($A_r = -10\%$, -8% , -10% at emission rates of 40, 100 and 200 g CH₄ h⁻¹, respectively) and Hi Flow ($A_r = -18\%$, -16% , -18%) repeatedly underestimate the emission, but the dynamic chamber is the most accurate for measurement. For the far field methods, the bLs method underestimated emissions ($A_r = +6\%$, -6% , -7%) while the GP method significantly overestimated the emissions ($A_r = +86\%$, $+57\%$, $+29\%$) despite using the same meteorological 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. In all cases the accuracy in the emission estimate increased with emission rate apart from the Hi Flow. The Bacharach Hi Flow system is designed to measure emission from 50 g CH₄ h⁻¹ to 9 kg CH₄ h⁻¹ to an accuracy of $\pm 10\%$. All flow rates presented here are at the lowest range that the Hi Flow can measure and it is likely that the uncertainty in the systems sensors that measures between 40 and 400 g CH₄ h⁻¹ is of negligible difference.

The method that improves the most as the emission rate increases is the GP method, where accuracy increases from $+87\%$ to $+29\%$ as the emission rate increased from 40 to 200 g CH₄ h⁻¹. This improvement in emission is likely caused by the increased size of the plume and the ability of GP model to parameterize the concentration at distances from the centerline of the plume. Although not explicitly stated, the parameterization of the lateral dispersion in the GP model is the same at 100 m as at 5 m which is unlikely. Other controlled release experiments using the GP approach show similar uncertainties, one experiment reported average emissions calculated using a GP model less than 20% (release rates were not published), with the uncertainty mainly driven by atmospheric variability (Caulton et al., 2019). Another showed uncertainties of $\pm 50\%$ for triplicate measurements of emissions between 90 and 970 g CH₄ h⁻¹ (Caulton et al., 2018).

Data do not exist on controlled release experiments using a dynamic chamber. One study suggested a theoretical emissions uncertainty in the dynamic chamber approach of $\pm 7\%$ (Riddick et al., 2019a), with the largest source of uncertainty caused by the measurement of the flow rate of air through the chamber. Other sources of uncertainty for the dynamic chamber methods are relatively negligible as the methane quantification of the background gas and the gas at steady state (assuming complete mixing of the gas in the chamber) using the GC is highly accurate over a large concentration range and the volume of the chamber fixed by a plastic structure.

A controlled release has been conducted for the bLs model, but only for an emission from an area source (Ro et al., 2011) at the surface and not analogous to the emissions of this study. Ro et al. (2011) estimated the bLs uncertainty at $\pm 25\%$ for a gas emitted at an unspecified rate from a 27 m² emission area. As with the GP approach, the bLs model's main source uncertainty is the parameterization of the atmospheric stability (Riddick et al., 2012; Flesch et al., 1995; Ro et al., 2011). The main advantage of the bLs model over the GP at these short distances is it calculates the lateral dispersion of gas for individual particles, while the GP uses an averaged dispersion parameter.

The emission estimates quantified using direct methods, dynamic chamber and Hi Flow sampler, have a lower S.D. than the far-field methods (Figure 2B). The S.D. of direct measurement methods remain relatively constant for emissions between 40 and 200 g CH₄ h⁻¹ and reflects the relative simplicity of the methods. Assuming all other parameters are measured correctly, for direct methods the variability in emission estimate is a function of how well the CH₄ is mixed into the air in the chamber during the measurement.

Variability in the far field emission estimates is much larger and reflects the relative complexity of inferring emissions. Variability in wind speed, wind direction and atmospheric stability over the 20-minute averaging period are likely to propagate through to large variability in the emission estimate. It may be reasonable to suggest that the variability in bLs calculated emission less than for the GP method because of the added parametrization available (roughness length and gas species). In summary, the penalty of downwind measurement is a higher uncertainty in individual measurements, but this appears to be corrected for by the bLs model through repeat measurements where uncertainty is corrected for by the stochastic nature of particle movement modelling.”

At L363:

“It is, however, concerning that many of the methods show a bias in measurement results and in particular the GP model (Figure 3). In most studies, it is assumed that in taking multiple measurements the average uncertainty will be reduced to an aggregate, unbiased emission estimate. Taking the GP emission estimates as an example, the individual calculated emissions are all overestimates of the true emission, therefore, suggesting a fundamental shortcoming in the method (Figure 3). These measurements were taken four days apart in similar environmental conditions (all PGSC C) with wind direction being the only difference between measurements, which can be seen from the correlation between the uncertainty and horizontal distance from plume center (Figure 3B). As mentioned above, it is likely that this is due to the lateral dispersion in the GP approach being parametrized incorrectly, i.e. using values that were defined for distances of 100 m. This suggests that using the GP approach for distances less than 100 m, it is not correct to assume that repeat measurements will remove bias in the calculated average emission.”

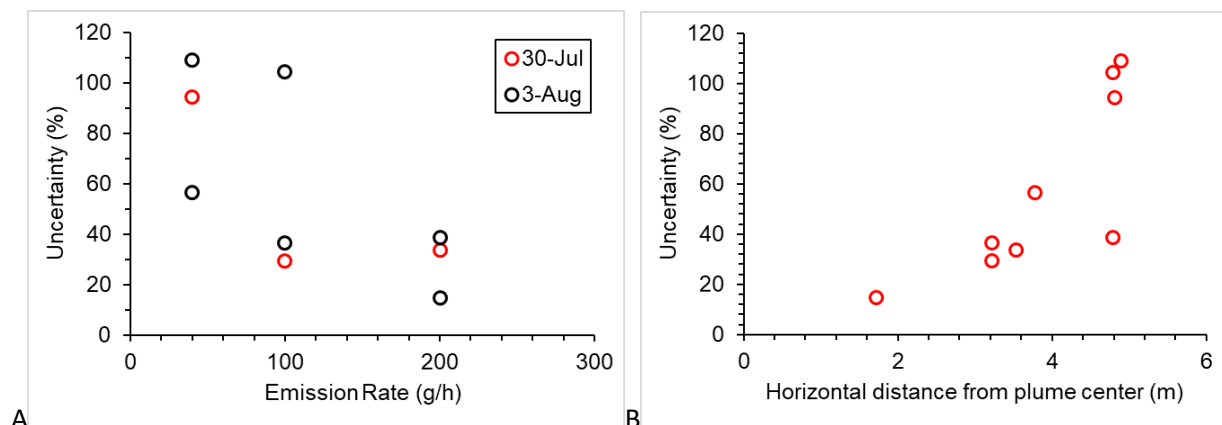


Figure 1 A) Individual uncertainty in Gaussian Plume measurements at 40, 100 and 200 g CH₄ h⁻¹ and B) Individual uncertainties plotted against the horizontal distance from the plume center (m)”

Below are some additional detailed comments:

Reviewer 1 Specific comment 2:

line 2: what is considered a "small point source"?

Response to reviewer:

In general, a point source is considered to be an emission from an aperture. In this case, the hole was a 6mm diameter tube. We have included this detail in the methods section.

Changes to the manuscript:

At L86:

“(diameter 6 mm)”

Reviewer 1 Specific comment 2:

line 7: how is a point source defined? At what scale?

Response to reviewer:

A point source is considered to be an emission from an aperture. In this case, the hole was a 6mm diameter tube. We have included this detail in the methods section.

Changes to the manuscript:

At L86:

“(diameter 6 mm)”

Reviewer 1 Specific comment 3:

line 13: not clear if static chambers are tested in this study.

Response to reviewer:

At L256 of the manuscript we describe one of the shortcomings of the static chamber as being inherently dangerous and “As such, we have not presented our measurement data and strongly encourage the use of an alternative method.”.

Measurements were taken but we chose not to release the data. Instead, we included a statement that we encourage the use of any other method. To clarify this we have added a caveat to the aims of the study.

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 Specific comment 3:

line 16: why only 200 g/h? confusing because in the text, there are three flow rates mentioned.

Response to reviewer:

Corrected as suggested.

Changes to the manuscript:

At L16:

“40 to 200 g CH₄ h⁻¹”

Reviewer 1 Specific comment 4:

line 40: "very small" sounds arbitrary. where does the 0.6 mg/hr come from? Kang et al. (2014) also measured negative emission rates. There also are a wide range of published measurements using this approach for much smaller fluxes found in natural environments.

Response to reviewer:

As suggested, we have edited the static chamber paragraph to remove the emission data.

Changes to the manuscript:

At L93:

“The static chambers method is relatively simple, where a container of a known volume (V , m^3) is placed 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). The static chamber method requires no power, apart from batteries to run a fan in the chamber and is very portable. The major shortcoming of this method is that large emission sources can result in the concentration inside the chamber exceeding the CH_4 lower explosive limit (LEL).”

Reviewer 1 Specific comment 5:

line 41: replace "place" with "placed"

Response to reviewer:

Corrected, as suggested.

Reviewer 1 Specific comment 6:

line 43: The static chamber method does not require a gas chromatograph. In El Hachem and Kang (2022) published in Science of the Total Environment, they do not use a gas chromatograph.

Response to reviewer:

As suggested, we have removed the GC comment and added using other sensors to the discussion.

At L93:

“The static chambers method is relatively simple, where a container of a known volume (V , m^3) is placed 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). The static chamber method requires no power, apart from batteries to run a fan in the chamber and is very portable. The major shortcoming of this method is that large emission sources can result in the concentration inside the chamber exceeding the CH_4 lower explosive limit (LEL).”

At L235

“To address the first shortcoming, a trace gas analyzer could be used to measure the concentrations inside the chamber. As trace gas analyzers use a pump to draw air into the measurement cavity, the analyzer could be arranged in one of two ways. Both introduce additional uncertainty into the quantification. If the gas is removed from the

chamber (i.e. the analyzer outlet is vented outside the chamber), the static chamber becomes a dynamic chamber and the analyzer flow rate must be accounted for in the quantification. If the measured gas is reintroduced to the chamber (i.e. the analyzer outlet is vented back to the chamber), a gas of lower concentration is being continually added to the “closed” system and it is therefore unclear how much uncertainty is caused by this cycling. Furthermore, the linear response of a portable trace gas analyzer, e.g. the ABB GLA131-GGA Greenhouse Gas Analyzer (<https://new.abb.com/>), is 100 ppm. Using the lowest emission rate in the study, $40 \text{ g CH}_4 \text{ h}^{-1}$, and the largest chamber, 0.5 m^3 , the concentration inside the chamber will exceed the linear range within 7 seconds. It is unlikely that gas will mix entirely throughout the chamber in 7 seconds and emission estimates are unlikely to be accurate. Another alternative could be using a lower precision sensor with a larger detection range, such as the SGX INIR- ME100 (<https://sgx.cdistore.com/>) that can measure from 200 ppm to 100% methane bv, but safety issues remain.”

Reviewer 1 Specific comment 7:

line 50: what is the source of this air? Is it background air?

Response to reviewer:

Yes, this is background air that was also sampled using a GC. A description of the sampling of air inside and outside the chamber has been added to the manuscript.

Changes to the manuscript:

At L132:

“When the chamber reached steady state, three air samples were taken from inside the chamber. A background air sample was taken outside the chamber as the chamber approached steady state. The methane concentration in all air samples was measured using a gas chromatography.”

Reviewer 1 Specific comment 8:

line 53: what is the background methane concentrations in the air? And what air is the authors referring to?

Response to reviewer:

A sample of air was taken outside the chamber as the it approached steady state. This is the sample of background air that was measured and used in the calculation.

Changes to the manuscript:

At L132:

“When the chamber reached steady state, three air samples were taken from inside the chamber. A background air sample was taken outside the chamber as the chamber approached steady state. The methane concentration in all air samples was measured using a gas chromatography.”

Reviewer 1 Specific comment 9:

line 56: how much power is required? What type of power source is needed?

Response to reviewer:

In this case, 120 V mains power was used, however, this could be anything supplying a 12 V.

Changes to the manuscript:

At L141:

“As the experiment was conducted at METEC, 120 V mains power was used, however, in a remote location power can be supplied by anything capable of delivering a stable 12 V output.”

Reviewer 1 Specific comment 10:

line 62: what is the current commercial HiFlow sampler? I see in the next lines that you mention the Bacharach. But I've heard that it's been discontinued. Are there others that are currently commercially available? In the previous sentence, the authors write "typical rates are 300 l/min" but that implies there are multiple types of samplers. If there is just one, why not just report the on high flow rate?

Response to reviewer:

This refers to the industry standard Bacharach HiFlow, even though the instrument's production has been discontinued. The Bacharach HiFlow has multiple flow rates and varies throughout the emission quantification, hence the inexact value. To remain instrument agnostic we have removed this information from this section.

Changes to the manuscript:

At L149:

“A Hi Flow sampler draws high volumes of air into a measurement chamber, where the concentration of CH₄ in the air is measured and the emission rate calculated (Equation 3). The Bacharach Hi Flow Sampler (Heath Consultants Inc., www.heathus.com) is the only current industry standard Hi Flow sampler, it draws air at between 226 and 297 l min⁻¹ and can measure CH₄ emissions between 50 g CH₄ h⁻¹ to 9 kg CH₄ h⁻¹ to an accuracy of ± 10% (Connolly et al.,

2019). A recent study commissioned by the California Air Resources Board developed open-source architecture for a new Hi Flow unit which is capable of replacing the current Bacharach Hi Flow Sampler (Vaughn et al., 2022).”

Reviewer 1 Specific comment 11:

line 87: the inputs to the bLS model appears to be the same as the GP model? What are the exact meteorology and micrometeorology parameters needed for the bLS model?

Response to reviewer:

The inputs used here were wind speed, wind direction, Pasquill-Gifford stability class and roughness length, as presented at L216. The roughness length was set at 2.3 cm to represent the short grass of the fetch.

Changes to the manuscript:

At L220:

“The roughness length was set at 2.3 cm to represent the short grass of the fetch. Again, it assumed that the experiments are conducted as close as possible to the source without direct access to the emission point.”

Reviewer 1 Specific comment 12:

line 87: where is this gas concentration taken?

Response to reviewer:

The height of the measurement, 1.5 m, has been included at L191.

Changes to the manuscript:

At L191:

“1.5 m above ground level,”

Reviewer 1 Specific comment 13:

line 94: isn't complex topography and buildings also an issue for the GP model?

Response to reviewer:

This isn't an input to the GP model, therefore not an issue.

Changes to the manuscript:

At L198:

“Complex topography, such as building and trees, are not parameterized or accounted for by the GP model.”

Reviewer 1 Specific comment 14:

line 96: what kind of micrometeorology data is needed?

Response to reviewer:

The inputs used here were Pasquill-Gifford stability class and roughness length, as presented at L220. The roughness length was set at 2.3 cm to represent the short grass of the fetch.

Changes to the manuscript:

At L220:

“The roughness length was set at 2.3 cm to represent the short grass of the fetch.”

Reviewer 1 Specific comment 15:

line 100: for higher emission rates, wouldn't it be easier to do downwind measurements such that site access is less of a concern?

Response to reviewer:

This sentence is misleading and has been reworded.

Changes to manuscript

At L58:

“In general, as access becomes more restricted, emission rates larger, or safety concerns increase (such as the co-emission of harmful gases), the method used to estimate the CH₄ emission rate of a source must be carefully considered.”

Reviewer 1 Specific comment 16:

line 100: what are the safety concerns here? just explosion risk due to high methane concentrations? What about H₂S (See El Hachem and Kang, 2022)?

Response to reviewer:

A significant safety concern for abandoned oil and gas would be the exposure to other co-emitted gases. The measurement of H₂S seems particularly hazardous without adequate PPE, even at low concentrations, 10 ppm, H₂S can cause respiratory failure. However, in this study we focus solely on quantifying methane emissions using typical PPE (FRs and 4 gas monitors) and do not comment on quantifying other gases. It is financially and logistically unrealistic to expect a team measuring many sites (many groups are aiming to measure 100s of sites) to put on self-contained breathing apparatus for each measurement when a safer alternative is available.

Changes to the manuscript:

At L 55:

“In general, as access becomes more restricted, emission rates larger, or safety concerns increase (such as the co-emission of harmful gases), the method used to estimate the CH₄ emission rate of a source must be carefully considered. From experience and the response of a 4-gas monitor, working close enough to measure emissions greater than 200 g CH₄ h⁻¹ for many of these methods (especially the chambers and Hi Flow) can be unsafe, therefore this study is limited to quantifying CH₄ emissions between the lowest flow METEC can produce (40 g CH₄ h⁻¹) and the highest flow we feel comfortable measuring with these methods (200 g CH₄ h⁻¹). Putting these emission ranges into real-world context, the maximum emission from unplugged and abandoned wells was measured at 177 g CH₄ h⁻¹ in West Virginia (Riddick et al., 2019a), 175 g CH₄ h⁻¹ in Pennsylvania (Pekney et al., 2018), 146 g CH₄ h⁻¹ across the US (Townsend-Small et al., 2016) and 35 g CH₄ h⁻¹ in the UK (Boothroyd et al., 2016). As most of the methods presented here require access to the source, we considered 200 g CH₄ h⁻¹ to be a sensible limit to the emission rate and is larger than the emissions observed by many previous studies. Therefore, the scope of this study is limited to estimating CH₄ emissions from a single point source that we would realistically be able to approach and measure, i.e. between 40 and 200 g CH₄ h⁻¹.”

At L71:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 Specific comment 17:

line 102: what is meant by "able to approach"? How close to the single point source in meters?

Response to reviewer:

From experience, it becomes difficult to work effectively in areas near NG emissions greater than 200 g CH₄ h⁻¹. As a rule of thumb at this emission rate, we suggest access to areas closer than 10 m of the source be restricted. As most of the methods require access to the source, we considered this to be a sensible limit to emission.

Changes to the manuscript:

At L55

“In general, as access becomes more restricted, emission rates larger, or safety concerns increase (such as the co-emission of harmful gases), the method used to estimate the CH₄ emission rate of a source must be carefully considered. From experience and the response of a 4-gas monitor, working close enough to measure emissions greater than 200 g CH₄ h⁻¹ for many of these methods (especially the chambers and Hi Flow) can be unsafe, therefore this study is limited to quantifying CH₄ emissions between the lowest flow METEC can produce (40 g CH₄ h⁻¹) and the highest flow we feel comfortable measuring with these methods (200 g CH₄ h⁻¹). Putting these emission ranges into real-world context, the maximum emission from unplugged and abandoned wells was measured at 177 g CH₄ h⁻¹ in West Virginia (Riddick et al., 2019a), 175 g CH₄ h⁻¹ in Pennsylvania (Pekney et al., 2018), 146 g CH₄ h⁻¹ across the US (Townsend-Small et al., 2016) and 35 g CH₄ h⁻¹ in the UK (Boothroyd et al., 2016). As most of the methods presented here require access to the source, we considered 200 g CH₄ h⁻¹ to be a sensible limit to the emission rate and is larger than the emissions observed by many previous studies. Therefore, the scope of this study is limited to estimating CH₄ emissions from a single point source that we would realistically be able to approach and measure, i.e. between 40 and 200 g CH₄ h⁻¹.”

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 Specific comment 18:

line 103: what are the emission rates considered? In the abstract, it was only for 200 g/h but there are three mentioned later. Need to be consistent throughout.

Response to reviewer:

Have amended to bracket the emissions throughout the manuscript.

Reviewer 1 Specific comment 19:

line 108: why the cut off at 200 g/h? There should be a paragraph on the literature for tests at >200 g/h and describe why those studies are not applicable here.

Response to reviewer:

This was a health and safety regulation at METEC. From experience, it becomes difficult to work effectively in areas near NG emissions greater than 200 g CH₄ h⁻¹. As a rule of thumb at this emission rate, we suggest access to areas closer than 10 m of the source be restricted. As most of the methods require access to the source, we considered this to be a sensible limit to emission. This upper limit is also larger than most of the emissions observed from abandoned oil and gas wells and emissions have been presented in the text.

Changes to the manuscript:

At L56:

“From experience and the response of a 4-gas monitor, working close enough to measure emissions greater than 200 g CH₄ h⁻¹ for many of these methods (especially the chambers and Hi Flow) can be unsafe, therefore this study is limited to quantifying CH₄ emissions between the lowest flow METEC can produce (40 g CH₄ h⁻¹) and the highest flow we feel comfortable measuring with these methods (200 g CH₄ h⁻¹). Putting these emission ranges into real-world context, the maximum emission from unplugged and abandoned wells was measured at 177 g CH₄ h⁻¹ in West Virginia (Riddick et al., 2019a), 175 g CH₄ h⁻¹ in Pennsylvania (Pekney et al., 2018), 146 g CH₄ h⁻¹ across the US (Townsend-Small et al., 2016) and 35 g CH₄ h⁻¹ in the UK (Boothroyd et al., 2016). As most of the methods presented here require access to the source, we considered 200 g CH₄ h⁻¹ to be a sensible limit to the emission rate and is larger than the emissions observed by many previous studies. Therefore, the scope of this study is limited to estimating CH₄ emissions from a single point source that we would realistically be able to approach and measure, i.e. between 40 and 200 g CH₄ h⁻¹.”

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 Specific comment 20:

line 110: the exact dates and times in which each experiment took place needs to be provided.

Response to reviewer:

Have been added as Supplementary Material Section 4

Reviewer 1 Specific comment 21:

line 111: it's unclear which exact experiments were done. it would be helpful if the authors could provide a spreadsheet with all the tests that were conducted in the supporting information.

Response to reviewer:

Have been added as Supplementary Material Section 4

Reviewer 1 Specific comment 22:

line 115: how were the emission rates set? What is the type of flow controller? What are the gases that are used?

Response to reviewer:

At the METEC site, compressed natural gas, with methane compositions ranging from 85 to 95% vol, is supplied from two 145 L cylinders. Flown rates are controlled using a pressure regulator and precision orifices.

Changes to the manuscript:

At L82

“At the METEC site, compressed natural gas, with methane compositions ranging from 85 to 95%vol, is supplied from two 145 L cylinders and flow rates controlled using a pressure regulator and precision orifices.”

Reviewer 1 Specific comment 23:

line 115-116: how do you define/determine what is safe or not?

Response to reviewer:

From experience, it becomes difficult to work effectively in areas near NG emissions greater than $200 \text{ g CH}_4 \text{ h}^{-1}$. As a rule of thumb at this emission rate, we suggest access to areas closer than 10 m of the source be restricted this is often evident from the response of the 4-gas monitor. As most of the methods require access to the source, we considered this to be a sensible limit to emission. The 4-gas monitor indicated low oxygen while conducting the static chamber measurements and the measurements we halted.

Changes to the manuscript:

At L56:

“From experience and the response of a 4-gas monitor, working close enough to measure emissions greater than $200 \text{ g CH}_4 \text{ h}^{-1}$ for many of these methods (especially the chambers and Hi Flow) can be unsafe, therefore this study

is limited to quantifying CH₄ emissions between the lowest flow METEC can produce (40 g CH₄ h⁻¹) and the highest flow we feel comfortable measuring with these methods (200 g CH₄ h⁻¹).”

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 Specific comment 24:

line 116: why not lower than 40 g/hr?

Response to reviewer:

The lowest emission rate METEC can produce is 40 g/hr.

Reviewer 1 Specific comment 25:

line 125: it's unclear if a static chamber measurement was done?

Response to reviewer:

The static chamber measurements were not reported. We felt the method was intrinsically unsafe when other methods were safer. During the experiment the 4-gas monitor on our researcher indicated low oxygen while conducting the static chamber measurements and the measurements we halted.

Changes to the manuscript:

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely, wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 Specific comment 26:

line 127: what is the location and size of the fan inside the chamber selected? and how was this selected?

Response to reviewer:

We used a fan similar to the one presented in Riddick et al. (2019). This was selected to ensure that the air was fully mixed throughout the cylindrical chamber.

Changes to the manuscript:

At L102:

“A fan was secured inside the chamber and used to circulate the air following the method of (Riddick et al. 2019a) to ensure the air inside the chamber was fully mixed.”

Reviewer 1 Specific comment 27:

line 128: What is meant by "three further air samples? Further to what?

Response to reviewer:

Should have said that a sample was taken at time t_0 .

Changes to the manuscript:

L 106:

“When the chamber is sealed with the ground, following Riddick et al. (2019a), an air sample is drawn using a gas syringe. During the experiment at least three further air samples are taken at regular intervals (Pihlatie et al., 2013; Collier et al., 2014),”

Reviewer 1 Specific comment 28:

line 132: what were the shapes of the chambers? what is the aspect ratio (height to diameter)?

Response to reviewer:

These were all rigid plastic cylindrical chambers, with heights approximately 1.5 times the diameter.

Changes to the manuscript:

At L107

“Two sizes of static chambers were used in this experiment (0.12 m³ and 0.5 m³; Figure 1), the chambers were made from rigid plastic cylindrical chambers, with heights approximately 1.5 times the chamber’s diameter. The chambers sizes was based on a measurable concentration change over time for given release rates, however, it is unlikely that the larger size is practical for field deployment.”

Reviewer 1 Specific comment 29:

line 133: how was the quality of the ground seal determined?

Response to reviewer:

Seal was made in line with the information in Riddick et al. (2019).

Changes to the manuscript:

L106

“, following Riddick et al. (2019a),”

Reviewer 1 Specific comment 30:

line 135: is this experiment a copy of Kang et al (2014), Pihlatie et al (2013), or Collier et al (2014)? Which one took four samples? Kang et al (2014) took 7 to 8 samples.

Response to reviewer:

As stated in L107 “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 (Pihlatie et al., 2013; Collier et al., 2014)”.

Reviewer 1 Specific comment 31:

line 135: there is a "s" missing. it should be "four samples".

Response to reviewer:

Corrected

Reviewer 1 Specific comment 32:

line 137: were there duplicates and blanks taken?

Response to reviewer:

Blanks (field or lab) were not required as all air samples were measured on the day of the experiment (L171 should say hours not days). Also, all experiments were repeated 3 times.

Changes to the manuscript:

At L118

“All samples were measured within two hours of collection. All experiments were repeated three times.”

Reviewer 1 Specific comment 33:

line 142: does this imply that emission rates were calculated for test even if only three samples were collected?

Response to reviewer:

Emission rates were calculated from the four measurements.

Reviewer 1 Specific comment 34:

line 146: why was this chamber size selected? what is the shape and aspect ratio? How useful is this size for field measurements of oil and gas wells?

Response to reviewer:

These chambers sizes were chosen as they are what we had to hand at METEC and calculated as a safe size to use. These sizes are impractical for field deployment.

Changes to the manuscript:

L107:

“Two sizes of static chambers were used in this experiment (0.12 m³ and 0.5 m³; Figure 1), the chambers were made from rigid plastic cylindrical chambers, with heights approximately 1.5 times the chamber’s diameter. The chambers sizes was based on a measurable concentration change over time for given release rates, however, it is unlikely that the larger size is practical for field deployment.”

Reviewer 1 Specific comment 35:

line 146: what type of plastic is used?

Response to reviewer:

Rigid plastic

Changes to the manuscript:

At L110

“rigid plastic”

Reviewer 1 Specific comment 36:

line 149: what is the detection limit of the HXG-2D? What are the methane concentrations observed inside the chamber?

Response to reviewer:

Added to the manuscript

Changes to the manuscript:

At L122:

“detection limit 10 ppm and range 0 to 40,000 ppm”

Reviewer 1 Specific comment 37:

line 150-151: blanks and duplicates taken?

Response to reviewer:

Blanks (field or lab) were not required as all air samples were measured on the day of the experiment (L186 should say hours not days). Also, all experiments were repeated 3 times.

Changes to the manuscript:

At L148

“As with the samples from the static chamber, the vials were measured within two hours of collection. All experiments were repeated three times.”

Reviewer 1 Specific comment 38:

line 155: how big is the hose end? Was the source enclosed by the Hi-Flow sampler?

Response to reviewer:

The Hi Flow has a bag, this was placed over the leak.

Changes to the manuscript:

At L162:

“the bag containing the hose”

Reviewer 1 Specific comment 39:

line 167: any concerns with topography and large objects (e.g., buildings, trees, and other infrastructure)?

Response to reviewer:

No, METEC is a big open flat site.

Reviewer 1 Specific comment 40:

line 177: what is the minimum distance between the sources and the detector?

Response to reviewer:

5 m, as stated on Line 191

Reviewer 1 Specific comment 41:

line 181: how different were the environments in which the experiments were conducted? Importantly, were experiments described in sections 2.1 to 2.5 all conducted on the same day. If they were all conducted on different days, then the uncertainties calculated cannot be directly compared.

Response to reviewer:

This study is the first to test the uncertainty of each of these methods in a controlled test. Therefore, there is no evidence to suggest that the uncertainty will change on different days in different environmental conditions.

Reviewer 1 Specific comment 42:

line 190-191: The collection of gas vials is not a requirement of the static chamber methodology.

Response to reviewer:

This is true, but this is the method we are testing.

Using a trace gas analyzer to measure the concentrations inside the chamber is questionable. If the gas is removed from the chamber (i.e. the analyzer outlet is vented outside the chamber) this system then becomes a dynamic chamber. If the measured gas is reintroduced to the chamber (i.e. the analyzer outlet is vented back to the chamber) this means gas of a lower concentration is being continually added to the “closed” system. It is unclear of the uncertainty caused by this cycling. To my knowledge, using a gas analyzer in series with the static chamber has never been tested.

Change to the manuscript:

At L239:

“To address the first shortcoming, a trace gas analyzer could be used to measure the concentrations inside the chamber. As trace gas analyzers use a pump to draw air into the measurement cavity, the analyzer could be arranged in one of two ways. Both introduce additional uncertainty into the quantification. If the gas is removed from the chamber (i.e. the analyzer outlet is vented outside the chamber), the static chamber becomes a dynamic chamber and the analyzer flow rate must be accounted for in the quantification. If the measured gas is reintroduced to the chamber (i.e. the analyzer outlet is vented back to the chamber), a gas of lower concentration is being continually added to the “closed” system and it is therefore unclear how much uncertainty is caused by this cycling. Furthermore, the linear response of a portable trace gas analyzer, e.g. the ABB GLA131-GGA Greenhouse Gas Analyzer (<https://new.abb.com/>), is 100 ppm. Using the lowest emission rate in the study, 40 g CH₄ h⁻¹, and the largest chamber, 0.5 m³, the concentration inside the chamber will exceed the linear range within 7 seconds. It is unlikely that gas will mix entirely throughout the chamber in 7 seconds and emission estimates are unlikely to be accurate. Another alternative could be using a lower precision sensor with a larger detection range, such as the SGX INIR- ME100 (<https://sgx.cdistore.com/>) that can measure from 200 ppm to 100% methane bv, but safety issues remain.”

At L257:

“The static chamber could be automated to release gas when CH₄ concentration inside the chamber approaches LEL to prevent chamber becoming explosive. The major shortcoming of this strategy is that the automation of a chamber takes away the operator’s control of when gas is released, which could happen at an inconvenient during measurement. If an automated system is used for collecting gas of unknown composition self-contained breathing apparatus should be worn.”

Reviewer 1 Specific comment 43:

line 193: the static chamber can be used with a methane analyzer (e.g., El Hachem and Kang, 2022), overcoming the first and second shortcoming.

Response to reviewer:

As the method in El Hachem and Kang, 2022 is presented, a GasScouter is used to measure the concentration of gas within a chamber. It does not comment on whether the outlet of the GasScouter is returned to the chamber, therefore, it is not clear if this is actually a static chamber.

The static chamber is inherently dangerous and we cannot advocate it’s use in measuring unknown compositions of gas at unknown flow rates. It is unreasonable to assume that measurement teams would want to use self-contained breathing apparatus for every well they measure at if there are safer alternative methods available.

Reviewer 1 Specific comment 44:

line 196-198: El Hachem and Kang (2022) conducted measurements from H₂S-emitting wells using a self-contained breathing apparatus. There are many options available in industry to ensure safe working conditions when toxic gases are present.

Response to reviewer:

It is unreasonable to expect full breathing apparatus as mandatory at all measurement locations. As such we have added a caveat to the manuscript.

Changes to the manuscript:

At L74:

“We add the caveat that we will only present data from measurement methodologies can be conducted safely wearing PPE as regulated at the Colorado State University Methane Emissions Technology Evaluation Center (METEC) facility in Fort Collins, CO, USA (steel toe boot, FR overalls, hard hat, safety glasses and 4-gas monitor).”

Reviewer 1 Specific comment 43:

line 198: what about for measuring low emitting sources?

Response to reviewer:

Gas composition is still unknown even for low methane emissions; therefore the risks remain. This is a major shortcoming and other safer methods should be used.

Reviewer 1 Specific comment 44:

line 200: why isn't "cost" italicized like the rest?

Response to reviewer:

Changed as suggested.

Reviewer 1 Specific comment 45:

Table 1: I'm surprised that the HiFlow sampler is only \$5k. Is this correct? The static and dynamic chamber measurements conducted here use a GC, which is around \$50k. So it's definitely not free. Even just getting the gas concentrations analyzed elsewhere is not free.

Response to reviewer:

Have added text to that effect.

Change to manuscript

At L272

◇ Cost of GC analysis will vary by laboratory.

Reviewer 1 Specific comment 46:

Table 1: why is there no time for measurement and analysis for the static chamber? same for the accuracy.

Response to reviewer:

As mentioned above the static chamber is not presented here as we found it inherently dangerous.

At L271

“- the static chamber data is not presented here as the method was found to be inherently dangerous.”

Reviewer 1 Specific comment 47:

line 209: there are other ways to reduce the potential of CH₄ concentrations in the chamber reaching explosive levels when using static chambers.

Response to reviewer:

Have added some text to the discussion about using automated opening chambers.

At L251:

“The static chamber could be automated to release gas when CH₄ concentration inside the chamber approaches LEL to prevent chamber becoming explosive. The major shortcoming of this strategy is that the automation of a chamber takes away the operator’s control of when gas is released, which could happen at an inconvenient during measurement. If an automated system is used for collecting gas of unknown composition self-contained breathing apparatus should be worn.”

Reviewer 1 Specific comment 48:

line 210: the need for a power source is another important shortcoming of the dynamic chamber method.

Response to reviewer:

Noted

Changes to the manuscript:

L277

“and a 12 V power source for the pump”

Reviewer 1 Specific comment 49:

line 266-267: what is the dynamic chamber the HiFlow more accurate than?

Response to reviewer:

This sentence doesn't make sense and has been removed.

Reviewer 1 Specific comment 50:

line 273: who is this decision-making paradigm for?

Response to reviewer:

On reflection we have removed the decision making paradigm.

Reviewer 1 Specific comment 51:

line 282: what are the conditions in this study? This needs to be better described to assess the applicability of the results.

Response to reviewer:

As described above, this study is the first to test the uncertainty of each of these methods in a controlled test. Therefore, there is no evidence to suggest that the uncertainty will change in different environmental conditions.

Reviewer 1 Specific comment 52:

Figure 4. Many estimates (e.g., the USEPA's GHGI) involve wells with no measurements. It's not possible for all wells to be measured. So if there is no trace gas analyzer, there is always the emission factor approach. But of course, that's not a good predictor of the emissions at a given well but over some large population, it may be representative. So this brings us back to the question of who this figure is for. This figure needs more context in the caption and the text.

Response to reviewer:

This aim of this work is to provide evidence as to how well a method performs quantifying emissions between 40 g CH₄ h⁻¹ and 200 g CH₄ h⁻¹.

As you say, emission factors do not quantify emissions from an individual source, therefore are unlikely to be representative.

We suggest this work is of interest to those currently quantifying methane emissions, those that plan to quantify methane emission and regulators that have justify decisions on policy.