

Evaluating the ~~feasibility~~accuracy of ~~using~~ downwind methods ~~for to~~ quantifying point source oil and gas emissions using continuous monitoring fence-line sensors

Mercy Mbua<sup>\*,1</sup>, Stuart N. Riddick<sup>1,2</sup>, Elijah Kiplimo<sup>1</sup>, Kira B. Shonkwiler<sup>1</sup>, Anna Hodshire<sup>1</sup> and Daniel Zimmerle<sup>1</sup>

<sup>1</sup>*The Energy Institute, Colorado State University, CO, 80524, Fort Collins, USA*

<sup>2</sup>*Department of Science, Engineering and Aviation, University of the Highlands and Islands Perth, Crieff Road, Perth, PH1 2NX, UK*

\*Correspondence to: Mercy Mbua (Mercy.Mbua@colostate.edu)

**Abstract.** The ~~dependable~~accurate reporting of methane (CH<sub>4</sub>) emissions from point sources, such as fugitive leaks from oil and gas infrastructure, is important for profit maximization (retaining more hydrocarbons), evaluating climate ~~change~~ impacts, assessing CH<sub>4</sub> fees for regulatory programs, and validating CH<sub>4</sub> methane intensity in differentiated gas programs. Currently, there are disagreements between emissions reported by different quantification techniques for the same sources. It has been suggested that downwind CH<sub>4</sub> quantification methods using CH<sub>4</sub> measurements on the fence-line of production facilities could be used to generate emission estimates from oil and gas operations at the site level, but it is currently unclear how accurate the quantified emissions are. To investigate ~~model~~downwind methods' accuracy, this study uses fence-line simulated data collected during controlled release experiments as input for closed-path eddy covariance, aerodynamic flux gradient, ~~and~~ the Gaussian plume inverse methods, and the backward Lagrangian stochastic model in a range of atmospheric conditions. ~~The results show that both the eddy covariance and aerodynamic flux gradient methods underestimated emissions in all experiments. Although calculated emissions had significant uncertainty, the Gaussian plume inversion method performed better. The uncertainty was found to have no significant correlation with most measurement variables (i.e. downwind measurement distance, wind speed, atmospheric stability, or emission height), which indicates that the Gaussian method can randomly either underestimate or overestimate emissions. For eddy covariance, downwind measurement distance and percent error had negative correlation indicating that far away emissions sources were likely underestimated or be undetected. The study concludes that using fence line measurement data as input to eddy covariance, aerodynamic flux gradient or Gaussian plume inverse method to quantify CH<sub>4</sub> emissions from an oil and gas production site is unlikely to generate representative emission estimates.~~ Generally, results show that flux quantification methods provide more reasonable estimates compared to point-source specific models especially when multiple releases are happening at the facility level. The closed-path eddy covariance quantified emissions with a mean relative factor (estimated emission over actual emission) of 0.7 to 1 for single-release single-point emissions, and within a mean relative of 1 and 2.4 for multi-release single-point emissions. The aerodynamic flux gradient method quantified emissions within a mean relative factor of 1.3 to 1.7 for single-release single-point emissions, and between 2.4 and 3.3 for multi-release single-point emissions. The Gaussian plume inverse model quantified emissions within a mean relative factor of between 2.4 and 2.6 for single-release single-point emissions, but largely overestimated emissions when multiple releases were happening; mean relative factor between 16 and 25. Similarly to the Gaussian plume inverse method, the backward

Lagrangian stochastic model for point sources using WindTrax quantified within a mean relative factor of between 0.8 to 1 for single-release single-point emissions, but largely overestimated emissions for multi-release single-point emissions; mean relative factor of 3.9 and 11958. As continuous monitoring of oil and gas sites can involve complex emissions where plumes are not defined due to multiple sources, this study shows that common downwind point source dispersion models could largely overestimate emissions. This study recommends more testing of flux quantification models for oil and gas continuous monitoring quantification.

**Keywords:** Continuous monitoring; oil and gas; point source; closed-path eddy covariance; aerodynamic flux gradient; Gaussian plume inverse method; backward Lagrangian stochastic model

## **1 Introduction**

~~Methane (CH<sub>4</sub>), the primary component of natural gas (NG), is a potent greenhouse gas with a global warming potential of 27 carbon dioxide (CO<sub>2</sub>) equivalent over 100 years (US EPA, 2016). Methane emissions reduction is a key part of global initiatives to reduce climate change (Chung, 2021). The 2021 Global Methane Assessment by the Climate and Clean Air Coalitions (CCAC, 2024) and the United Nations Environment Programme (UN Environment Programme, 2024) state that reducing CH<sub>4</sub> emissions from anthropogenic sources by 45% in 2030 would result in avoiding a global atmospheric temperature increase of 0.3°C in 2045 (Chung, 2021). Such measures would align with the Paris Agreement goal of limiting global temperature rise to 1.5°C by 2030 (United Nations Climate Change, 2015). The US is one of the countries that reports its total greenhouse gas emissions to the Intergovernmental Panel on Climate Change as part of the Paris Agreement (United Nations Climate Change, 2015).~~

Reducing methane (CH<sub>4</sub>) emissions from oil and gas systems is necessary for adhering to regulations and voluntary reporting frameworks such as the Oil & Gas Methane Partnership 2.0 (OGMP 2.0). The OGMP 2.0 provides a comprehensive measurement-based international reporting framework allowing companies to stay ahead of regulatory compliance requirements, meet investor and market pressure, have an enhanced corporate image, and prevent revenue loss by lowering their emissions.

~~In the US, Currently, the amount of CH<sub>4</sub> emitted from US oil and gas production are compiled is calculated by the US Environmental Protection Agency (EPA) under Subpart W. Typically, companies use using a bottom-up inventory approach. The inventory approach multiplies where emission factors (CH<sub>4</sub> emissions per equipment e.g., separator or emissions per event e.g., liquid unloading) are multiplied by activity factors (total number of pieces of equipment or events (OAR US EPA, 2023)) to generate emissions. This quantification approach has several shortcomings, including: 1. It separately calculates CH<sub>4</sub> emissions from natural gas and petroleum systems, which practically are not independent systems, and can result in bias based on changes in gas to oil ratios throughout a basin (Riddick et al., 2024a); 2. Some emission factors used are outdated (Riddick et al., 2024b) and others do not account for the temporal and spatial variation in emissions (Riddick and Mauzerall, 2023); and 3. Emission factors do not account for the long-tail distributions in emissions distributions (Riddick et al., 2024b). Recently, mechanistic models, such as the Colorado State University's Mechanistic Air Emissions Simulator (MAES), have been developed to address shortcomings in bottom-up CH<sub>4</sub> reporting (Colorado State University, 2021), but these still depend on direct measurements to inform emission factors.~~

Top-down methods, including using aircraft such as Bridger Photonics LiDAR (Light Detection and Ranging; 90% detection limit of  $\sim 2 \text{ kg h}^{-1}$ ) (Johnson et al., 2021) and satellites such as Carbon Mapper (predicted 90% detection limit of about  $100 \text{ kg h}^{-1}$ ) (“Carbon Mapper - Science & Technology,” n.d.); can also be used to infer emissions. ~~For example, Carbon Mapper satellites can locate and quantify  $\text{CH}_4$  emissions using absorption spectra taken from space (Carbon Mapper, 2024).~~ However, these survey methods only quantify emissions over a very short period of time ( $< 10 \text{ s}$ ) and observations are typically made during the day which can often coincide with maintenance activities that can bias emissions and result in overestimation (Riddick et al., 2024a; Zimmerle et al., 2024). Additionally, different top-down technologies measuring the same source have disagreed in their reported emissions which has called into question the credibility of these methods (Brown et al., 2023; Conrad et al., 2023). As a result, ensuring accuracy in models and technologies used in  $\text{CH}_4$  emissions quantification has been a complex issue.

~~The accurate reporting of  $\text{CH}_4$  from fugitive emissions at oil and gas production sites is important for evaluating potential effects on climate change, correctly assessing  $\text{CH}_4$  fees on companies as part of the Methane Emissions Reduction Program created under the 2022 Inflation Reduction Act (OA US EPA, 2023), and validating  $\text{CH}_4$  content of reported differentiated gas composition where NG companies differentiate their market products based on the environmental impact (CO2EFFICIENT, 2022). Direct measurements have been recommended to augment/update emissions factors used in bottom-up inventories and for better understanding temporal/spatial variability of emissions (Riddick et al., 2024).~~

~~Downwind methods are widely used to directly measure  $\text{CH}_4$  emissions from area and point sources at site/basin levels due to their low cost and wide coverage within a short time (Caulton et al., 2018; Heimbürger et al., 2017; Riddick et al., 2020, 2022a; Sonderfeld et al., 2017). Commonly used downwind quantification methods include the Gaussian plume inversion method, eddy covariance, backward Lagrangian stochastic models, aerodynamic flux gradient, mass balance method, the EPA Other Test Method (OTM 33) and the Gaussian puff modelling approach (Denmead, 2008; Edie et al., 2020; Foster-Wittig et al., 2015; Jia et al., 2023; Kamp et al., 2020; Nemitz et al., 2018; Shaw et al., 2021).~~

Currently, fence-line methods are used to detect, localize and quantify emissions. This approach uses point sensors fixed to the fence-line of the production site and emissions detected when the measured concentration exceeds a threshold, localized by triangulating multiple detections and quantified using a simple dispersion modelling framework, usually based on a Gaussian plume inverse approach (Bell et al., 2023; Day et al., 2024; Jia et al., 2023; Riddick et al., 2022a). The detection and localization of simulated fugitive emissions have been successful, with controlled release testing against point sensors and scanning/imaging solutions reporting a 90% probability of detection for emissions of between  $3.9$  and  $18.2 \text{ kg CH}_4 \text{ h}^{-1}$  (Ilonze et al., 2024). Major shortcomings have been identified using a fence-line approach with quantified emissions reported at between a factor of 0.2 to 42 times for emissions between  $0.1$  and  $1 \text{ kg CH}_4 \text{ h}^{-1}$ , and between 0.08 and 18 times ~~for~~ emissions greater than  $1 \text{ kg CH}_4 \text{ h}^{-1}$  (Ilonze et al., 2024). As a result, questions have arisen if other approaches, such as the eddy covariance (EC) or aerodynamic flux gradient (AFG) would generate more accurate results. These methods have been suggested as they have been used to quantify emissions from other sectors, i.e. agriculture (Denmead, 2008; Morin, 2019) and landfills (Xu et al., 2014), as well as have been used to quantify emissions in large downwind areas (Vogel et al., 2024). ~~and~~

Such quantification does not require assumptions made on downwind dispersion coefficients or micrometeorology that are often required for dispersion modelling (Denmead, 2008).

Due to interest in using a subset of these methods to quantify emissions from oil and production sites, this study will evaluate the quantification accuracy of the ~~closed-path EC~~ eddy covariance, aerodynamic flux gradient, AFG, and Gaussian plume inverse ~~model (GPIM), and the backward Lagrangian stochastic model (bLs)~~ methods for oil and gas point source quantification.

Eddy covariance is a vertical flux gradient measurement that measures CH<sub>4</sub> emissions based on the covariance between CH<sub>4</sub> concentrations measured using a fast-response analyzer (> 10 Hz) and vertical wind vector measured by a fast-response sonic anemometer (>10 Hz) (Figure 1A; Morin, 2019). It is typically implemented over long homogeneous fetches where eddy mixing scale is a small fraction of the distance from the site providing more predictable vertical transport. Dumortier et al., 2019 used EC to estimate known point source emissions at a cow's muzzle height and reported the model could estimate emissions between 90 and 113% of the true emission. Dumortier et al., 2019 stated the optimal controls for point source quantification and footprint modelling are using running mean, 15-minute averaging periods, no application of Foken and Wichura (1996) stationarity filter and use of the Kormann and Meixner (2001) footprint function. The study tested the model using an artificial CH<sub>4</sub> source at 0.8 m, programmed to emit when winds were coming from the source direction ( $\pm 45^\circ$ ), and when friction velocity ( $u_*$ ) was above 0.13 m s<sup>-1</sup>. In Dumortier et al. (2019)'s point-source testing, they noted that amplitude resolution, skewness and kurtosis tests were disabled as they deleted almost all periods involving the artificial source in the footprint. Rey-Sanchez et al. (2022) studied the accuracy of Hsieh model (Hsieh et al., 2000), the Kljun model (Kljun et al., 2015) and the K & M model (Kormann and Meixner, 2001b) in calculating the footprint of point source hot spots using footprint-weighted flux maps. The study reported the K & M model to be the most accurate. Polonik et al. (2019) compared five gas analyzers, two open-paths, two enclosed-path and one closed-path analyzer for carbon dioxide EC measurements. The study noted that while open-path sensors minimize spectral attenuation and require smaller spectral correction factors compared to sensors with an inlet tube such as a closed-path sensor, open-path sensors risk data loss in non-ideal conditions like precipitation, fog, dust or dew. The main challenge of applying EC for continuous monitoring of oil and gas sites is instrument limitations (requires deployment of multiple sensors throughout a facility; sensor cost is a factor) and statistical tests as well as quality controls could filter out some of the data.

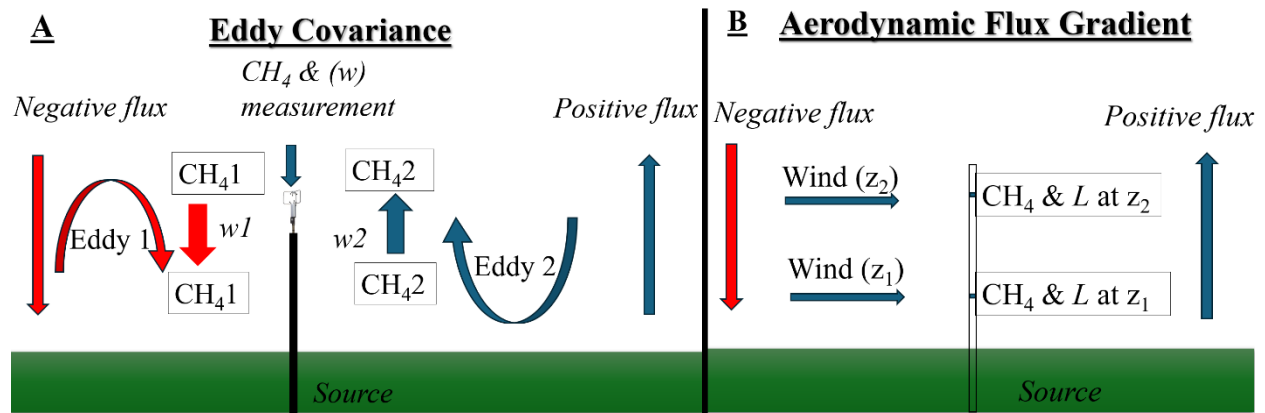


Figure 1: Illustrations of eddy covariance (A) and flux gradient measurements (B) where  $CH_4$  is methane concentrations,  $w$  is the vertical wind speed,  $L$  is the Monin-Obukhov length (measure of atmospheric stability), and  $z$  is the measurement height.

The aerodynamic flux gradient AFG method quantifies  $CH_4$  emissions from a source by comparing  $CH_4$  concentrations at two heights (Figure 1B; Querino et al., 2011). Kamp et al. (2020) used the method to calculate ammonia fluxes over a grass field using a single analyzer by alternating two heights and reported 7% mean relative difference in flux in this approach compared to continuous measurements at two heights. Generally, the AFG approach is designed for homogeneous sources where footprints at different sensor heights would not affect quantification results, and its applicability to point source quantification, currently, is limited.

The Gaussian Plume Inverse GPIM method calculates  $CH_4$  emission rate as a function of mole fraction at a point in space ( $x, y, z$ ), as a function of the downwind distance, perpendicular distance (crosswind), mean wind speed and atmospheric stability (Figure 2A; Jia et al., 2023; Riddick et al., 2022b). This method has been used to quantify emissions from oil and gas production sites especially for survey solutions (Riddick et al., 2022b). For a single point-source, Riddick et al. (2022b) reported absolute uncertainties of between 40.7 and 60%. Foster-Wittig et al. (2015) using controlled single point source tests reported average errors of between -5 to 6%. The limitations of the GPIM method has been used to quantify emissions from oil and gas production sites (Caulton et al., 2014; Riddick et al., 2022b) but it assumes a homogeneous emission source, steady-state flow, and uniform dispersion of gas in an open area free of obstructions (Hutchinson et al., 2017) (Hutchinson et al., 2017).

The bLs model adapted in WindTrax can simulate the transport of gases from point sources that emit them (Figure 2B; Crenna, 2006). The model releases individual particles and follows them along their unique path in air by mimicking random, turbulent motion of the atmosphere. Tagliaferri et al. (2023) investigated the validity of WindTrax in quantifying emissions from complex sources and reported the model to be reliable under neutral conditions, underestimated emission rates during unstable stratification and overestimated emissions during stable conditions. Similarly to the GPIM method, the model assumes free flow of air in the absence of obstructions and uses time-averaged data as input.

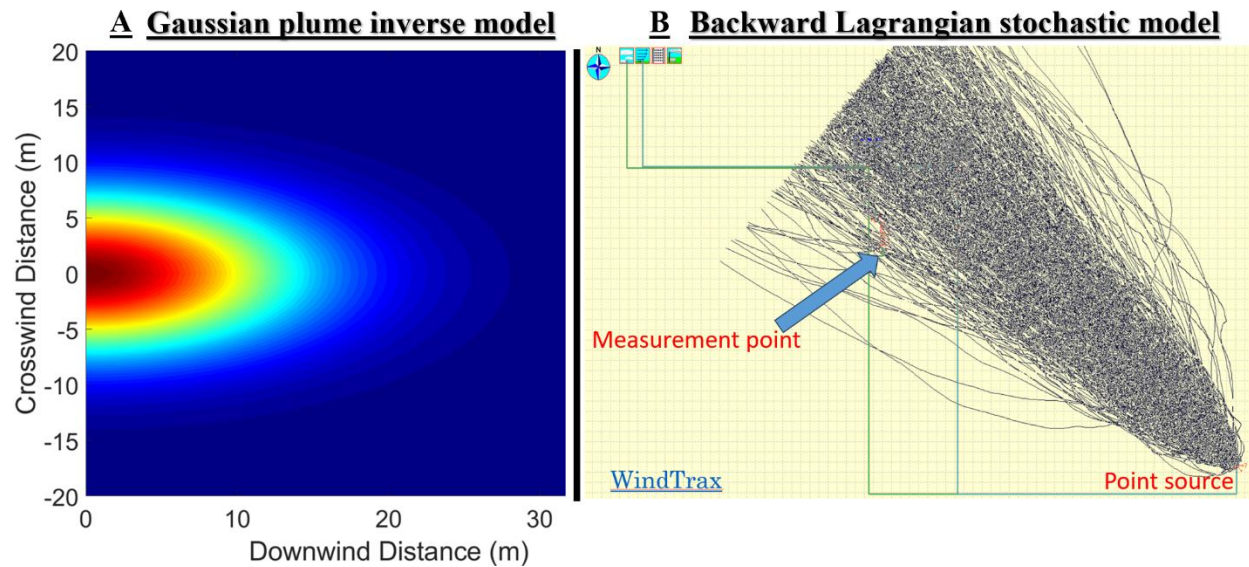


Figure 2. A: An illustration of a plume that follows a Gaussian plume inverse model where emission rate can be inferred from concentrations at different downwind distances and crosswind distances. B: An illustration of how the backward Lagrangian stochastic model traces particles to the source. These approaches were developed to quantify emissions from single point or area emission sources and have not been tested against a controlled release to evaluate their quantification performance. The aerodynamic flux gradient and eddy covariance, for example, have been used to measure trace gas, e.g., nitrogen oxide and carbon dioxide, fluxes from large croplands (Kamp et al., 2020).

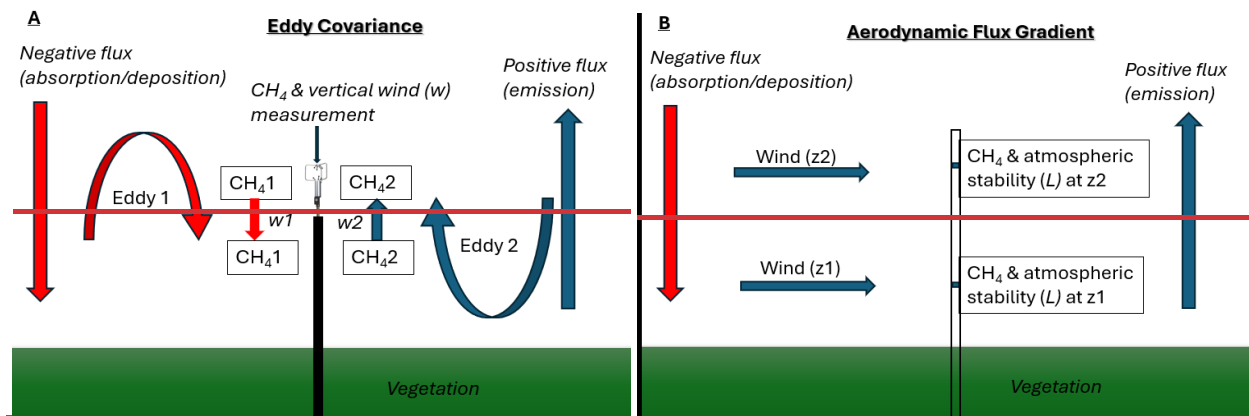


Figure 1: Illustrations of eddy covariance (A) and flux gradient measurements (B) where  $CH_4$  is methane concentrations,  $w$  is the vertical wind speed,  $L$  is the Monin-Obukhov length (measure of atmospheric stability), and  $z$  is the measurement height.

The Gaussian plume inversion method has been used to quantify emissions from oil and gas production sites (Caulton et al., 2014; Riddick et al., 2022b) but it assumes a homogenous, steady-state flow, uniform dispersion of gas in an open area free of obstructions (Hutchinson et al., 2017).

Continuous monitoring of  $CH_4$  emissions using fence line sensors requires proper quantification of intermittent and persistent releases from oil and gas during all release (complex emission profiles) and atmospheric conditions (unstable, neutral and stable). Oil and gas emissions are characterized by intermittent, non-uniform, single

or multiple point source emissions, varying in leak size, location, height and distance between the source and sensor, and are typically in complex aerodynamic environments (i.e. not flat). An ideal quantification model should always quantify emissions and should capture short and long-lasting emission events. Most models have been validated to work best during neutral conditions for single point sources. However, it is important to test and apply these models during non-neutral conditions as well as these are part of real-world conditions where continuous monitoring is applied. In this study, we evaluate if using a readily available CH<sub>4</sub> cavity ring down analyzer for models' quantification such as the closed-path EC is a feasible solution to quantify point source emissions.

This study aims to inform the feasibility of downwind quantification models in oil and gas settings by investigating which models are likely to work most of the time with instrumentation that is typically available for fence-line deployment. Fence-line sensor deployments involve multiple sensors, continuously running in all conditions and providing emissions data. The need for accurate CH<sub>4</sub> quantification and reporting necessitates evaluating the performance of these downwind quantification approaches in different controlled release and characterized meteorological conditions, to ensure credibility. Using robust releases and environmental conditions, t  
This study aims to investigate the performance of these methods in quantifying emissions for known gas release rates and evaluating uncertainties that could result in incorrect CH<sub>4</sub> reporting. Specifically, the study will (1) evaluate the overall quantification accuracy of closed-path eddy covariance EC, aerodynamic flux gradient AFG, bLs model, and the Gaussian plume inverse GPIM method in quantifying single-release single-point and multi-release single-point emissions single-point and multi-point emissions that simulate oil and gas emissions, (2) evaluate the probability of these models quantifying within a defined range (i.e.  $\pm 30\%$ ), and (3) determine the mean relative factor (estimated emissions over actual emission) for these models ~~investigate which variables have the largest effect on quantification uncertainty.~~

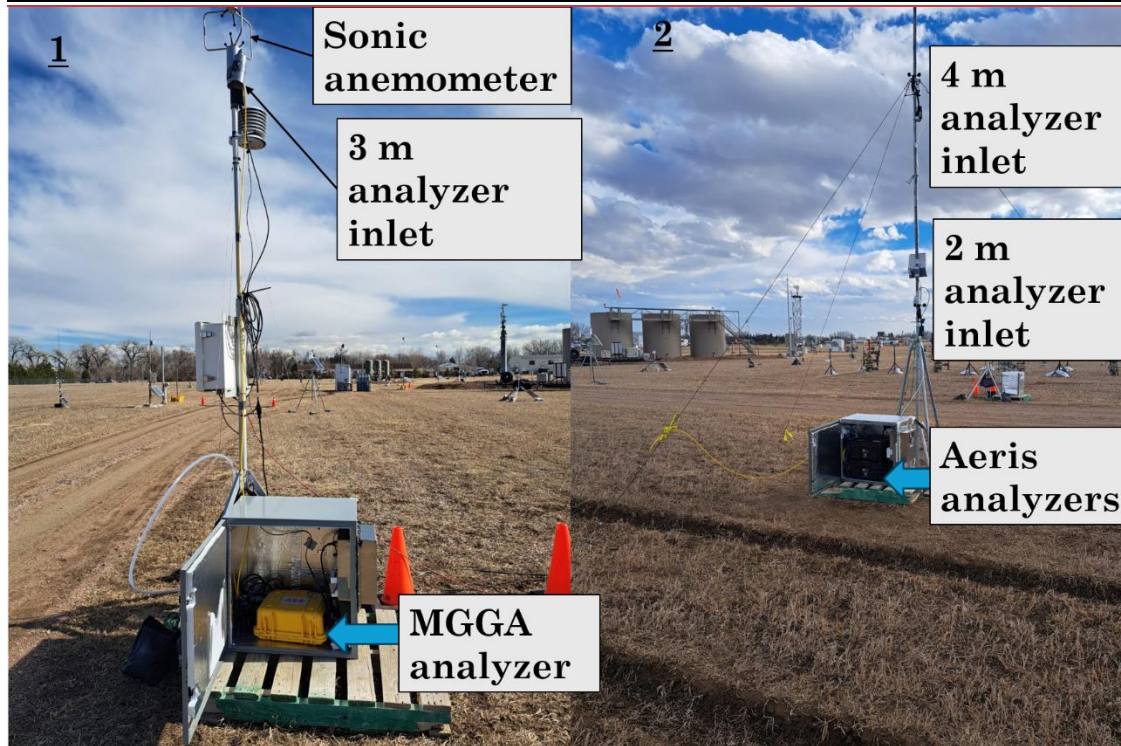
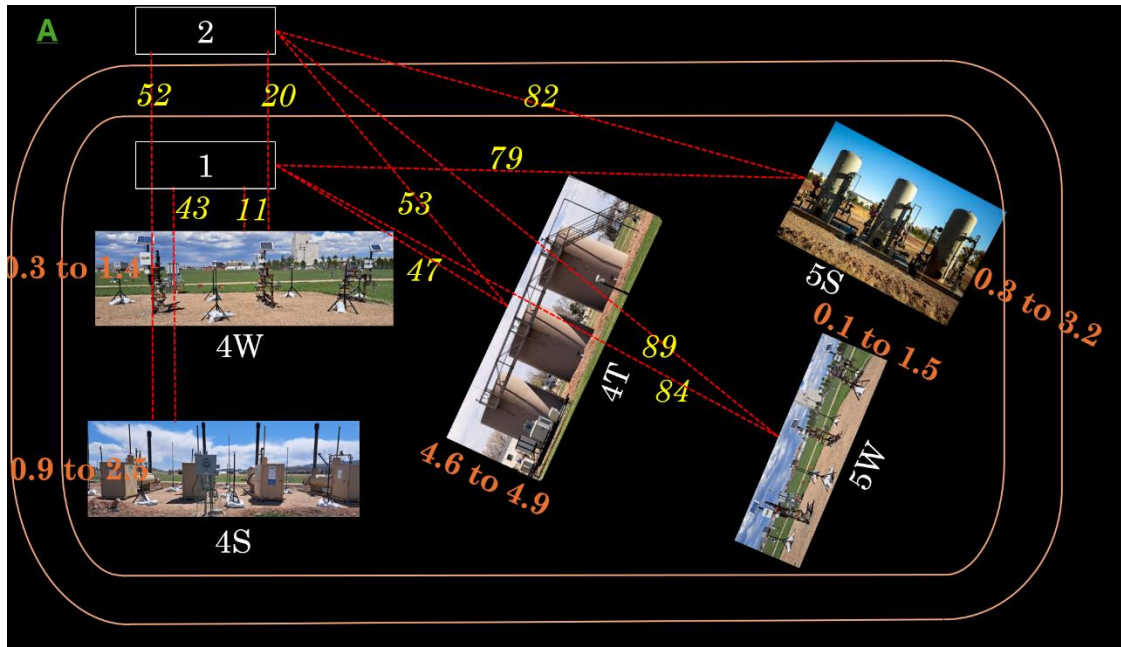
## 2 Methods

### 2.1 Experimental Setup

Controlled release experiments were conducted at the Colorado State University's Methane Emissions Technology Evaluation Center (METEC) in Fort Collins, CO, (USA, 65 miles north of Denver) between February 8, and March 20~~7~~, 2024. The METEC center is a simulated oil and gas facility that does controlled testing for emissions leak detection and quantification technology development, field demonstration, leak detection protocol and best practices development (METEC, 2025). The weather conditions during the test period were mostly sunny but precipitation was also observed (32 sunny, 7 snowy, 12 rainy, 7 cloudy and 1 foggy day; Supplementary Information Section 1). Wind speeds were between 0 and 25 m s<sup>-1</sup> and temperatures ranged between -15 and +19 °C (Supplementary Information Section 1). Two stationary masts holding the instrumentation were setup on the North-West corner of METEC to take advantage of the predominant wind direction, avoid the largest aerodynamic obstructions and to simulate the likely placement of a fence-line instrument (Figure 23A; Day et al., 2024; Riddick et al., 2022a). Fence-line sensors are typically placed within the oil and gas perimeter (~30 m) (Riddick et al., 2022a). This study collected data for what we considered as both close and far away releases; distances between 9 and 94 m.

Methane concentration data for closed-path EC, GPIM and bLs methods were collected through an inlet tubing (3.275 mm inner diameter) at 3 m height, connected to the ABB (Zurich, Switzerland) GLA131 Series Microportable Greenhouse Gas Analyzer (MGGA) set to sample at 10 Hz. The MGGA is a closed-path greenhouse gas analyzer with a ~3.2 lpm pump flowrate, 10 cm cell length, 1 inch cell diameter (~0.23 standard cubic centimeters per minute (sccm) effective volume), and 0.4 s gas flow response time. The inlet tubing was collocated with an R. M. Young (Traverse City, MI, USA) 81000 sonic anemometer (R.M. Young Company, 2023) which measured micrometeorology at 10 Hz (Figure 3-1). The northward, eastward and vertical separation of the inlet tubing from the sonic anemometer was 0, 0, -10 cm, respectively. For AFG, CH<sub>4</sub> concentration data was collected at 2 and 4 m using two Aeris (Hayward, CA, USA) MIRA Ultra Series analyzers connected to tubing with a 3.275 mm inner diameter (Figure 3-2). As we had only one sonic anemometer, data from the sonic anemometer collocated with the MGGA were used for the AFG quantification. The two sampling points are 9.4 m apart.

~~2 and 4 m To calculate emissions using the two sampling inlets were mounted at 2 and 4 m heights on mast 2 and connected to the inlets of A. The analyzers were housed in a temperature-controlled unit and sampled at 5 Hz. Data from the 2-m analyzer were also used as input for the Gaussian Plume Inverse method analysis. To collect CH<sub>4</sub> concentration data for the eddy covariance method, the inlet tubing of the ABB (Zurich, Switzerland) GLA131 Series Microportable Greenhouse Gas Analyzer (MGGA) sampling at 10 Hz was collocated with an R. M. Young (Traverse City, MI, USA) 81000 sonic anemometer (R.M. Young Company, 2023) which measured micrometeorology at 10 Hz, 3 m height above ground level on mast 1 (Figure 3B).~~



**Figure 3: A:** Map illustration of major pieces of equipment and the measurements points at Colorado State University's Methane Emissions Technology Evaluation Center (METEC) in Fort Collins, CO, USA. Equipment 4S denotes horizontal separators, 4W are well heads, 4T are tanks, 5S are vertical separators and 5W are well heads. The number 1 is the measurement point for the Microportable Greenhouse Gas Analyzer for closed-path eddy covariance, Gaussian plume inverse and backward Lagrangian stochastic model quantification. The inlet tubing and the sonic anemometer are at 3 m height. The number 2 is the measurement point for the Aeris analyzers at 2 and 4 m heights for aerodynamic flux gradient sampling. The red dotted lines with yellow numbers show the average distances (meters) between emission equipment and measurement point. The orange number show the range of emission heights (meters) for each equipment. The analyzers were hosted in a temperature-controlled box. The two sampling points are 9.4 m apart.

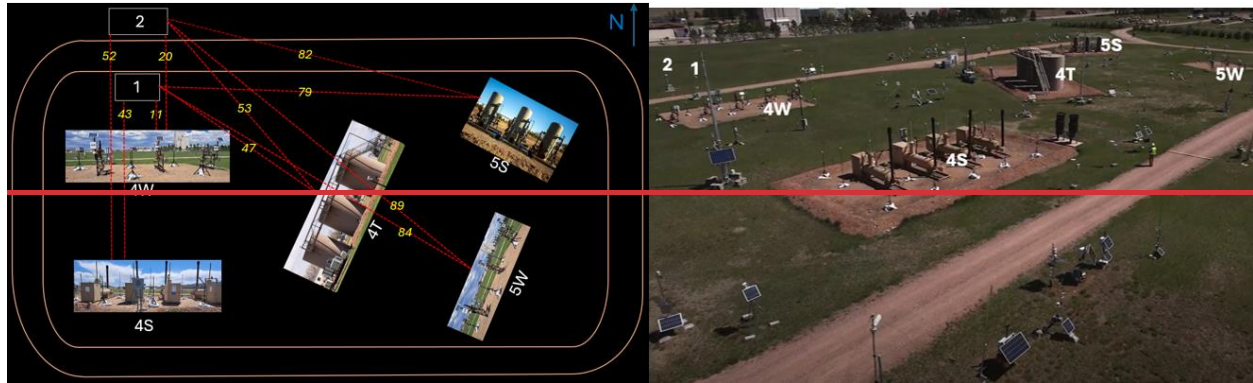


Figure 2: Left pane: Map illustration of major pieces of equipment and the measurements points at Colorado State University's Methane Emissions Technology Evaluation Center (METEC) in Fort Collins, CO, USA. 4S denotes the location of horizontal separators, 4W are well heads, 4T are tanks, 5S are vertical separators and 5W are well heads. 1 is the measurement point for the Microportable Greenhouse Gas Analyzer and 2 is the measurement point for the Aeris analyzers. The red dotted lines with yellow numbers show the average distances (meters) between emission equipment and measurement points. Right pane: Image of METEC showing relative heights of equipment ("METEC | Colorado State University," 2024).

To calculate emissions using the aerodynamic flux gradient approach, two sampling inlets were mounted at 2 and 4 m heights on mast 2 and connected to the inlets of two Aeris (Hayward, CA, USA) MIRA Ultra Series analyzers (Figure 3A). The analyzers were housed in a temperature-controlled unit and sampled at 5 Hz. Data from the 2 m analyzer were also used as input for the Gaussian Plume Inverse method analysis. To collect  $\text{CH}_4$  concentration data for the eddy covariance method, the inlet tubing of the ABB (Zurich, Switzerland) GLA131 Series Microportable Greenhouse Gas Analyzer (MGGA) sampling at 10 Hz was collocated with an R. M. Young (Traverse City, MI, USA) 81000 sonic anemometer (R.M. Young Company, 2023) which measured micrometeorology at 10 Hz, 3 m height above ground level on mast 1 (Figure 3B).

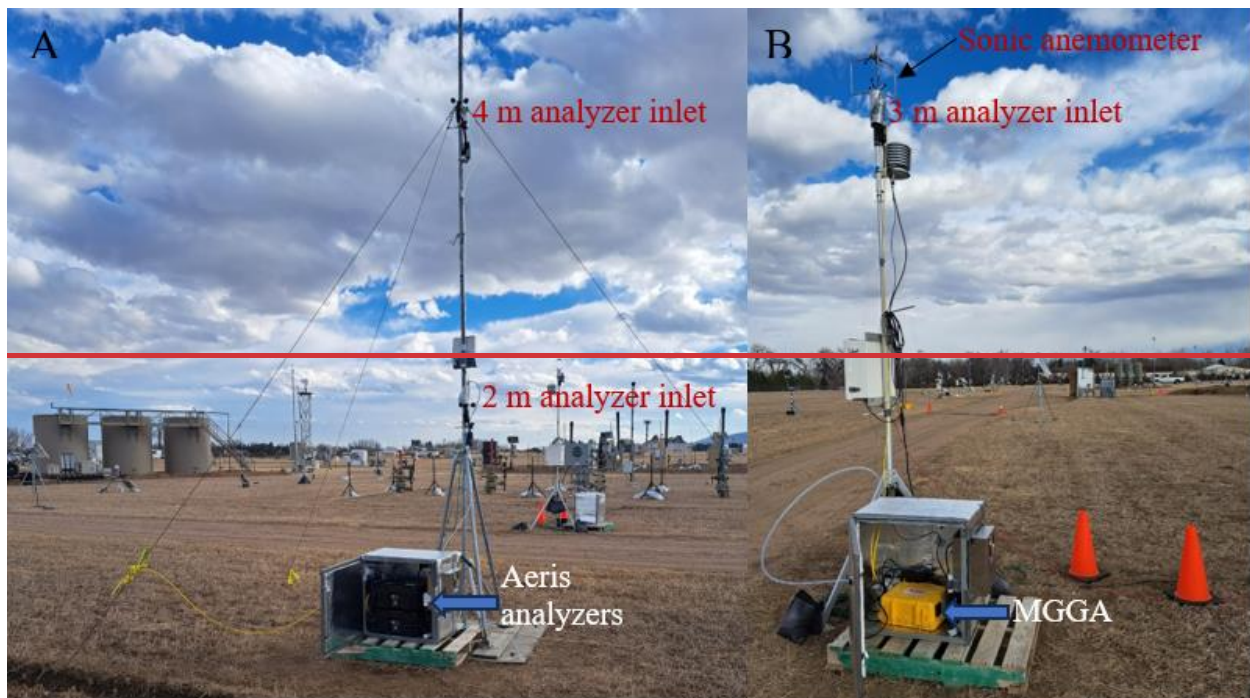


Figure 3: A is the aerodynamic flux gradient and Gaussian plume inverse sampling points and B is the eddy covariance sampling point. The two sampling points are 9.4 m apart.

## 2.2 Controlled Methane Releases

Controlled releases were part of the METEC Spring 2024 Advancing Development of Emissions Detection (ADED) Campaign conducted between February 6 and April 29, 2024 (Colorado State University, 2024). At METEC, natural gas of known  $\text{CH}_4$  content was released from above-ground emission points attached to equipment typically present in an oil and gas facility (tanks, separators and well pads). The gas release rates ranged between  $0.0195 \text{ kg h}^{-1}$  and  $8.75 \text{ kg h}^{-1}$ , and the release durations ranged from 10 seconds to 8 hours, simulating both fugitive and large emission events. The releases were run both during the day and night. The distance from the release points to the measurement points ranged between 9 and 94 m, and emission heights were between 0.14 and 64.9 m (Figure 23A). Emission points simulate the realistic size and locations of typical emissions from components such as the thief hatches, pressure relief valves, flanges, bradenheads, pressure transducers, Kimray valves and vents. The releases included both single-point emissions (single releases) and multi-point emission events (multiple simultaneous releases).

## 2.3 Calculation of Roughness Length

Surface roughness length ( $z_0$ ) was calculated from friction velocity (Supplementary Information Section 2a: Equations 1 and 2) by splitting the high frequency sonic anemometer data into 15-minute tables and filtering for those in neutral conditions,  $|L| > 500$  (Supplementary Information Section 2a: Equation 3). The overall roughness length selected as the median of all the calculated  $z_0$  was 0.1 m (Rey-Sanchez et al., 2022).

## 2.4 Models Quantification

### 2.4.1 Eddy Covariance

#### 2.4.1.1 Data pre-processing

Evaluating the MGGA CH<sub>4</sub> data showed that actual sampling was between 4 and 12 Hz frequency (highest sampling at 6 Hz), even though it had been set to sample at 10 Hz (Supplementary Information Section 2b). To account for this sampling variability, data were filtered to when sampling was equal/greater than 8 Hz. Data where the frequency were greater than 8 Hz were down sampled to 8 Hz. The sonic anemometer meteorological data (horizontal wind vectors (u, v), vertical wind vector (w), temperature (T), and pressure (P)) actual sampling varied between 7 and 9 Hz with the most frequent frequency at 8 Hz (Supplementary Information Section 2b). As the MGGA gas analyzer and sonic anemometer were not designed to clock synchronously, using the MGGA CH<sub>4</sub> clock time as a reference, meteorological data from the sonic anemometer were matched to the MGGA CH<sub>4</sub> data using linear interpolation to generate concentration-meteorological 8 Hz data.

The aggregated concentration-meteorological data were then merged with METEC's release data and metadata, and release event tables created. Release event tables were aggregated tables of concentration, meteorology and release (emission source location, duration and rate) information for all defined release events at METEC. The concentration-meteorological -release event data were then separated into single-release and multi-release events. Single-release events were when there was a single emission point at the site level, while multi-release events were when there was more than one emission point at the site level. The concentration-meteorological-release event tables were split into 5, 10 and 15-minute release event tables (i.e. there was a continuous release in the duration). Based on the bearing of the emission point to the measurement point and the average wind direction in the duration, the data was further filtered to downwind data,  $\pm 5^\circ$ ,  $\pm 10^\circ$ ,  $\pm 20^\circ$ , and  $\pm 45^\circ$ .

#### 2.4.1.2 Flux calculation

Turbulent fluxes were calculated using the open software EddyPro® version 7 (LI-COR, Nebraska, USA, n.d.). Acquisition frequency was set at 8 Hz, while file duration and the flux interval were set at 5, 10, and 15 minutes, respectively, depending on the file being processed. Table 1 shows the instruments input to the software.

Table 1. Anemometer and Gas Analyzer Input into EddyPro

| <u>Anemometer</u>               |                  | <u>Gas Analyzer</u>             |                            |
|---------------------------------|------------------|---------------------------------|----------------------------|
| <u>Information</u>              |                  | <u>Information</u>              |                            |
| <u>Manufacturer</u>             | <u>Young</u>     | <u>Manufacturer</u>             | <u>Other</u>               |
| <u>Model</u>                    | <u>81000</u>     | <u>Model</u>                    | <u>Generic closed path</u> |
| <u>Height</u>                   | <u>3 m</u>       | <u>Tube length</u>              | <u>300 cm</u>              |
| <u>Wind data format</u>         | <u>u, v, w</u>   | <u>Tube inner diameter</u>      | <u>3.275 mm</u>            |
| <u>North alignment</u>          |                  | <u>Nominal tube flow rate</u>   | <u>3.2 l/m</u>             |
| <u>North off-set</u>            | <u>0.0</u>       | <u>Northward separation</u>     | <u>0.00 cm</u>             |
| <u>Northward separation</u>     | <u>Reference</u> | <u>Eastward separation</u>      | <u>0.00 cm</u>             |
| <u>Eastward separation</u>      | <u>Reference</u> | <u>Vertical separation</u>      | <u>-10.00 cm</u>           |
| <u>Vertical separation</u>      | <u>Reference</u> | <u>Longitudinal path length</u> | <u>10.00 cm</u>            |
| <u>Longitudinal path length</u> |                  | <u>Transversal path length</u>  | <u>2.54 cm</u>             |
| <u>Transversal path length</u>  |                  | <u>Time response</u>            | <u>0.4 s</u>               |

In raw data processing, axis rotations for tilt correction under wind speed measurement offsets was checked. Under turbulent fluctuations, double rotation and block average detrend methods were used. Covariance maximization with default was used for time lag detection; time lags detection was checked. Compensation for density fluctuations (Webb-Pearman-Leuning terms) was unchecked as the MGGA analyzer synchronously reported dry CH<sub>4</sub> and water mole fractions, cell temperature and pressure. Mauder and Foken (2004) (0-1-2 system) were used for quality check. All statistical tests for raw data screening, Vickers and Mahrt (1997)– spike count/removal, amplitude resolution, drop-outs, absolute limits, skewness and kurtosis, discontinuities, time lags, angle of attack and steadiness of horizontal wind were checked. The default values for all these tests were used. Similarly, default settings for spectral analysis and corrections were used. Analytic correction of high-pass filtering effects (Moncrieff et al., 2005) for low frequency range; and correction of low-pass filtering effects (Fratini et al., 2012 - In situ analytic) and instruments separation (Horst and Lenschow, 2009 - only crosswind and vertical) in the high frequency range were used.

#### 2.4.1.3 Post-processing

During post-processing, flux data were filtered based on (1) quality flags, Mauder and Foken (2004) (0-1-2 system), and (2) surface friction velocity ( $u^* > 0.13$  m/s). Data that were flagged “2” were first filtered out as they were considered poor quality fluxes (LICOR, 2025), and the remaining dataset were filtered for high turbulence data. All data were filtered out as low quality and no further post-processing was done.

#### 2.3 Data Processing

Methane concentrations data from the analyzers were aggregated with the meteorological data from the sonic anemometer. For aerodynamic flux gradient and Gaussian plume inverse method data were averaged to 1 Hz, for the eddy covariance the raw CH<sub>4</sub> 10 Hz data was used. The aggregated meteorological concentration data were then merged with METEC’s release data and metadata, and event tables created. The meteorological concentration release event data were then separated into single point and multi point events. The event tables were split into 20 minute

emission events for aerodynamic flux gradient and Gaussian plume inverse method as they are dependent on atmospheric stability that is typically determined in time durations of 15 to 30 minutes. Shorter duration measurements (i.e. <15 minutes) may not represent the mean atmospheric state, while longer periods (> 30 minutes) may cause errors especially during rapid transitions in weather conditions (Crenna, 2006). 30 minute events were used for eddy covariance processing following published typical averaging times of eddy covariance measurements (Nemitz et al., 2018), and its quantification is assumed to be independent of atmospheric stability (Denmead, 2008). For eddy covariance and aerodynamic flux gradient, Monin Obukhov length ( $L$ ) was calculated as the measure of atmospheric stability for every 20 or 30 minute time period, depending on the method, using output from the sonic anemometer.  $L$  was calculated from the surface friction velocity ( $u_*$ ,  $\text{m s}^{-1}$ ), mean potential temperature ( $\theta$ , K), von Kármán's constant ( $k$ , 0.41), gravitational acceleration ( $g$ ,  $9.8 \text{ m s}^{-2}$ ) and the surface (kinematic) turbulent flux of sensible heat  $w'\theta'$  (Eq. 1 and 2) (Kljun et al., 2015; Stull, 1988).

$$L = -\frac{u_*^3 \theta}{k_\theta g w' \theta'} \quad (1)$$

$$u_* = \left[ (\overline{u'w'})^2 + (\overline{v'w'})^2 \right]^{1/4} \quad (2)$$

For the Gaussian method, atmospheric stability was calculated based on the EPA standard operating procedure for point source Gaussian method (US EPA, 2013). The average local wind stability class ( $pgi$ ) was calculated as the average of atmospheric stability determined using the standard deviation of the wind direction, and the stability calculated from turbulent intensity (ratio of the standard deviation of the wind speed to the average wind speed). The dispersion coefficients used for Gaussian quantification were extracted from the EPA operating procedure that provided coefficients for distances ranging from 1 to 200 m from source (US EPA, 2013).

The wind direction ( $WD$ ) and speed ( $WS$ ) were calculated from the wind vectors  $u$  and  $v$ , based on the manufacturer's configuration:  $+u$  values = wind from the east,  $+v$  values = wind from the north, and  $+w$  = updraft (Eq. 3 and 4).

$$WD = \text{mod}(90 - \text{atan2d}(v, u), 360) \quad (3)$$

$$WS = \sqrt{u^2 + v^2} \quad (4)$$

The bearing of each release point to the masts' location was calculated using the latitudes and longitudes of the release points provided in the METEC metadata. This bearing was used to determine when the masts were downwind of the release points during the 20/30 minute period. The models' quantification accuracies were tested in three downwind ranges:  $\pm 10^\circ$ ,  $\pm 20^\circ$ , and  $\pm 30^\circ$ . A mast was considered downwind when the wind direction was within the specified range for 30% of the 20/30 minute duration. Results for the 20-degree range are presented in the Results section, while the 10 and 30 degree results are included in the Supplementary Material. The 30% threshold was chosen to ensure sufficient data points for evaluating the models. The data were categorized into single release single emission (single emission at the site and the mast was downwind of the release point), multi-release single emission (multiple emissions

at the site level, but the mast was downwind of a single release point), and multi-release multi-emission (multiple emissions at the site level, but the mast was downwind of more than one release point).

## 2.4 Methane Emissions Quantification

### 2.4.1 Background Concentration

Background concentration was determined for each of the sensors to calculate CH<sub>4</sub> enhancement. Due to inherent variation in sensors that were used in this study, CH<sub>4</sub> background was calculated for each sensor separately. CH<sub>4</sub> background was calculated as the average of the lowest 5<sup>th</sup> percentile of all continuous concentration readings (US EPA, 2013). Methane enhancement was determined as CH<sub>4</sub> concentration measurement minus the background concentration measurement.

### 2.4.2 Eddy Covariance

Emissions were quantified using the eddy covariance method for all three emissions scenarios (single release single emission, multi-release single emission and multi-release multi-emission). Methane flux ( $F$ , kg m<sup>-2</sup> s<sup>-1</sup>) was calculated as the covariance between the vertical wind speed ( $w$ , m s<sup>-1</sup>) and CH<sub>4</sub> enhancement ( $c$ , g m<sup>-3</sup>) over 30 minutes (Eq. 5; Denmead, 2008).

$$F = w'c' \quad (5)$$

### 2.4.2.3 Aerodynamic Flux Gradient

Methane concentration data from the 2 and 4 m analyzers and meteorology data from the sonic anemometer were averaged to 1 Hz and then aggregated. Similarly to EC pre-processing, the aggregated concentration-meteorological data were merged with METEC's release data and metadata, and release event tables created. The concentration-meteorological-release event data were then separated into single-release and multi-release events. For single-release events, the concentration-meteorological-release event tables were split into 5, 10 and 15-minute release event tables. Based on the bearing of the emission point to the measurement point and the average wind direction in the duration, the data was further filtered to downwind data,  $\pm 5^\circ$ ,  $\pm 10^\circ$ ,  $\pm 20^\circ$  and  $\pm 45^\circ$ . Multi-release events were further classified into multi-release single-point emissions (i.e., there were multiple emissions at the site level, but the mast was downwind of a single source) and multi-release multi-point emissions (i.e. there were multiple emissions at the site level and the mast was downwind of more than one source). As we were limited to data from a single sonic anemometer for footprint calculation, we only calculated emissions when the mast was downwind of a single source in single-release single-point emission and multi-release single-point scenarios. Flux determination and measurement footprint calculation are discussed in section 2.4.3.

Aerodynamic flux gradient quantification was also tested in all three cases. Methane flux ( $F$ , kg m<sup>-2</sup> s<sup>-1</sup>) were then calculated using the AFG equation based on surface friction velocity ( $u_*$ , m s<sup>-1</sup>), von Kármán's constant ( $k_s$ , 0.41), the difference in the average CH<sub>4</sub> enhancement between the higher and lower height ( $g$ , m<sup>-3</sup>), natural log of the higher and lower height, and stability correction factors  $\Psi$  (Supplementary Information Equation Eq. 6; Denmead, 2008; Kamp et al., 2020).

#### 2.4.4 Determining the Area of Vertical Flux Contribution

Eddy covariance and aerodynamic flux gradient measurements at a point  $(0, 0, z)$  generate vertical fluxes in  $\text{kg m}^{-2}\text{s}^{-1}$ . In this study, these fluxes represent emissions from single point or multi point sources distributed over an area ( $\text{m}^2$ ).

#### 2.4.3 Footprints Calculation

Eddy covariance and AFG footprints were calculated using the Kljun et al. (2015) footprint model. Even though Rey-Sanchez et al. (2022) reported the Kljun et al. (2015) footprint model to be less accurate compared to the Kormann and Meixner (2001), Kormann and Meixner (2001) was too complex for our study because it required multiple sonic anemometers or tracer release experiments to calculate the exponential wind velocity power law, and the exponential eddy diffusivity power law for site specific data.

Our study was limited to a single sonic anemometer, and this provided enough inputs for the Kljun et al. (2015) footprint model. The default pixel size  $2 * 2 \text{ m}$  was used in this study, and determine the area that contributed 80% ( $r = 80, 10 \leq r \leq 90$ ) of the vertical flux measured by the eddy covariance and aerodynamic flux gradient systems. In previous studies, 80% footprints have been used due to the difficulty of reproducing 90% of the sources under neutral and stable conditions, where footprints tend to be long. The difference between the 80% and 90% contours is typically excessively large, despite minimal flux contributions in that area (Rey-Sanchez et al., 2022). The Kljun et al. (2015) model calculates footprint as a function of effective height ( $z_m$  = sensor height ( $z$ ) — displacement height ( $m$ )), roughness length ( $z_o, m$ ) / mean wind speed ( $u_{mean}, \text{m s}^{-1}$  — used in this study), height of the boundary layer ( $h, m$ ), Obukhov length ( $L, m$ ), standard deviation of the lateral velocity ( $\sigma_v, \text{m s}^{-1}$ ), and friction velocity ( $u^*, \text{m s}^{-1}$ ) (Kljun et al., 2015). The roughness sublayer in the model was set to 1 (footprint is calculated even if  $z_m$  is within the roughness layer). The area of vertical flux contribution was calculated as the polygon area covered by the contour. Due to the limitations of the flux footprint model for the measurement height and stability (Kljun et al., 2015), 20/30 minute files flagged by the footprint model when  $z_m/L < 15.5$ , were excluded from further analysis. This study first calculated the area that contributed 90% of the vertical flux; and based on the location of the point source, the source was determined if it was within the 90% footprint area. Point source emissions of sources within this region were then calculated based on the approach by Dumortier et al. (2019). This approach assumes all measured flux is equal to flux resulting from a single point source. In case of the mast being downwind of more than one source, more sonic anemometers are needed to solve the two unknown point source fluxes.

#### 2.4.4.5 Gaussian Plume Inverse Method

##### 2.4.4.1 Data pre-processing

Methane concentration data from the MGGA analyzer and meteorology data from the sonic anemometer were averaged to 1 Hz and pre-processed similarly to the AFG method. For continuous monitoring sensors, background concentration can be determined from  $\text{CH}_4$  concentrations measured by a sensor upwind of the emission source, or by sampling when the wind is blowing away from the source. However, for continuous monitoring sensors, using an upwind sensor has the limitation of missing downwind background noise resulting from emissions in the

preceding emission event where there is residual CH<sub>4</sub> in air especially during stable conditions, and capturing sensor drift in the downwind sensor. In this study, background CH<sub>4</sub> was calculated as the average of the lowest 5<sup>th</sup> percentile, 5 minutes before each release started. In cases where this background was greater than the mean CH<sub>4</sub> concentration in the quantifying duration, the minimum CH<sub>4</sub> concentration for that duration was used as the background. Methane enhancement was then calculated as CH<sub>4</sub> concentration minus the background.

#### 2.4.4.2 Quantification

For single-release tables, the measurement point was downwind of a single source (single-release single-point emission), hence the tables were quantified as they were using the standard GPIM equation (Supplementary Information Section 2a: Equation 7). However, for multi-release events, the tables were further processed as the GPIM method is designed to quantify a single point source at a time. For multi-release events, the number of emission points in the downwind tables were used to further classify the tables into multi-release single-point emissions (i.e. there were multiple emissions at the site level, but the mast was downwind of a single source), and multi-release multi-point emissions (i.e. there were multiple emissions at the site level and the mast was downwind of more than one emission source). The GPIM method was only used for the multi-release single-point emissions.

The Gaussian plume inverse method was used to quantify single release single emission and multi-release single emission. The quantified emission ( $Q$ , kg h<sup>-1</sup>) was calculated from the CH<sub>4</sub> enhancement ( $X$ , g m<sup>-3</sup>), wind speed ( $u$ , m s<sup>-1</sup>), horizontal dispersion coefficient ( $\sigma_y$ , m), vertical dispersion coefficient ( $\sigma_z$ , m), crosswind distance ( $y$ , m), sampling height ( $z$ , m), emission height ( $h_s$ , m), and the height of the boundary layer (Equation 7; Riddick et al., 2022b).

$$X(x, y, z) = \frac{Q}{2\pi u \sigma_y \sigma_z} e^{-\frac{y^2}{2\sigma_y^2}} \left( e^{-\frac{(z-h_s)^2}{2\sigma_z^2}} + e^{-\frac{(z+h_s)^2}{2\sigma_z^2}} + e^{-\frac{(z-2h+h_s)^2}{2\sigma_z^2}} + e^{-\frac{(z+2h-h_s)^2}{2\sigma_z^2}} + e^{-\frac{(z-2h-h_s)^2}{2\sigma_z^2}} \right) \quad (7)$$

#### 2.4.5 Backward Lagrangian stochastic model

Pre-processed data from the GPIM method was used for bLs quantification. Quantification was done using the open-source software WindTrax 2.0 (Crenna, 2006; WindTrax 2.0, n.d.). For every 5-, 10- and 15-minute duration in the  $\pm 5^\circ$ ,  $\pm 10^\circ$ , and  $\pm 20^\circ$ , respectively, inputs included roughness length ( $z_0$ ), Monin-Obukhov length ( $L$ ), mean (wind speed, wind direction, concentration, pressure, temperature), background concentration, source height, and distance from the emission point to sensor. WindTrax is also designed to quantify a single point source at a time, and hence, was only used to quantify single-point single emissions and multi-point single emissions.

### 3 Results

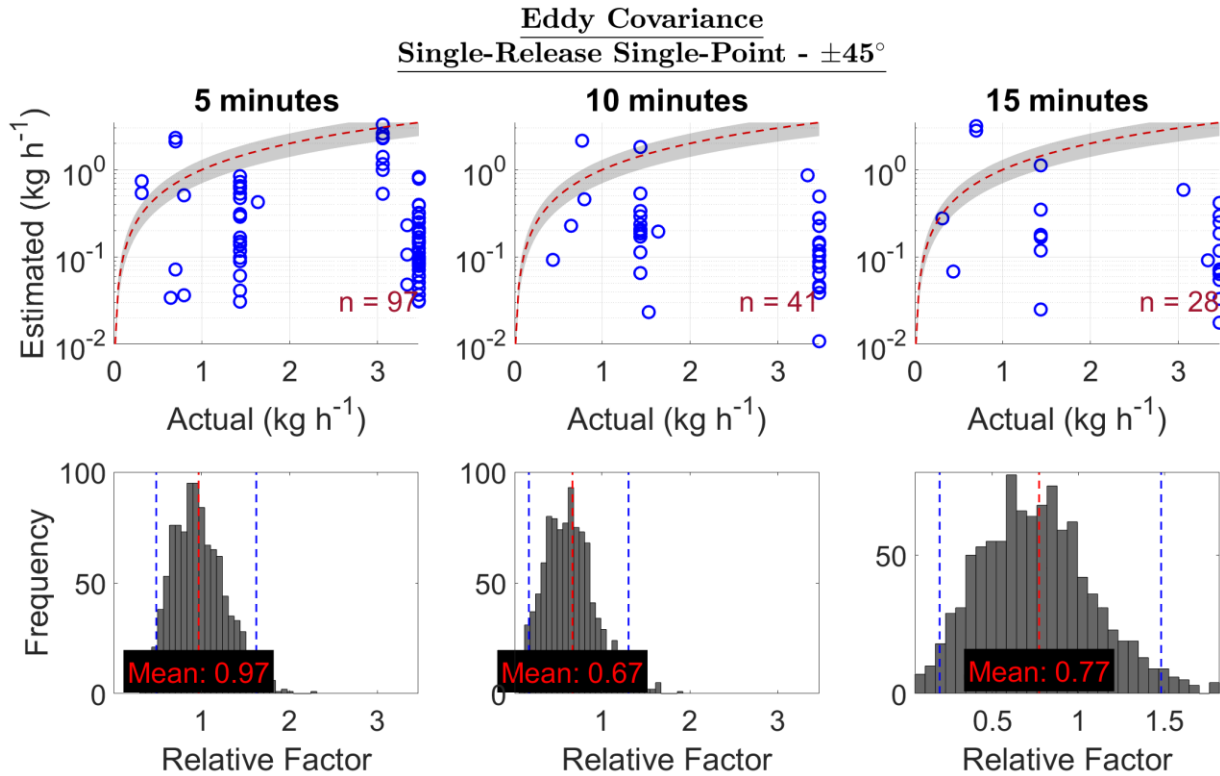
#### 3.1 Methane Emission Quantification

##### 3.1.1 Eddy Covariance

##### 3.1.1 Single-Release Single-Point

For single-release single-point (SRSP) emissions, the closed-path EC quantified emissions correctly within a mean relative factor (MRF) of between 0.67 and 0.97 at  $\pm 45^\circ$  wind sector range (Figure 4). The MRF was 0.97, 0.67 and 0.77 for a 97, 41 and 28 sample size at an averaging period of 5, 10 and 15 minutes, respectively (Figure 4). At

$\pm 5^\circ$  wind sector range, the sample size was 1, 2 and 3 at 5, 10 and 15-minutes averaging periods, hence, no reasonable quantification results (Supplementary Information Section 3.1.1). At  $\pm 10^\circ$  wind sector range, the MRF was 0.60, 0.95, and 1.71 for a 15, 6 and 6 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.1.1). At  $\pm 20^\circ$  wind sector range, the MRF was 0.53, 0.86, and 1.10 for a 40, 16 and 14 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.1.1).



**Figure 4.** Top plot: Estimated emission vs actual emission (kg h<sup>-1</sup>) for a single-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is n. Bottom plot: A bootstrap of mean relative factor (MRF: estimated emissions divided by actual controlled emission) for a single-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.

### 3.1.2 Multi-Release Single-Point

For multi-release single-point (MRSP) emissions, the closed-path EC quantified emissions correctly within an MRF of between 1.02 and 2.43 at  $\pm 45^\circ$  wind sector range (Figure 5). The MRF was 1.02, 2.43 and 1.88, for a 355, 183, and 110 sample size at an averaging period of 5, 10 and 15 minutes, respectively (Figure 5). At  $\pm 5^\circ$  wind sector range, the MRF was 1.68, 5.21 and 5.16 for a 61, 34 and 23 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.1.2). At  $\pm 10^\circ$  wind sector range, the MRF was 2.75, 3.32, and 4.11 for a 124, 70 and 44 sample size, averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.1.2). At  $\pm 20^\circ$  wind sector range, the MRF 2.08, 2.89 and 2.70 for a 284, 143 and 80 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.1.2).

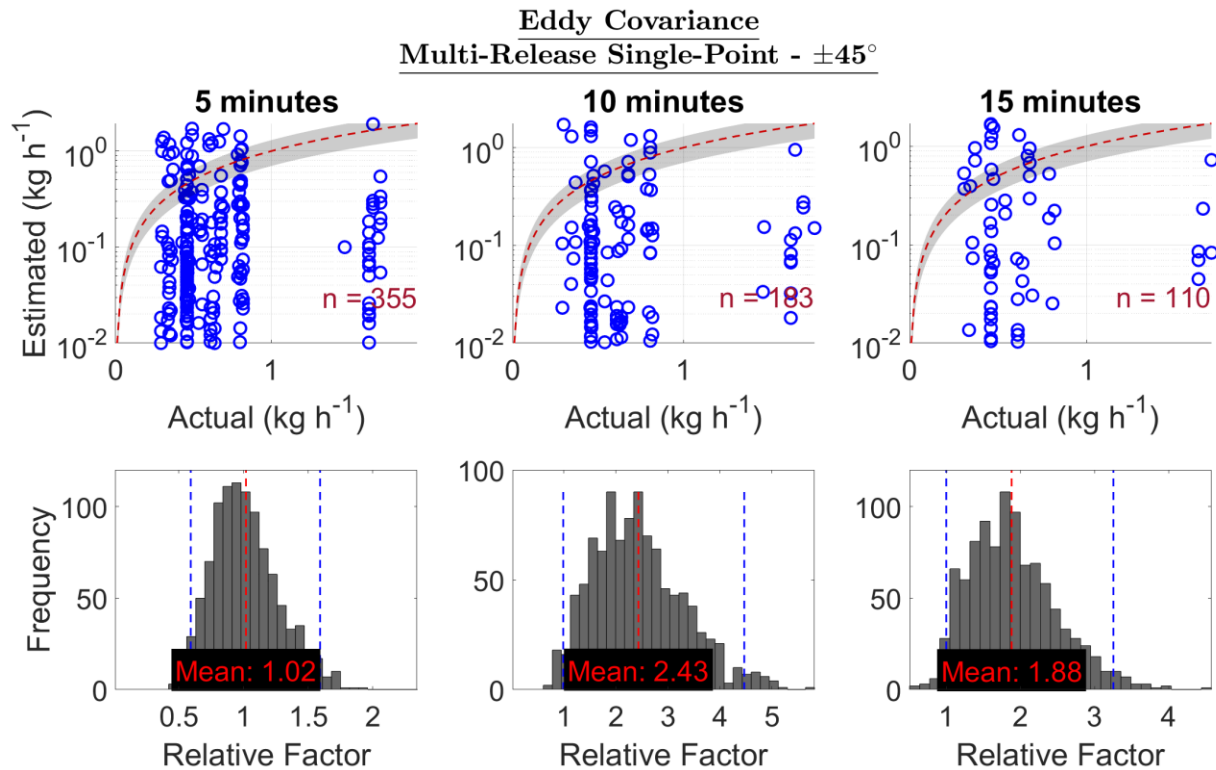


Figure 5. Top plot: Estimated emission vs actual emission ( $\text{kg h}^{-1}$ ) for a multi-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is n. Bottom plot: A bootstrap of mean relative factor (MRF: estimated emissions divided by actual controlled emission) for a multi-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.

~~For stable, continuous 30 minute release events, emissions calculated using the eddy covariance method were an underestimate for single release single emission, multi release single emission and multi release multi emission events (Figure 4). All data points were below the 1:1 line. A plot of the quantified emission versus controlled release ( $\text{kg h}^{-1}$ ) did not show a linear correlation ( $R^2$  between 0.03 and 0.36), as all emissions were largely underestimated. The eddy covariance method reported emissions of between 0 and  $0.5 \text{ kg h}^{-1}$  overall, despite actual emissions being between 0 and about  $7 \text{ kg h}^{-1}$  (Figure 4). The underestimation was consistent across all downwind ranges, 10, 20 and 30 degrees (Supplementary Material Section 2.1).~~

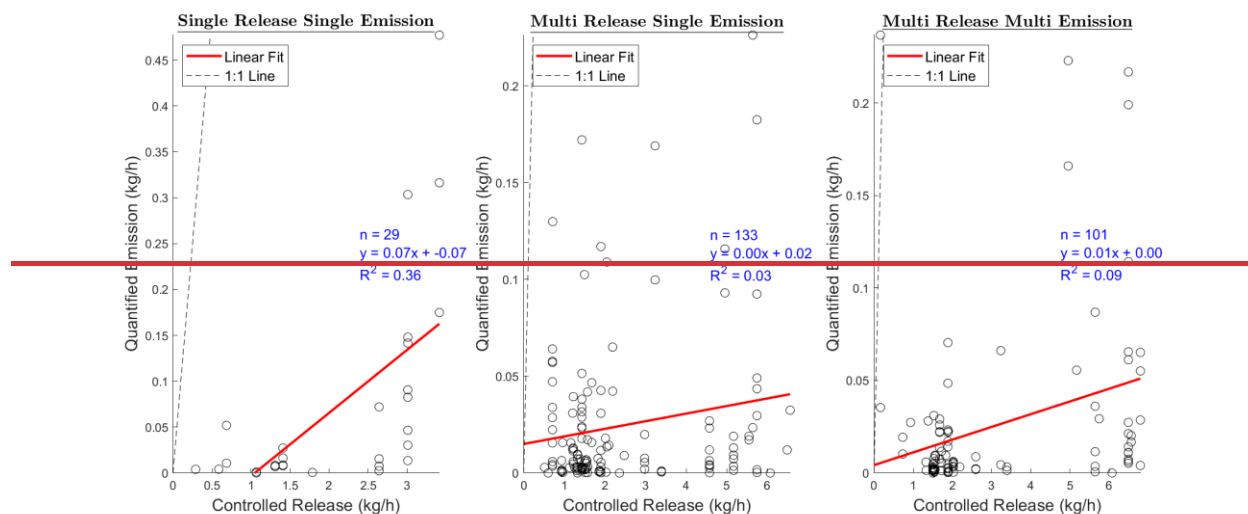


Figure 4: Quantified emission calculated using the eddy covariance method. Left pane shows a scatter plot of quantified emission versus total controlled release for a single release at the site level and the mast was downwind of the release point. Center pane shows a scatter plot of quantified emission versus total controlled release for multiple releases at the site level, but the mast was downwind of a single release point. Right pane shows a scatter plot of quantified emission versus total controlled release for multiple releases at the site level and the mast was downwind of more than one release point. The dashed line represents the 1:1 line (points below the line were underestimated); the red line is the linear regression fit of the data, and n is the number of data points.

### 3.1.2 Aerodynamic Flux Gradient

#### 3.2.1 Single-Release Single-Point

For SRSP emissions, the MRF for AFG was 1.41, 1.67 and 1.30 at  $\pm 45^\circ$  wind sector range, for a 112, 56 and 34 sample size at an averaging period of 5, 10 and 15 minutes (Figure 6). At  $\pm 5^\circ$  wind sector range, the sample size was 3 at 5-minutes and 0 at 10 and 15-minutes averaging periods, hence, no reasonable quantification results (Supplementary Information Section 3.2.1). Similarly, at  $\pm 10^\circ$  wind sector range, the sample size was 7, 1 and 2, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.2.1). For the  $\pm 20^\circ$  wind sector range, the MRF was 0.75, 0.48 and 1.58 for a sample size of 26, 8 and 4, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.2.1). These results show that close-to-stable MRF for ARF is achieved over a wide sector range,  $\pm 45^\circ$ .

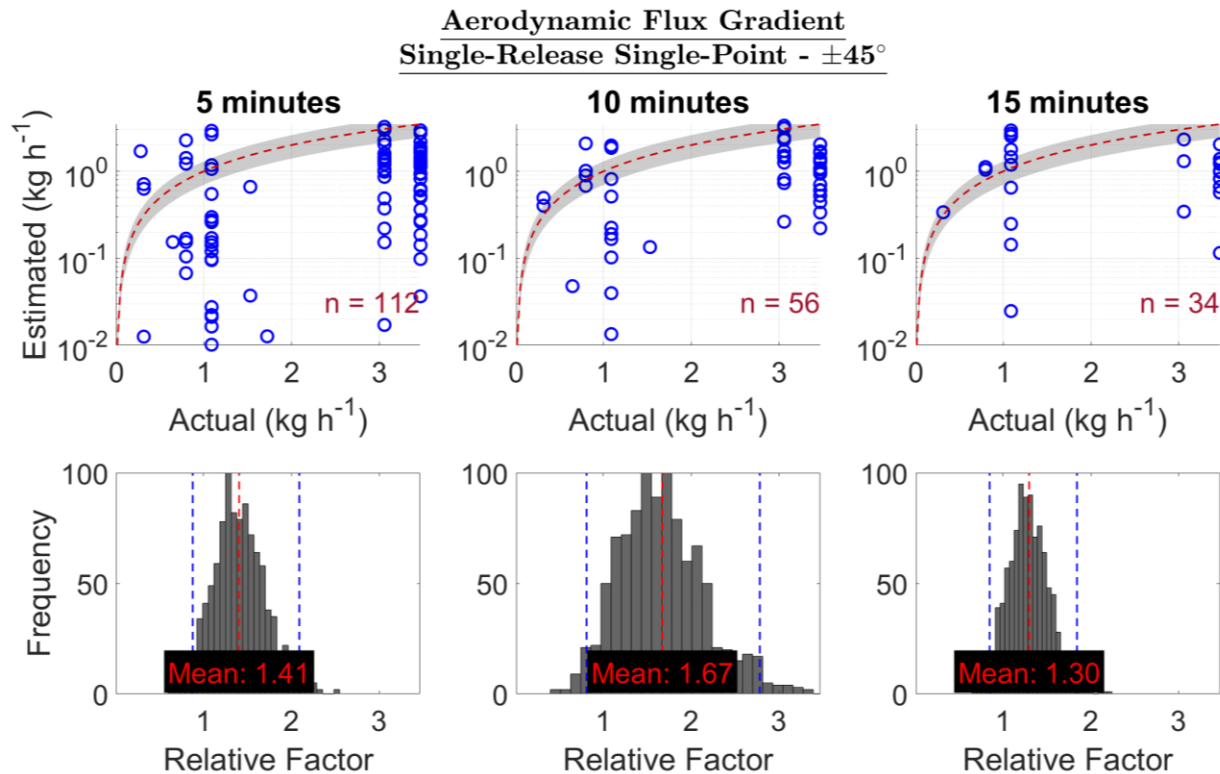


Figure 6. Top plot: Estimated emission vs actual emission ( $\text{kg h}^{-1}$ ) for a multi-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is  $n$ . Bottom plot: A bootstrap of mean relative factor (MRF; estimated emissions divided by actual controlled emission) for a multi-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.

### 3.2.2 Multi-Release Single-Point

For MRSP emissions, generally, the MRF for AFG was between 2 and 9 for all wind sector ranges and averaging periods. The MRF was 3.32, 2.40 and 2.48 at  $\pm 45^\circ$  wind sector range, for a 278, 146 and 94 sample size at an averaging period of 5, 10 and 15 minutes (Figure 7). At  $\pm 5^\circ$  wind sector range, the MRF was 8.84, 2.51 and 2.93 for a 36, 20 and 13 sample size at 5, 10 and 15-minutes averaging periods, respectively (Supplementary Information Section 3.2.2). At  $\pm 10^\circ$  wind sector range, the MRF was 6.12, 2.29 and 2.61 for a 76, 40 and 26 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.2.2). For the  $\pm 20^\circ$  wind sector range, the MRF was 5.16, 2.24 and 4.69 for a 142, 74, and 42 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.2.2). These results show that close-to-stable MRF for ARF is achieved for longer-averaging (10 to 15 minutes) and wide sector ranges of  $\pm 45^\circ$ .

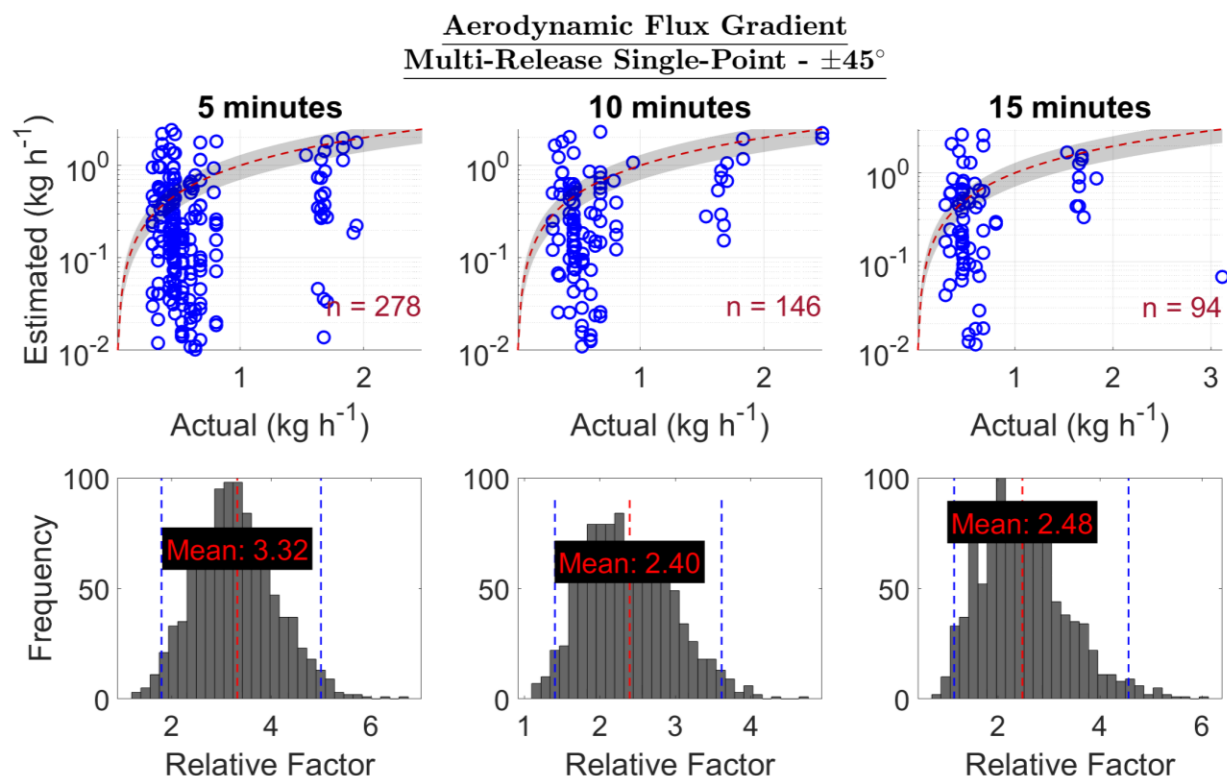


Figure 7. Top plot: Estimated emission vs actual emission ( $\text{kg h}^{-1}$ ) for a multi-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is  $n$ . Bottom plot: A bootstrap of mean relative factor (MRF; estimated emissions divided by actual controlled emission) for a multi-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.

### 3.4 Gaussian Plume Inverse Method

#### 3.4.1 Single-Release Single-Point

Generally, the GPIM quantified SRSP emissions between an MRF of 1.85 and 443.54 for all wind sector ranges and averaging periods. The MRF was 2.58, 2.37 and 2.63 for a 79, 41, and 27 sample size, at 5, 10 and 15-minutes averaging period, respectively at  $\pm 10^\circ$  wind sector range (Figure 8). At  $\pm 5^\circ$  wind sector range, the MRF was 2.26, 1.85 and 3.04 for a sample size of 31, 22 and 17 at 5, 10 and 15-minutes averaging period; and at  $\pm 20^\circ$  wind sector range, the MRF was 443.54, 3.36 and 3.14 for 165, 88, and 57 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.3.1). The GPIM MRF is more stable at narrow wind-sector ranges, and over a long averaging period, 10 to 15 minutes.

**Gaussian Plume Inverse Method**  
**Single-Release Single-Point -  $\pm 10^\circ$**

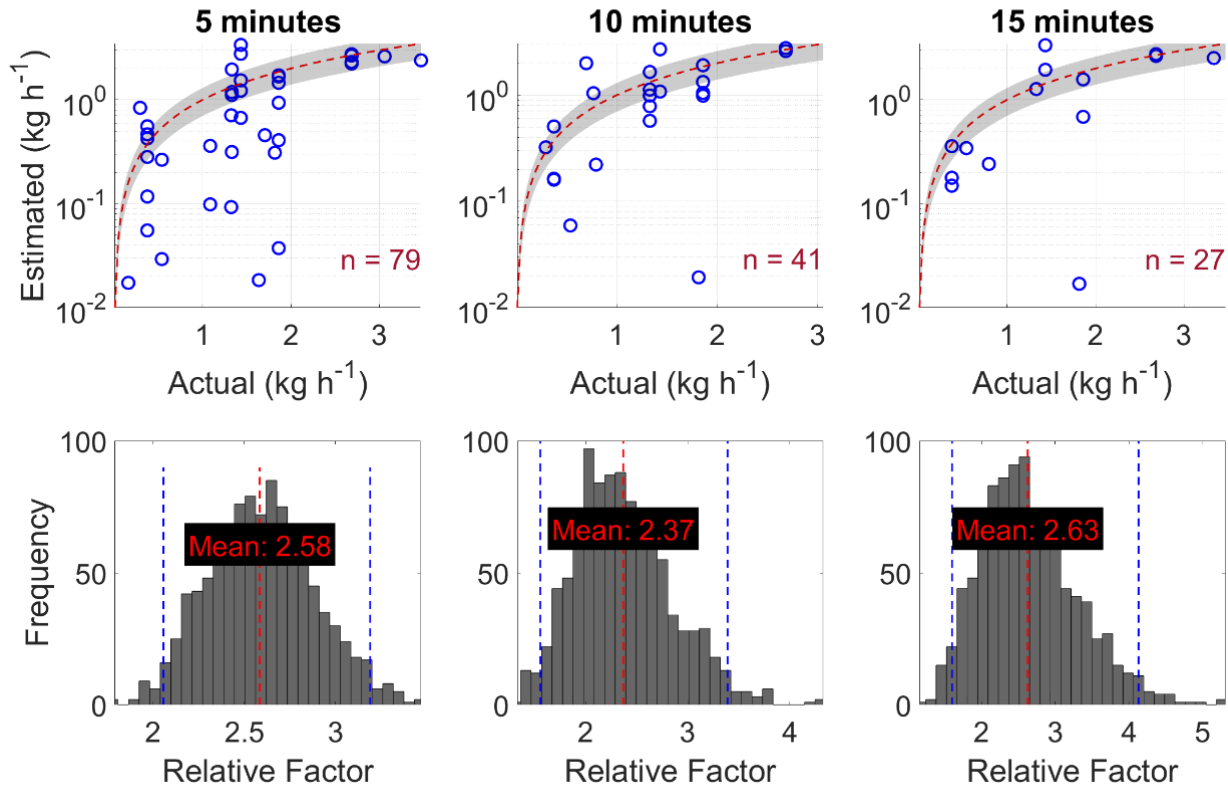


Figure 8. Top plot: Estimated emission vs actual emission ( $\text{kg h}^{-1}$ ) for a multi-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is  $n$ . Bottom plot: A bootstrap of mean relative factor (MRF: estimated emissions divided by actual controlled emission) for a multi-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.

### 3.4.2 Multi-Release Single-Point

The MRF GPIM results for MRSP emissions were between 15.72 and 29.01 for all wind sector ranges and averaging periods. The MRF was 15.72, 24.95 and 16.97 for an 827, 430, and 256 sample size, at 5, 10 and 15-minutes averaging period, respectively at  $\pm 10^\circ$  wind sector range (Figure 9). At  $\pm 5^\circ$  wind sector range, the MRF was 26.77, 25.04 and 29.01 for a sample size of 398, 189 and 132 sample size at 5, 10 and 15-minutes averaging period; and at  $\pm 20^\circ$  wind sector range, the MRF was 18.15, 23.07 and 19.96 for a 1273, 656, and 407 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.3.2). Generally, the GPIM overestimated MSRP emissions by up to a magnitude of 20.

**Gaussian Plume Inverse Method**  
**Multi-Release Single-Point -  $\pm 10^\circ$**

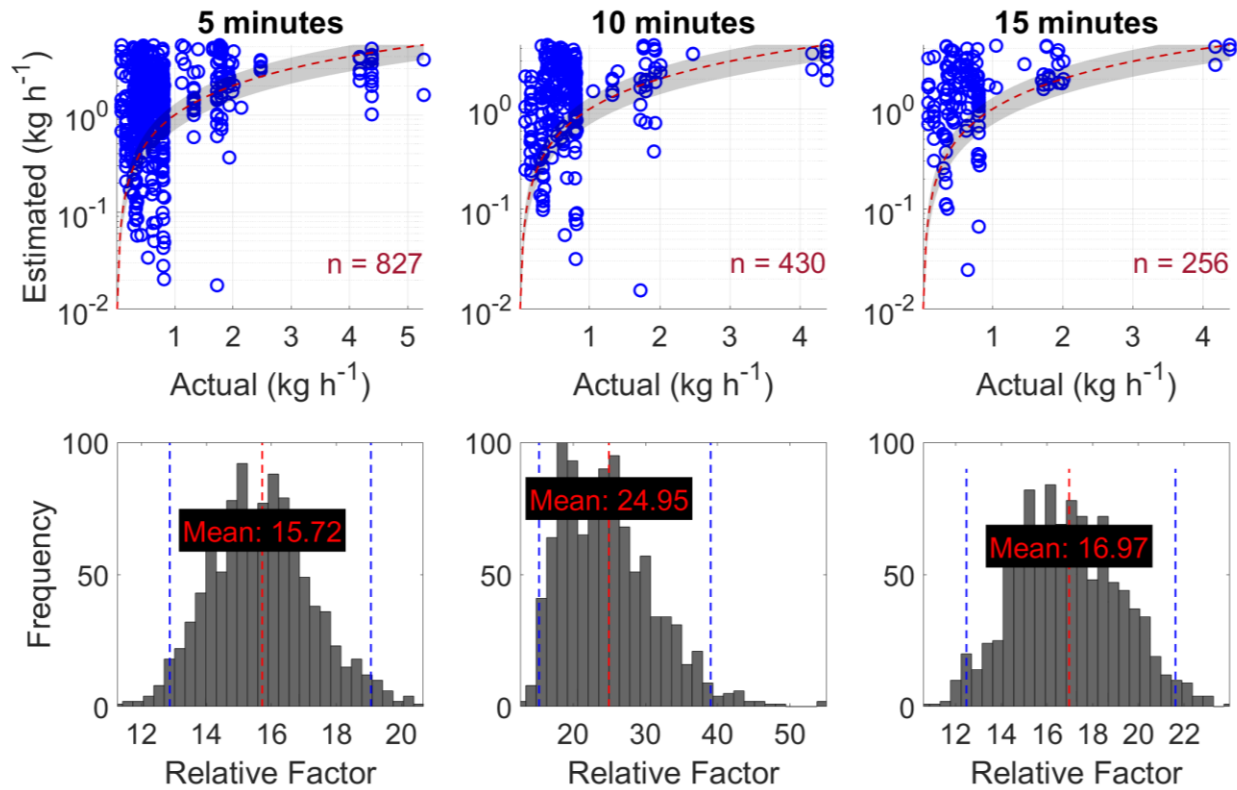


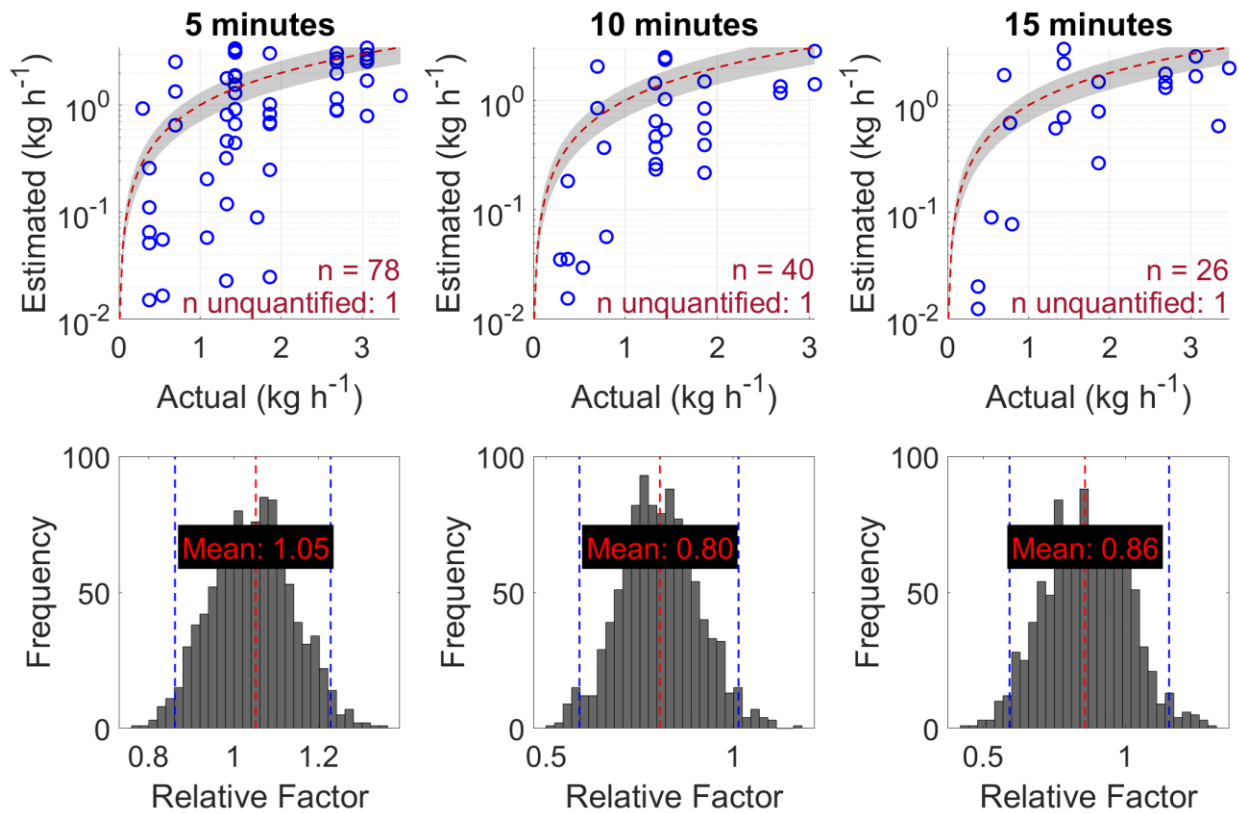
Figure 9. Top plot: Estimated emission vs actual emission (kg h<sup>-1</sup>) for a multi-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is n. Bottom plot: A bootstrap of mean relative factor (MRF: estimated emissions divided by actual controlled emission) for a multi-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.

### 3.4 Backward Lagrangian Stochastic Model

#### 3.4.1 Single-Release Single-Point

For SRSP emissions, the bLs method estimated emissions between 0.68 and 1.34 MRF. At  $\pm 10^\circ$  wind sector range, the MRF was 1.05, 0.80 and 0.86 for 78, 40 and 26 sample size, at 5, 10 and 15-minutes averaging period, respectively (Figure 10). At  $\pm 5^\circ$  wind sector range, the MRF was 0.72, 0.68 and 0.82 for a sample size of 31, 22 and 17 at 5, 10 and 15-minutes averaging period; and at  $\pm 20^\circ$  wind sector range, the MRF was 1.34, 1.34 and 1.33 for a 131, 70 and 49 sample size, at averaging periods of 5, 10 and 15-minutes, respectively (Supplementary Information Section 3.4.1). Comparing bLs method to GPIM for SRSP emissions as they are both point-source methods, the MRF for bLs was closer 1 indicating higher possibility of correct quantification.

**Backward Lagrangian Stochastic Model**  
**Single-Release Single-Point -  $\pm 10^\circ$**

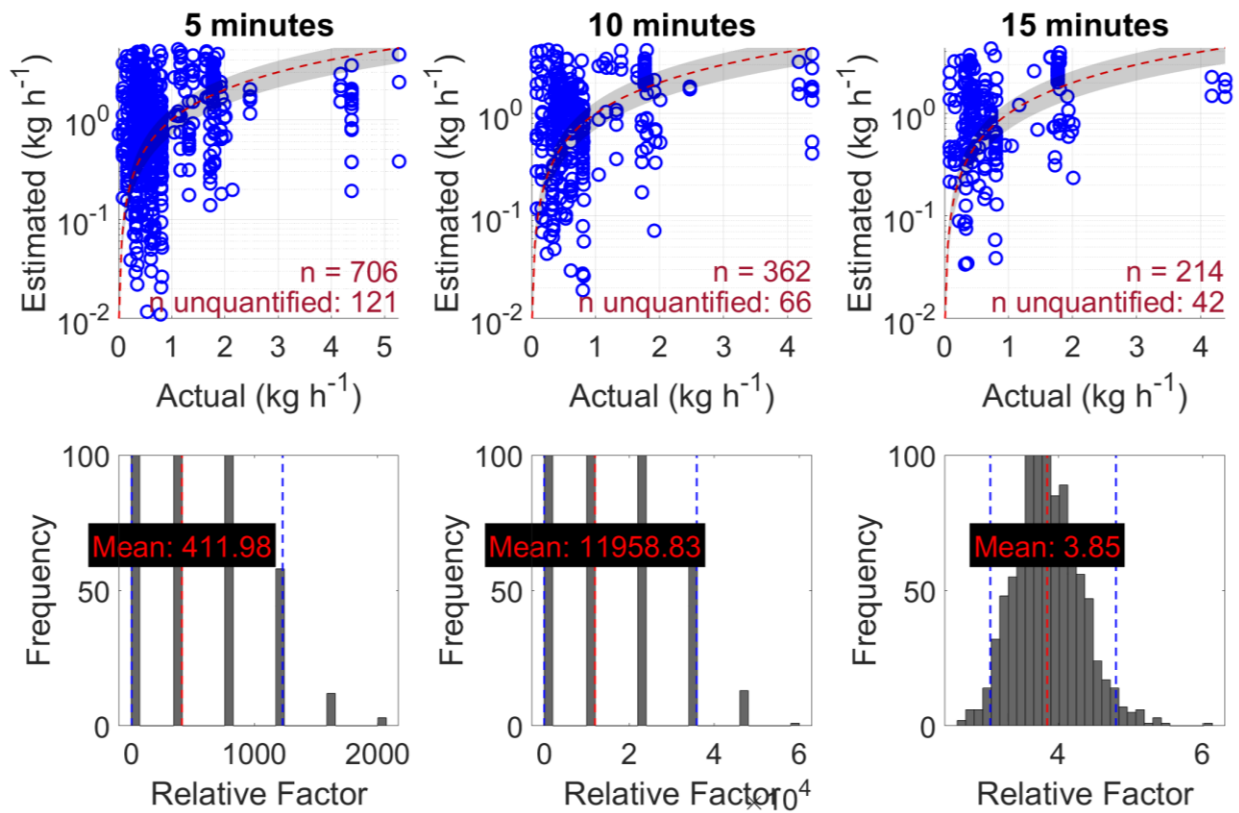


**Figure 10.** Top plot: Estimated emission vs actual emission ( $\text{kg h}^{-1}$ ) for a multi-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is n, and “n unquantified” is the number of points WindTrax reported -9999 (i.e. could not quantify). Bottom plot: A bootstrap of mean relative factor (MRF: estimated emissions divided by actual controlled emission) for a multi-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.

### 3.4.2 Multi-Release Single-Point

For MRSP emissions, the bLs method largely overestimated emissions at between an MRF of 3.85 and 12239.20. At  $\pm 10^\circ$  wind sector range, the MRF was 411.98, 11958.83 and 3.85 for a 706, 362 and 214 sample size, at 5, 10 and 15-minutes averaging period, respectively (Figure 11). At  $\pm 5^\circ$  wind sector range, the MRF was 7.04 and 6.81 and 10 and 15-minutes averaging periods, 186 and 126 sample sizes; and 12239.20 and 5.08 for a 458 and 286 sample size, 10 and 15-minutes averaging period, at  $\pm 20^\circ$  wind sector range (Supplementary Information Section 3.4.2). These results show that the bLs model largely overestimated emissions when there were multiple releases at the site level.

**Backward Lagrangian Stochastic Model**  
**Multi-Release Single-Point -  $\pm 10^\circ$**



*Figure 11. Top plot: Estimated emission vs actual emission ( $\text{kg h}^{-1}$ ) for a multi-release single-point at site level,  $\pm 45^\circ$  wind sector range. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents  $\pm 30\%$  of the actual emission. The sample size is n, and “n unquantified” is the number of points WindTrax reported -9999 (i.e. could not quantify). Bottom plot: A bootstrap of mean relative factor (MRF: estimated emissions divided by actual controlled emission) for a multi-release single-point,  $\pm 45^\circ$  wind sector range. An MRF of less than 1 shows an overall underestimation of emissions while an MRF of greater than 1 shows an overall overestimation of emissions. The dotted blue lines are the lower confidence intervals (CI) and upper CI, 95% confidence intervals.*

The aerodynamic flux gradient method also largely underestimated emissions for single release single emission, multi release single emission and multi release multi emission (Figure 5). A plot of quantified emission versus actual release did not show a linear relationship ( $R^2$  between 0.01 and 0.39), and most data points were below the 1:1 line (Figure 5). The aerodynamic flux gradient quantified emissions were between 0 and about  $1.6 \text{ kg h}^{-1}$  despite actual emissions being between 0 and about  $7 \text{ kg h}^{-1}$  (Figure 5). The underestimation was also consistent across all downwind ranges, 10, 20 and 30 degrees (Supplementary Material Section 2.2).

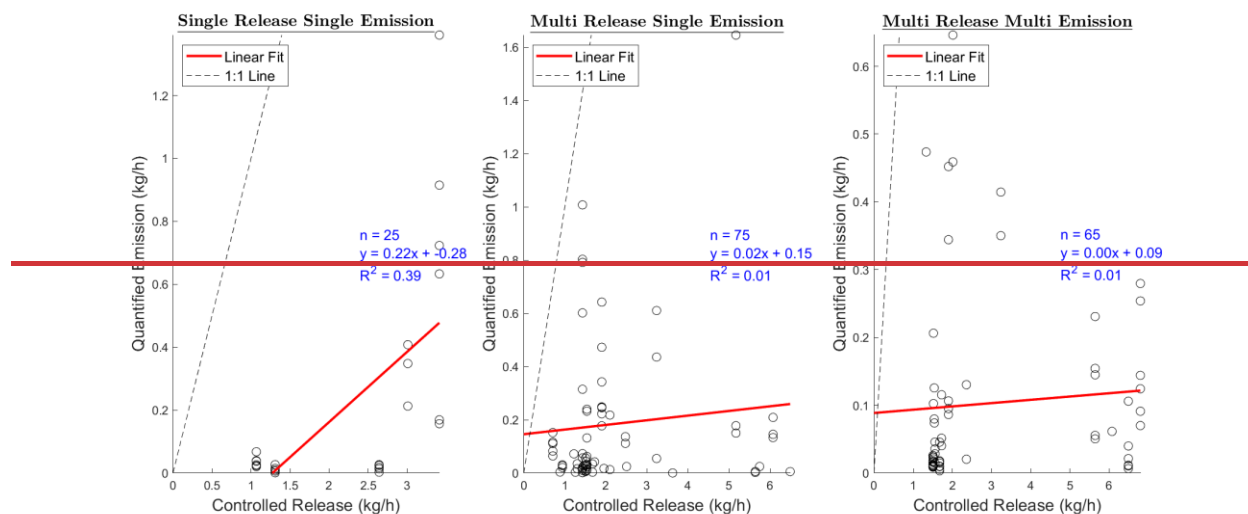


Figure 5: Quantified emission calculated using the aerodynamic flux gradient method. Left pane shows a scatter plot of quantified emission versus total controlled release for a single release at the site level and the mast was downwind of the release point. Center pane shows a scatter plot of quantified emission versus total controlled release for multiple releases at the site level, but the mast was downwind of a single release point. Right pane shows a scatter plot of quantified emission versus total controlled release for multiple releases at the site level and the mast was downwind of more than one release point. The dashed line represents the 1:1 line (points below the line were underestimated), the red line is the linear regression fit of the data, and n is the number of data points.

### 3.1.3 Gaussian Plume Inverse Method

The Gaussian plume inverse method was tested for single release single emission and multi release single emission as the method is only used for single point sources and preliminary results showed the method provided reasonable results within 20 degrees downwind range (Figure 6; Supplementary Material Section 1.3). For single release single emission, the method quantified emissions within a factor of 1.5 (Figure 6) and showed reasonably linear relationship ( $R^2$  of 0.65) (Figure 6). For multi release single emission, the gradient (m) of the linear regression was 0.95 and  $R^2$  of 0.21. This suggests that the linear relationship cannot be well explained due to a random scatter of calculated emissions.

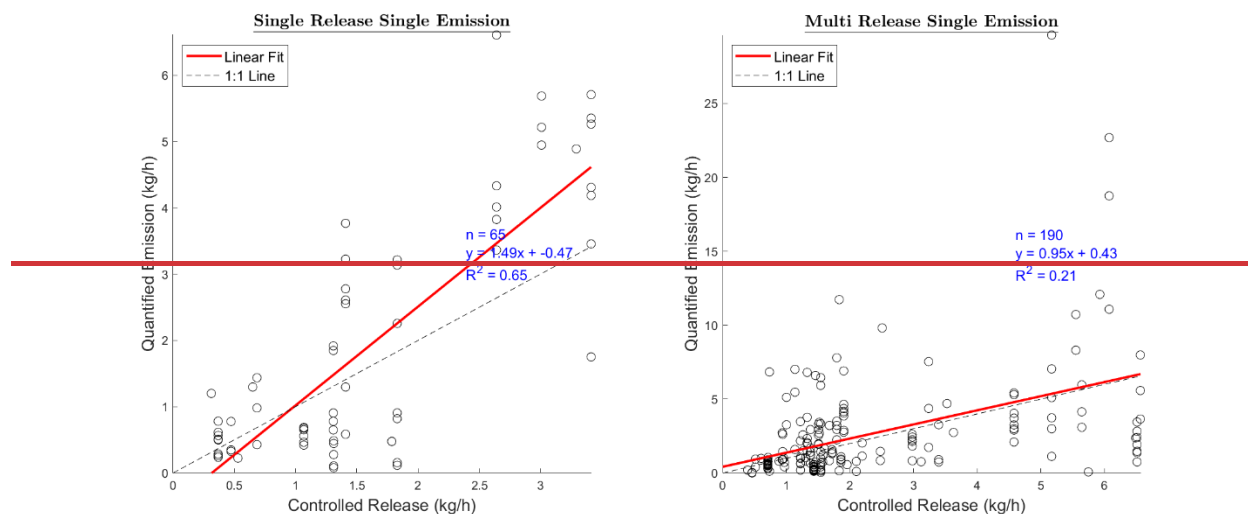


Figure 6: Quantified emission calculated using the Gaussian plume inverse method. Left pane shows a scatter plot of quantified emission versus total controlled release for a single release at the site level and the mast was downwind of the release point. Right pane shows a scatter plot of quantified emission versus total controlled release for multiple releases at the site level, but the mast was downwind of a single release point. The dashed line represents the 1:1 line (points below the line were underestimated), the red line is the linear regression fit of the data, and n is the number of data points.

### 3.2 Quantification within 30% Uncertainty

#### 3.2.1 Eddy Covariance

The eddy covariance method showed a very low probability of quantifying emissions within 30% uncertainty ( $\pm 30\%$ ) (Figure 7). Only a single measurement in the multi release multi emission category showed an approximately 0.01 probability of quantifying within 30% (Figure 7). The errors for eddy covariance were between 100 and 86% for single release single emission, between 100 and 82% for multi release single emission, and between 100 and about +30% for multi release multi emission (Figure 7). This shows that using eddy covariance to quantify single point and multi point emissions will largely underestimate emissions.

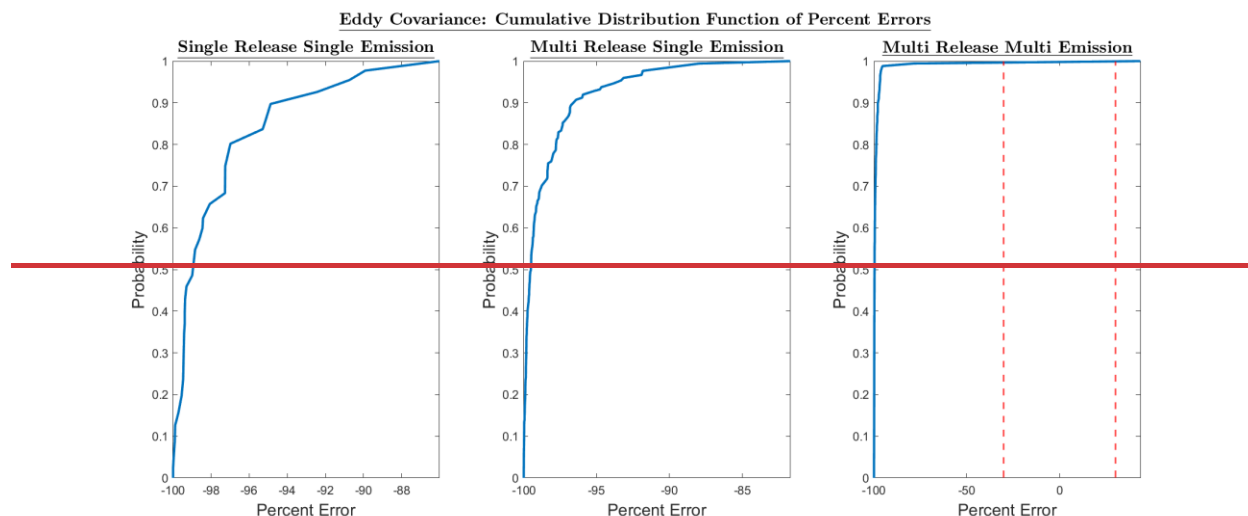


Figure 7: Cumulative distribution function (cdf) of percent errors for eddy covariance. Left pane shows a cdf plot for a single release at the site level and the mast was downwind of the release point. Center pane shows a cdf for multiple releases at the site level, but the mast was downwind of a single release point. Right pane shows a cdf for multiple releases at the site level and the mast was downwind of more than one release point. The area bounded by the red dotted line shows the region within  $\pm 30$  uncertainty.

### 3.2.2 Aerodynamic Flux Gradient

The aerodynamic flux gradient also showed a very low probability of quantifying within 30% uncertainty (Figure 8). In the multi release single emission category results indicate a 0.02 probability of quantifying within 30% (Figure 8) of the true value. The errors for aerodynamic flux gradient were between -100 and -60% for single release single emission, between -100 and 0% for multi release single emission, and between -100 and -70% for multi release multi emission (Figure 8). These data show that the aerodynamic flux gradient will underestimate a point emission. Similar to eddy covariance, quantifying an emission within 30% uncertainty using aerodynamic flux gradient for point sources is highly unlikely.

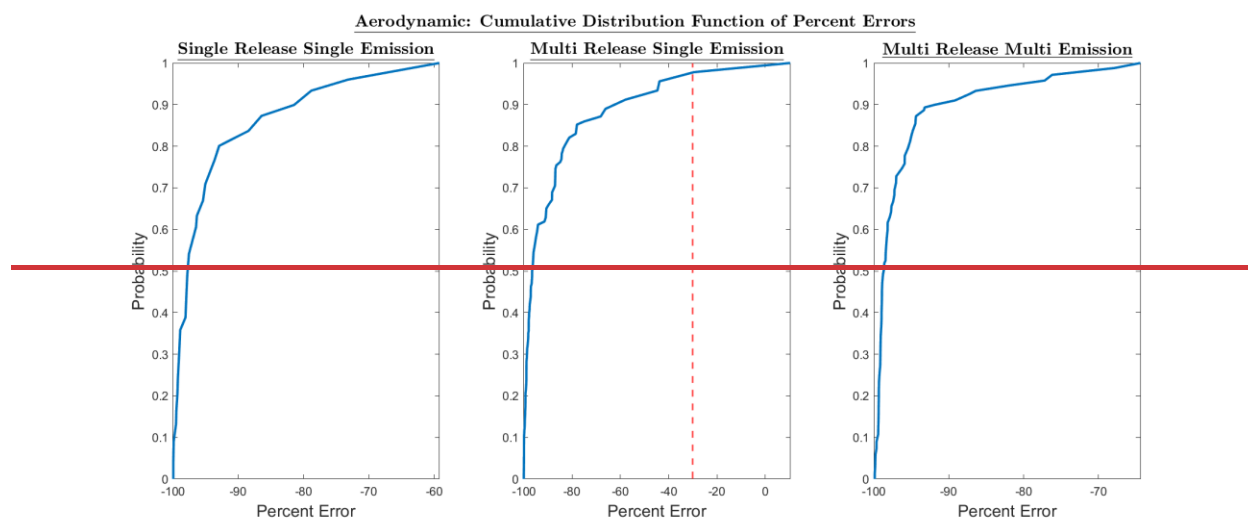


Figure 8: Cumulative distribution function (cdf) of percent errors for aerodynamic flux gradient method. Left pane shows a cdf plot for a single release at the site level and the mast was downwind of the release point. Center pane

shows a cdf for multiple releases at the site level, but the mast was downwind of a single release point. Right pane shows a cdf for multiple releases at the site level and the mast was downwind of more than one release point. The area bounded by the red dotted line shows the region within  $\pm 30$  uncertainty.

### 3.2.3 Gaussian Plume Inverse Method

The Gaussian plume inverse method showed a higher probability of quantifying an emission correctly within 30% uncertainty than eddy covariance and aerodynamic flux gradient methods (Figure 9);  $\sim 0.12$  for the single release single emission and  $\sim 0.25$  for the multi release single emission categories (Figure 9). Percent errors of the Gaussian method calculated emissions are between  $-100$  and  $+250\%$  for single release single emission and between  $-100$  and  $+800\%$  for multi release single emission (Figure 9). This shows that even though the Gaussian method is designed for point sources, it is highly likely to miss, underestimate or overestimate an emission. Similar to eddy covariance and aerodynamic flux gradient, it is a challenge to correctly quantify a single emission event (single release or multiple release) using the Gaussian plume inverse method.

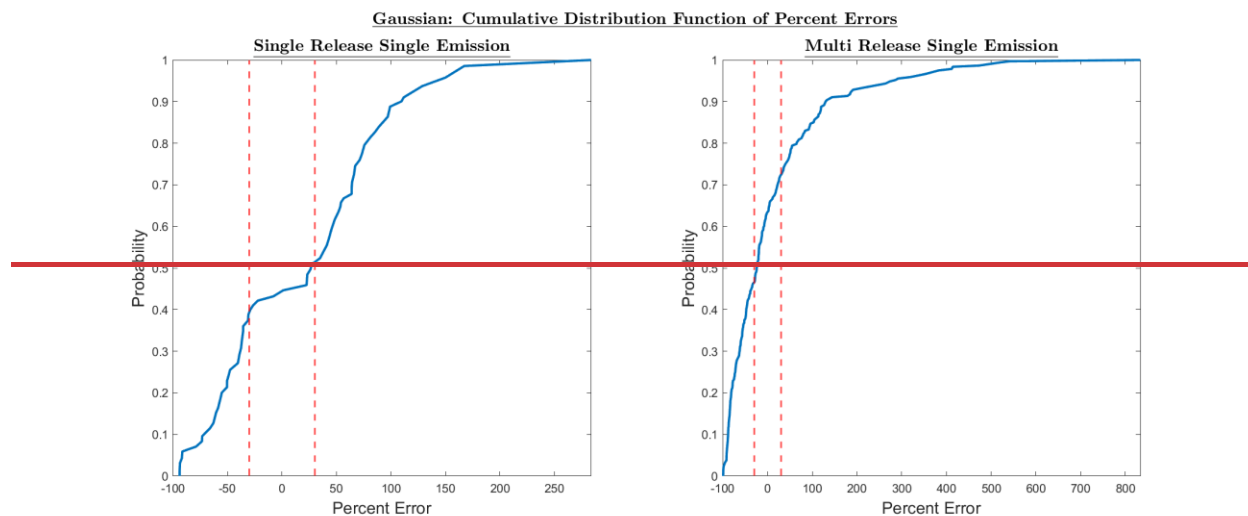


Figure 9: Cumulative distribution function (cdf) of percent errors for the Gaussian plume inverse method. Left pane shows a cdf for a single release at the site level and the mast was downwind of the release point. Right pane shows a cdf for multiple releases at the site level, but the mast was downwind of a single release point. The area bounded by the red dotted line shows the region within  $\pm 30$  uncertainty.

## 3.3 Variables Affecting Quantification

### 3.3.1 Eddy Covariance

A Spearman's rank correlation analysis of measurement and environmental variables (distance, controlled release, emission height, mean wind speed (WS), Monin Obukhov length ( $L$ ) and contribution area) to percent error in quantification as calculated by the eddy covariance method, showed that downwind distance had significant impact on quantification for the single release single emission ( $p = 4.73e-6$ ), multi release single emission ( $p = 2.66e-4$ ), and multi release multi emission ( $p = 2.00e-3$ ) categories for  $p < 0.01$  significance level (Figure 10). The correlation coefficients were  $-0.74$  for single release single emission,  $-0.31$  for multi release single emission, and  $-0.30$  for multi release multi emission. The negative correlation in all three categories suggests that the percent error became more negative as distance increased i.e., far away emission sources were likely underestimated or undetected. Also,

controlled release and emission height had significant impact on quantification only in the multi release single emission category,  $p = 2.00e-3$  and  $9.42e-3$  respectively, (Figure 10) but this correlation was inconsistent across the three categories. Due to inconsistent correlation, and the errors being close to  $\pm 100\%$ , the results show that generally, quantifying emissions using an eddy covariance approach will not work for emissions typically observed at oil and gas production sites.

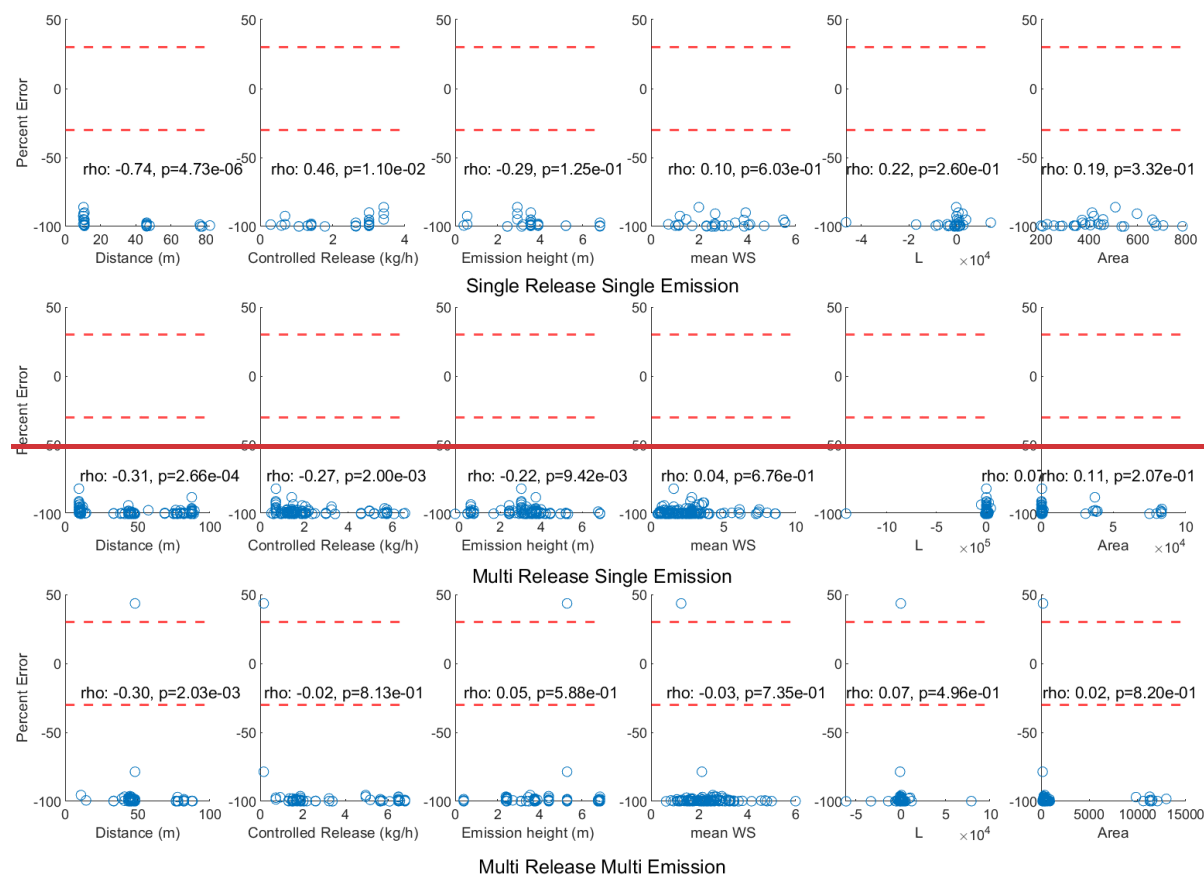


Figure 10: Correlation analysis for eddy covariance in the three release categories. The area bounded by the red dotted line shows the region within  $\pm 30$  uncertainty.

### 3.3.2 Aerodynamic Flux Gradient

A Spearman's rank correlation analysis between the environmental and measurement variables and emissions calculated using the aerodynamic flux gradient method showed that only emission height in the single release single emission category had significant impact on model quantification ( $p = 1.79e-3$ ) (Figure 11). The correlation between emission height and percent error in this category was  $-0.59$  suggesting percent error became more negative as emission height increased. However, the correlation between emission height and percent error in the multi release single emission and multi release multi emission categories is approximately zero, meaning no correlation. Similar to eddy covariance, there is inconsistent correlation, and most errors are close to  $\pm 100\%$  (Figure 11). The results show that generally, quantifying emissions using an aerodynamic flux gradient approach will not work for emissions typically observed at oil and gas production sites.

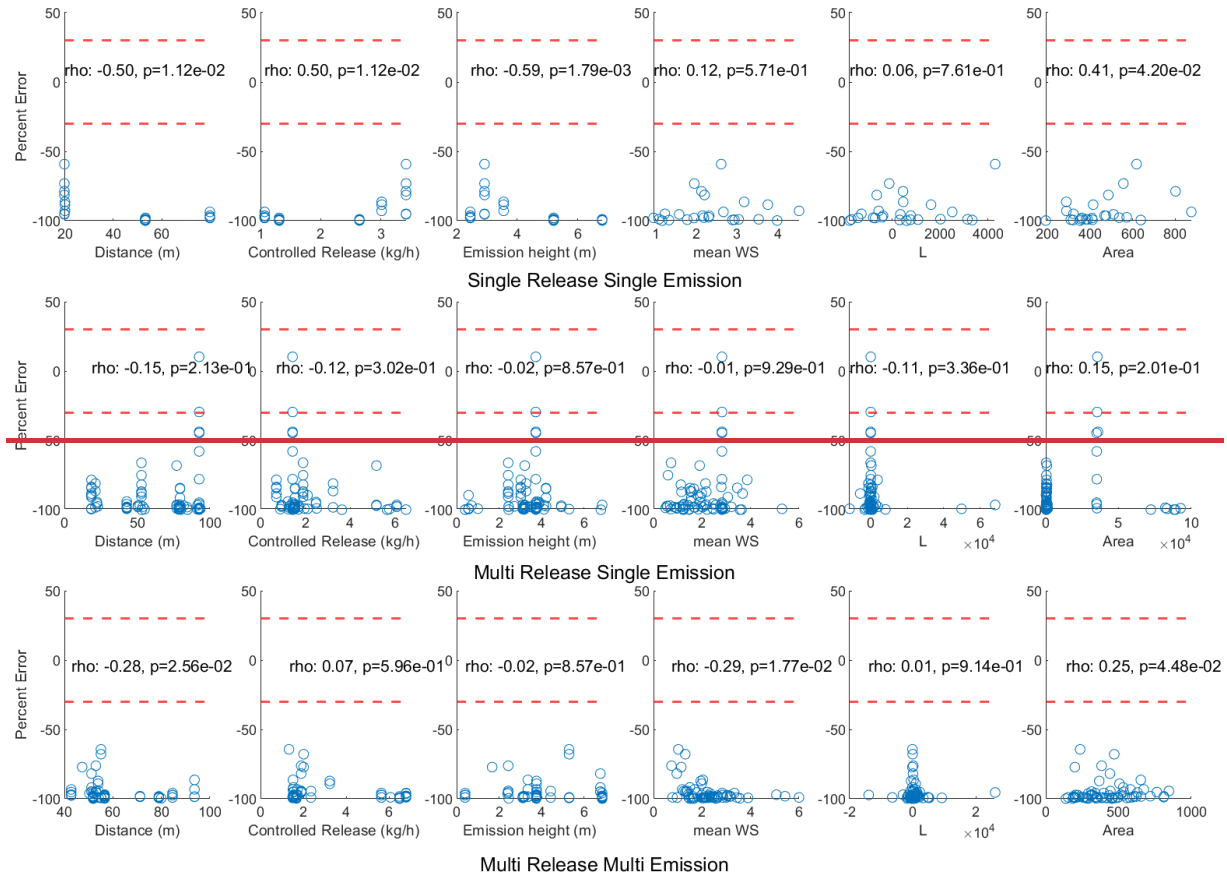


Figure 11: Correlation analysis for aerodynamic flux gradient in the three release categories. The area bounded by the red dotted line shows the region within  $\pm 30$  uncertainty.

### 3.3.3 Gaussian Plume Inverse Method

The Spearman's rank correlation analysis between the emissions calculated using the Gaussian plume inverse method and measurement/environmental variables showed that only the mean wind speed and atmospheric stability had significant impact on the model quantification (Figure 12). In the single release single emission category, mean wind speed and percent error had a positive correlation ( $0.44, p = 2.74e-4$ ) indicating that an increase in WS increased the model's positive error. However, in the multi release single emission category, the correlation is opposite (a negative correlation of  $-0.21, p = 3.71e-3$ ) (Figure 12). Atmospheric stability had significant impact on model quantification in the multi release single emission category ( $p = 9.15e-5$ ) but not in the single release single emission category (Figure 12). The correlation analysis for the Gaussian plume inverse model was inconsistent suggesting random errors in quantification. This shows that the model could either underestimate or overestimate an oil and gas emission at random.

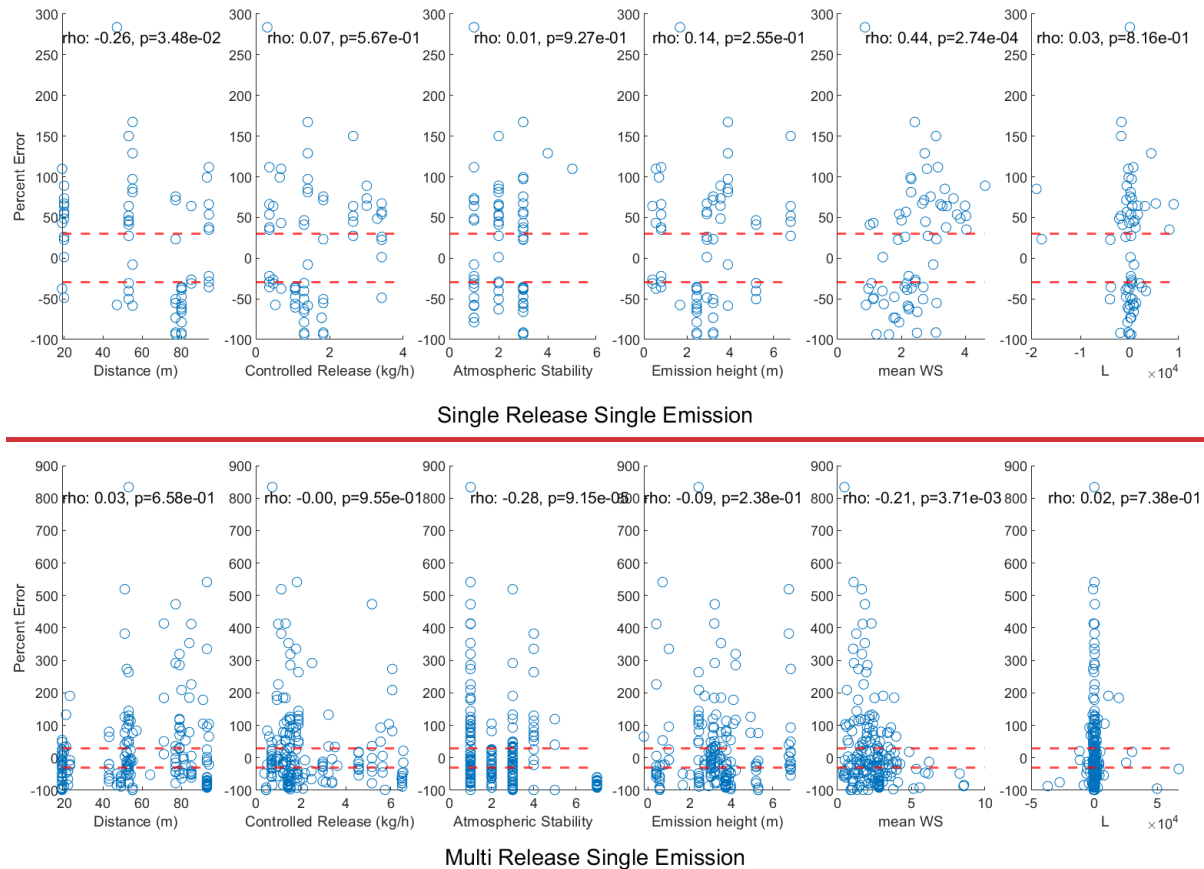


Figure 12: Correlation analysis for the Gaussian plume inverse method in the three release categories. The area bounded by the red dotted line shows the region within  $\pm 30\%$  uncertainty.

## 4 Discussion

Methane emissions quantification from oil and gas is a complex system comprising of gas emissions from different heights, different locations, encountering aerodynamic obstacles of different sizes, and of varying emissions duration, amongst others. The ability to precisely quantify an emission using data collected by a point sensor, downwind of a source is directly influenced by plume dynamics. The CH<sub>4</sub> plume downwind of a source will change in size and shape in different atmospheric conditions, in open areas versus areas with obstacles, diurnally, and in different seasons (Casal, 2008). In this study, the precision to which downwind methods/models (closed-path EC, AFG, GPIM and bLs) eddy covariance, aerodynamic flux gradient and Gaussian plume based could quantify the emission rate of point source(s) were tested in different atmospheric conditions (rain, sunny, snow, windy, calm etc.), and aerodynamic scenarios (emissions sources in open areas, behind obstacles, changing atmospheric stability, and day/night). As a result, testing the models' predicted emission rates to controlled release rates in different conditions introduced real-world scenarios that have not previously been tested, hence better understanding model uncertainty in the application of quantifying emissions from oil and gas production infrastructure.

### 4.1 Eddy Covariance

Eddy covariance was tested using a closed-path analyzer, cavity ring-down spectroscopy, with a 3.2 lpm pump flowrate and a 0.4 s gas flow response time. The closed-path EC estimated emissions between a factor of 0.67

and 0.97 for SRSP emissions, and between 1.02 and 2.43 for MRSP emissions at  $\pm 45^\circ$  wind sector range (Section 3.1). This was a wider uncertainty in estimated emissions than one reported by Dumortier et al. (2019), who estimated emissions at between 90 and 113% of true emission ( $\sim 1.5 \text{ kg day}^{-1}$ ) with concentrations between 2 and 3 ppm. Our study tested closed-path EC at emission rates between 0.005 and  $8.5 \text{ kg h}^{-1}$ .

Our study's results were when the data was filtered for frequencies greater than 8 Hz, hence, largely reducing the sampled emissions. The 10 Hz sampling frequency set for this instrument was not a true 10 Hz and this could have been due to the 0.4 gas flow response time that delayed analysis of the drawn air sample in the cavity, or the 3 lpm pump flow rate for a 3 m tubing that might have varied the effective sample turnover rate. The rest of the data were flagged as low quality by Mauder and Foken (2004) (0-1-2 system), which flags based on steady state and well-developed turbulence. This could have been due to low turbulence as experiments were carried out in winter, and instrumentation limitations (low pump flow rate and asynchronous configuration of the gas analyzer and meteorological instrument).

Continuous monitoring requires deployment of multiple sensors which create limitations of cost and requires instrumentation with a wide measurement range as concentrations for oil and gas emissions can range between 0 to 250 ppm, as in this study (Supplementary Information Section 2c). The currently available EC instruments have a narrow measurement range (LI-COR LI-7700 open path  $\text{CH}_4$  analyzer has a measurement range of 0 to 25 ppm at  $-25^\circ\text{C}$  and 0 to 40 ppm at  $25^\circ\text{C}$ ; PICARRO G2311-f, a closed-path analyzer has an operating range of 0 to 20 ppm). Also, the instrumentation should be environmentally robust and not lab-grade (be able to run smoothly in adverse weather conditions). Given these parameters, market available EC instruments that can currently be deployed in oil and gas are limited. The instrument used in this study is a field instrument (ABB MGGA GLA131 Series has a measurement range of 0 to 100 ppm for  $\text{CH}_4$  but can be extended to 0 to 1%).

Eddy covariance underestimated or failed to observe almost all emissions released during this study (linear regression  $m$  between 0 and 0.07, and  $R^2$  between 0.03 and 0.36) (Figure 4). The method measures  $\text{CH}_4$  atmospheric fluxes for area sources transferred as eddies of different sizes as caused by turbulence within the atmospheric boundary layer (Babaeian and Tuller, 2023). Assumptions governing eddy covariance include: (1) the terrain is homogenous and horizontal, (2)  $\text{CH}_4$  fluxes are turbulent, (3) measurements at a point are from an upwind area, (4) measurements are within the boundary layer and in the constant flux layer, (5) instruments can capture small fluctuations at high frequency, (6) fluctuations in air density are negligible (Babaeian and Tuller, 2023), and (7) upward fluxes represent emissions and downward fluxes represent depositions (Zinke et al., 2024). Nemitz et al., (2018) adds that eddy covariance is frequently deployed to target large fluxes in high emission ecosystems, which is not typical in oil and gas, and that data where wind direction includes obstructed wind sectors should be flagged (Nemitz et al., 2018).

For oil and gas point sources, the measured gas concentration is dependent on plume dynamics as opposed to mass transfer and eddy covariance methods using fence line measurements are unlikely to work because:

- Oil and gas point sources violate assumptions (1), (2), and (4) as these sources are heterogenous and emissions are collimated plumes instead of turbulent fluxes.

- As the measurement by a point sensor is dependent on being inside the plume, which changes in different atmospheric conditions, placing the sensor high enough, and/or far enough downwind, to where the flux layer is constant, is impractical.
- Even though current eddy covariance application assumes the vertical flux at a point is independent of atmospheric stability (Denmead, 2008), atmospheric stability has impact on point source gas dispersion at fence line distances and hence needs to be accounted for even for eddy measurements.
- Footprint models are designed for area sources that require horizontal homogeneity of the flow (Kljun et al., 2015). As a result, the area of contribution generated by the models do not accurately represent the area between the point sources and the measurement location at fence line distances.

In summary, this study shows that eddy covariance is not applicable for oil and gas point source quantification.

## 4.2 Aerodynamic Flux Gradient

Overall, the AFG method quantified emissions within an MRF of 1.3 to 1.7 for SRSP emissions and between an MRF of 2.4 and 3.3 for MRSP emissions (Section 3.2). The uncertainties in AFG were higher than EC especially for MRSP emissions but lower than the GPIM and bLs methods. The differences between AFG and EC could have been due to different instrumentation and analytical approaches that limited the exact comparison of the methods i.e. EC data was filtered for frequencies less than 8 Hz, and AFG instrumentations required both analyzers to be running, periods when one analyzer was down, were not tested. To our knowledge, this is the first study to test AFG for point source quantification and results are promising.

Compared to the EC method that requires a very fast analyzer which may be difficult to deploy in oil and gas, the AFG requires at least 2 analyzers sampling at 1 Hz frequency, which is currently possible with the range of sensors available in the market. The main advantage of flux gradient methods (EC and AFG) tested in this study is that they do not require background control as background  $\text{CH}_4$  concentration is a highly variable parameter that cannot be controlled in open air especially when multiple emissions are happening. The AFG method relies on differences in  $\text{CH}_4$  concentrations between two heights and this study shows that in complex sites where there are multiple sources, the method quantifies better than the point-source GPIM and bLs methods. The main limitation of the AFG and EC methods for point-source quantification is that when the measurement point is downwind of more than one source in a wind sector range, more than sonic anemometers are required to estimate the flux of each source for footprint calculation based on Dumortier et al. (2019)'s calculations.

Overall, aerodynamic flux gradient method underestimated the emission rate of all controlled releases during this experiment with high variability. The slope of the linear regression and  $R^2$  were both very small (linear regression  $m$  between 0 and 0.22, and  $R^2$  between 0.01 and 0.39) (Figure 5). The aerodynamic flux gradient model quantification is used to quantify emissions from area sources and relies on differences in  $\text{CH}_4$  concentrations between the higher and lower height, and stability correction factors. Assumptions of flux gradient approach using Monin-Obukhov similarity theory include: (1) measurements require steady state conditions of wind direction and speed, (2) measurements should be done above the roughness sub layer, (3) sufficiently large homogenous area for development of an adequately equilibrated layer of air, and for constant equilibrium during measurement (Prueger and Kustas, 2015), and (4) positive fluxes represent emissions and downward fluxes represent absorptions (Kamp et al., 2020).

Similar to eddy covariance, aerodynamic flux gradient methods at fence line distances are unlikely to work because point sources typical of oil and gas emissions violate the following assumptions:

- Obstacles at an oil and gas facility affects wind direction and speed, and these impacts may also vary substantially with small changes in wind direction. Therefore, wind conditions are unlikely to attain steady state during the measurement period, as directed by assumption (1) above.
- The emission height of oil and gas sources in typical upstream field conditions can be as low as 0.4 m and as high as 6.9 m and measurements are unlikely to be made by fence line sensor above the roughness sublayer (2 above), i.e. twice the height of the mean obstacle height for ~30 m downwind.
- Oil and gas sources are heterogeneous (i.e. varying source distance and height) and can last a short time (e.g. a short maintenance event) or a long time ('normal' fugitive emissions) hence, achieving constant equilibrium, as stated in (3) above, is unlikely.
- Footprint models used to generate the area of contribution between the source and the measurement location are designed for area sources with horizontal flow homogeneity (Kljun et al., 2015). Thus, the area of contribution generated for oil and gas point sources is likely inaccurate.

#### 4.3 Gaussian Plume Inverse Method

The GPIM method quantified emissions within an MRF of 2.4 and 2.6 for SRSP and between 15.7 and 25 for MRSP emissions (Section 3.3). The GPIM method is a point-source specific quantification approach and works best in open areas, free of obstacles, and when the background concentration is well defined. For multiple emissions, in aerodynamically complex environments, even though the sensor is downwind of a single source based on average wind direction, quantification is complexed by interference from other neighboring sources. The GPIM has previously been reported to quantify emissions within 40.7 and 60% error for a single point-source, (Riddick et al., 2022b). However, GPIM correct quantification has been suggested to be possible for longer distances where the plume is well mixed. This is typically a challenge for fence-line sensors that have to be deployed within the facility boundaries where large downwind distances may not be practical.

#### 4.3 Backward Lagrangian Stochastic Model

The bLs method quantified emissions within an MRF of 0.8 to 1.05 for SRSP emissions and between 3.85 and 11958.8 MRF for MRSP emissions (Section 3.4). Similarly to the GPIM method, the bLs method used in this study is a point-source specific quantification method that simulates transport of molecules in open area and where the background concentration is defined. In this case, as with the SRSP test scenario, the bLs approach quantified within 20% uncertainty. However, for MRSP emissions, the bLs largely overestimated emissions and this could have been due to the interference of neighboring sources, that even though the measurement point is downwind of a single source, actual plumes are not distinct and model-simulated plumes may not be representative. The point-source bLs approach in WindTrax is also not designed for more than one downwind source.

In contrast to the other methods in this study, the Gaussian plume inverse model both underestimated and overestimated emissions in this study. Linear regression gradient and coefficient of correlation ( $m$  between 0.95 and 1.49, and  $R^2$  between 0.21 and 0.65; Figure 6) was better than either eddy covariance or aerodynamic flux gradient. The main assumption of the Gaussian plume model is that  $CH_4$  emitted from a point source enters the air flow,

disperses vertically and laterally, forming a conical plume (Riddick et al., 2022b; US EPA, 2013). However, the formation of a conical plume is hindered at oil and gas facilities by obstacles (equipment) and is affected by atmospheric stability. Atmospheric stability in the Gaussian plume inverse model is based on Pasquill Gifford classification system which accounts for daytime solar insolation (slight, moderate and strong), nighttime cloud cover and surface wind speed at 10 m (Kahl and Chapman, 2018). Solar insolation and cloud cover are not typically measured, and if measured, dispersion parameter models currently available do not use this data, therefore, it is difficult to calculate for continuous fence-line measurements. The modified dispersion parameters developed by EPA (US EPA, 2013) only account for wind conditions i.e., speed and deviation in direction. As a result, plume dynamics during diverse atmospheric conditions such as during snow versus rain or sunny conditions are unaccounted for. In this study, despite the Gaussian model having been developed for point sources, the model did not show consistent correlation with the measurement and atmospheric variables. This showed that there are complexities in continuous monitoring quantification compared to survey solutions where the model is widely applied, that introduce significant uncertainties in quantification. It is suggested that one problem with the Gaussian plume model is that the dispersion coefficients are simply not representative as they were developed for longer distances, in different climatological conditions, and do not transfer well to current applications (Riddick et al., 2022a). We conclude that, while it is better suited than eddy covariance or aerodynamic flux gradient, a Gaussian plume inverse approach will likely have significant uncertainties when used to quantify emissions from oil and gas production sites using data collected at a fence line (~30 m away).

#### 4.4 Implications

In the recent years, there has been growing interest and need for accurate CH<sub>4</sub> quantification from oil and gas sites. This is generally done through survey methods and continuous monitoring using fence-line sensors. Continuous monitoring involves having stationary sensors measuring meteorology and CH<sub>4</sub> mixing ratios, which are then used to infer emission rates. For point sources, downwind methods such as the Gaussian plume inverse method have been widely used, especially for survey quantification. Continuous monitoring is relatively new but fast growing. This study's design replicated a continuous monitoring setup's downwind deployment distance, range of typical emission rates, emissions heights, and meteorological data acquisition.

Oil and gas point sources could either be single emissions or multiple emissions occurring concurrently. In cases of multiple emissions with more than one release point being downwind, the Gaussian model and the backward Lagrangian stochastic models are limited, as they can only quantify one source at a time (dispersion coefficients are generated as a function of emission height and source distance); and interference from neighboring emissions affects the underlying principles of dispersion on which these models were developed. As a result, flux quantification models used in other applications such as eddy covariance and aerodynamic flux gradient have been proposed as the solution. This study's results show that generally reasonable quantification estimates are achieved with flux approaches (eddy covariance and aerodynamic flux gradient), but these methods require more instrumentation effort (fast sampling analyzer for eddy covariance, and multiple collocated sensors for aerodynamic flux gradient). Even though the widely applied Gaussian plume inverse method and the backward Lagrangian stochastic models are widely used for single-point emissions, this study shows aerodynamic complexities, the

difficulty in defining the background, and interference from neighboring sources challenge the application of these models for fence-line continuous monitoring. This study recommends more testing of flux quantification models for oil and gas quantification as they could improve emissions quantification for leak repair prioritization and methane reporting.

~~However, as this study has shown, eddy covariance and flux gradient approaches are unlikely to quantify realistic emission estimates using fence line measurements. Here, we strongly advise that controlled tests under controlled environments are crucial to evaluate modelling approaches' precision and accuracy, and associated uncertainties before applying them in the real world. Even though these modelling approaches have been reported to work elsewhere (e.g., agricultural and landfill emissions), it does not necessary mean it could work in the intended area of application.~~

#### **Author contributions**

Mercy Mbua: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Writing original draft, Review, and Editing. Stuart N. Riddick: Conceptualization, Methodology, Review, Editing, Project Administration, Supervision, and Funding Acquisition. Elijah Kiplimo: Investigation, Review, and Editing. Kira Shonkwiler: Review and Editing. Anna Hodshire: Review and Editing. Daniel J. Zimmerle: Review, Editing, Funding acquisition, and Project administration.

#### **Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### **Acknowledgment**

This work is funded by the Office of Fossil Energy and Carbon Management within the Department of Energy as part of the Site-Air-Basin Emissions Reconciliation (SABER) Project #DE-FE0032288. Any opinions, findings, conclusions, or recommendations expressed herein are those of the authors and do not necessarily reflect the views of those providing technical input or financial support. The trade names mentioned herein are merely for identification purposes and do not constitute an endorsement by any entity involved in this study. The authors also thank Ryan Brouwer, Daniel Fleischmann, Ryan Buenger, and Wendy Hartzell for their assistance.

#### **Data availability**

Data sets for this research are available in the in-text data citation reference: Mbua, Mercy; Riddick, Stuart N.; Kiplimo, Elijah et al. (Forthcoming 2025). Evaluating the feasibility of using downwind methods to quantify point source oil and gas emissions using continuous monitoring fence-line sensors [Dataset]. Dryad. <https://doi.org/10.5061/dryad.hhmgqnkss>

~~Data sets for this research are available in the in-text data citation reference: Mercy, Mbua; Riddick, Stuart N.; Kiplimo, Elijah, and Zimmerle, Daniel J. Dataset for evaluating the accuracy of downwind methods for quantifying point source emissions. [Dataset]. Dryad.~~

#### **References**

Bell, C., Ilonze, C., Duggan, A., Zimmerle, D., 2023. Performance of continuous emission monitoring solutions under a single-blind controlled testing protocol. *Environ. Sci. Technol.* 57, 5794–5805. <https://doi.org/10.1021/acs.est.2c09235>.

Brown, J.A., Harrison, M.R., Rufael, T., Roman-White, S.A., Ross, G.B., George, F.C., Zimmerle, D., 2023. Informing methane emissions inventories using facility aerial measurements at midstream natural gas facilities. *Environ. Sci. Technol.* 57, 14539–14547. <https://doi.org/10.1021/acs.est.3c01321>.

Carbon Mapper - Science & Technology [WWW Document], n.d. URL <https://carbonmapper.org/work/science> (accessed 2.18.25).

Casal, J., 2008. Chapter 6 Atmospheric dispersion of toxic or flammable clouds, in: *Industrial Safety Series*. Elsevier, pp. 195–248. [https://doi.org/10.1016/S0921-9110\(08\)80008-0](https://doi.org/10.1016/S0921-9110(08)80008-0).

Colorado State University, 2021. Mechanistic Emission Estimation Tool (MEET) [WWW Document]. Energy Inst. URL <https://energy.colostate.edu/mechanistic-air-emissions-simulator/> (accessed 2.26.24).

Conrad, B.M., Tyner, D.R., Johnson, M.R., 2023. Robust probabilities of detection and quantification uncertainty for aerial methane detection: Examples for three airborne technologies. *Remote Sens. Environ.* 288, 113499. <https://doi.org/10.1016/j.rse.2023.113499>.

Crenna, B., 2006. An introduction to WindTrax. <http://thunderbeachscientific.com/downloads/introduction.pdf>.

Day, R.E., Emerson, E., Bell, C., Zimmerle, D., 2024. Point sensor networks struggle to detect and quantify short controlled releases at oil and gas sites. *Sensors* 24, 2419. <https://doi.org/10.3390/s24082419>.

Denmead, O.T., 2008. 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>.

Dumortier, P., Aubinet, M., Lebeau, F., Naiken, A., Heinesch, B., 2019. Point source emission estimation using eddy covariance: Validation using an artificial source experiment. *Agric. For. Meteorol.* 266–267, 148–156. <https://doi.org/10.1016/j.agrformet.2018.12.012>.

Foken, Th., Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. *Agric. For. Meteorol.* 78, 83–105. [https://doi.org/10.1016/0168-1923\(95\)02248-1](https://doi.org/10.1016/0168-1923(95)02248-1).

Foster-Wittig, T.A., Thoma, E.D., Albertson, J.D., 2015. Estimation of point source fugitive emission rates from a single sensor time series: A conditionally-sampled Gaussian plume reconstruction. *Atmos. Environ.* 115, 101–109. <https://doi.org/10.1016/j.atmosenv.2015.05.042>.

Fratini, G., Ibrom, A., Arriga, N., Burba, G., Papale, D., 2012. Relative humidity effects on water vapour fluxes measured with closed-path eddy-covariance systems with short sampling lines. *Agric. For. Meteorol.* 165, 53–63. <https://doi.org/10.1016/j.agrformet.2012.05.018>.

Horst, T.W., Lenschow, D.H., 2009. Attenuation of scalar fluxes measured with spatially-displaced sensors. *Boundary-Layer Meteorology*. 130, 275–300. <https://doi.org/10.1007/s10546-008-9348-0>.

Hsieh, C.-I., Katul, G., Chi, T., 2000. An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. *Adv. Water Resour.* 23, 765–772. [https://doi.org/10.1016/S0309-1708\(99\)00042-1](https://doi.org/10.1016/S0309-1708(99)00042-1).

Ilonze, C., Emerson, E., Duggan, A., Zimmerle, D., 2024. Assessing the progress of the performance of continuous monitoring solutions under a single-blind controlled testing protocol. *Environ. Sci. Technol.* 58, 10941–10955. <https://doi.org/10.1021/acs.est.3c08511>.

Johnson, M.R., Tyner, D.R., Szekeres, A.J., 2021. Blinded evaluation of airborne methane source detection using Bridger Photonics LiDAR. *Remote Sens. Environ.* 259, 112418. <https://doi.org/10.1016/j.rse.2021.112418>.

Kamp, J.N., Häni, C., Nyord, T., Feilberg, A., Sørensen, L.L., 2020. The aerodynamic gradient method: implications of non-simultaneous measurements at alternating heights. *Atmosphere* 11, 1067. <https://doi.org/10.3390/atmos11101067>.

Kljun, N., Calanca, P., Rotach, M.W., Schmid, H.P., 2015. A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP). *Geosci. Model Dev.* 8, 3695–3713. <https://doi.org/10.5194/gmd-8-3695-2015>.

Kormann, R., Meixner, F.X., 2001a. An analytical footprint model for non-neutral stratification. *Boundary-Layer Meteorology*. 99, 207–224. <https://doi.org/10.1023/A:1018991015119>.

LICOR, 2025. EddyPro\_Manual\_12025.pdf | Powered by Box [WWW Document]. URL <https://licor.app.boxenterprise.net/s/1ium2zmwm6hl36yz9bu4> (accessed 2.14.25).

LI-COR, Nebraska, USA, n.d. EddyPro 7 | Software Downloads [WWW Document]. URL <https://www.licor.com/support/EddyPro/software.html> (accessed 1.30.25).

Mauder, M., Foken, T., 2004. Documentation and instruction manual of the eddy covariance software package TK2. <https://epub.uni-bayreuth.de/id/eprint/884/1/ARBERG026.pdf>.

METEC, 2025. METEC | Colorado State University [WWW Document]. URL <https://metec.colostate.edu/> (accessed 2.18.25).

Moncrieff, J., Clement, R., Finnigan, J., Meyers, T., 2005. Averaging, Detrending, and Filtering of Eddy Covariance Time Series, in: Lee, X., Massman, W., Law, B. (Eds.), *Handbook of Micrometeorology, Atmospheric and Oceanographic Sciences Library*. Kluwer Academic Publishers, Dordrecht, pp. 7–31. [https://doi.org/10.1007/1-4020-2265-4\\_2](https://doi.org/10.1007/1-4020-2265-4_2).

Morin, T.H., 2019. Advances in the eddy covariance approach to CH<sub>4</sub> monitoring over two and a half decades. *J. Geophys. Res. Biogeosciences* 124, 453–460. <https://doi.org/10.1029/2018JG004796>.

Nemitz, E., Mammarella, I., Ibrom, A., Aurela, M., Burba, G.G., Dengel, S., Gielen, B., Grelle, A., Heinesch, B., Herbst, M., Hörtnagl, L., Klemetsson, L., Lindroth, A., Lohila, A., McDermitt, D.K., Meier, P., Merbold, L., Nelson, D., Nicolini, G., Nilsson, M.B., Peltola, O., Rinne, J., Zahniser, M., 2018. Standardisation of eddy-covariance flux measurements of methane and nitrous oxide. *Int. Agrophysics* 32, 517–549. <https://doi.org/10.1515/intag-2017-0042>.

Polonik, P., Chan, W.S., Billesbach, D.P., Burba, G., Li, J., Nottrott, A., Bogoev, I., Conrad, B., Biraud, S.C., 2019. Comparison of gas analyzers for eddy covariance: Effects of analyzer type and spectral corrections on fluxes. *Agric. For. Meteorol.* 272–273, 128–142. <https://doi.org/10.1016/j.agrformet.2019.02.010>.

Querino, C.A.S., Smeets, C.J.P.P., Vigano, I., Holzinger, R., Moura, V., Gatti, L.V., Martinewski, A., Manzi, A.O., De Araújo, A.C., Röckmann, T., 2011. Methane flux, vertical gradient and mixing ratio measurements in a tropical forest. *Atmospheric Chem. Phys.* 11, 7943–7953. <https://doi.org/10.5194/acp-11-7943-2011>.

Rey-Sanchez, C., Arias-Ortiz, A., Kasak, K., Chu, H., Szutu, D., Verfaillie, J., Baldocchi, D., 2022. Detecting hot spots of methane flux using footprint-weighted flux maps. *J. Geophys. Res. Biogeosciences* 127, e2022JG006977. <https://doi.org/10.1029/2022JG006977>.

Riddick, S.N., Ancona, R., Cheptonui, F., Bell, C.S., Duggan, A., Bennett, K.E., Zimmerle, D.J., 2022a. A cautionary report of calculating methane emissions using low-cost fence-line sensors. *Elem. Sci. Anthr.* 10, 00021. <https://doi.org/10.1525/elementa.2022.00021>.

Riddick, S.N., Cheptonui, F., Yuan, K., Mbua, M., Day, R., Vaughn, T.L., Duggan, A., Bennett, K.E., Zimmerle, D.J., 2022b. Estimating regional methane emission factors from energy and agricultural sector sources using a portable measurement system: Case study of the Denver–Julesburg basin. *Sensors* 22, 7410. <https://doi.org/10.3390/s22197410>.

Riddick, S.N., Mauzerall, D.L., 2023. Likely substantial underestimation of reported methane emissions from United Kingdom upstream oil and gas activities. *Energy Environ. Sci.* 16, 295–304. <https://doi.org/10.1039/D2EE03072A>.

Riddick, S.N., Mbua, M., Anand, A., Kiplimo, E., Santos, A., Upreti, A., Zimmerle, D.J., 2024a. Estimating total methane emissions from the Denver-Julesburg basin using bottom-up approaches. *Gases* 4, 236–252. <https://doi.org/10.3390/gases4030014>.

Riddick, S.N., Mbua, M., Santos, A., Hartzell, W., Zimmerle, D.J., 2024b. Potential underestimate in reported bottom-up methane emissions from oil and gas operations in the Delaware basin. *Atmosphere* 15, 202. <https://doi.org/10.3390/atmos15020202>.

Tagliaferri, F., Invernizzi, M., Capra, F., Sironi, S., 2023. Validation study of WindTrax reverse dispersion model coupled with a sensitivity analysis of model-specific settings. *Environ. Res.* 222, 115401. <https://doi.org/10.1016/j.envres.2023.115401>.

US EPA, OAR, 2023. Natural gas and petroleum systems in the GHG inventory: Additional information on the 1990-2021 GHG inventory (published April 2023) [WWW Document]. URL <https://www.epa.gov/ghgemissions/natural-gas-and-petroleum-systems-ghg-inventory-additional-information-1990-2021-ghg> (accessed 5.6.24).

Vickers, D., Mahrt, L., 1997. Quality control and flux sampling problems for tower and aircraft data. *J. Atmospheric Ocean. Technol.* 14, 512–526. [https://doi.org/10.1175/1520-0426\(1997\)014<0512:QCAFSP>2.0.CO;2](https://doi.org/10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2).

Vogel, E., Davis, K.J., Wu, K., Miles, N.L., Richardson, S.J., Gurney, K.R., Monteiro, V., Roest, G.S., Kenion, H.C.R., Horne, J.P., 2024. Using eddy-covariance to measure the effects of COVID-19 restrictions on CO<sub>2</sub> emissions in a neighborhood of Indianapolis, IN. *Carbon Manag.* 15, 2365900. <https://doi.org/10.1080/17583004.2024.2365900>.

WindTrax 2.0, n.d. thunder beach scientific [WWW Document]. URL <http://www.thunderbeachscientific.com/> (accessed 1.31.25).

Xu, L., Lin, X., Amen, J., Welding, K., McDermitt, D., 2014. Impact of changes in barometric pressure on landfill methane emission. *Glob. Biogeochem. Cycles* 28, 679–695. <https://doi.org/10.1002/2013GB004571>

- Zimmerle, D., Dileep, S., Quinn, C., 2024. Unaddressed uncertainties when scaling regional aircraft emission surveys to basin emission estimates. *Environ. Sci. Technol.* 58, 6575–6585. <https://doi.org/10.1021/acs.est.3c08972>.
- Babaeian, E., Tuller, M., 2023. Proximal sensing of evapotranspiration, in: *Encyclopedia of Soils in the Environment*. Elsevier, pp. 610–617. <https://doi.org/10.1016/B978-0-12-822974-3.00156-7>
- Bell, C., Ilonze, C., Duggan, A., Zimmerle, D., 2023. Performance of Continuous Emission Monitoring Solutions under a Single Blind Controlled Testing Protocol. *Environ. Sci. Technol.* 57, 5794–5805. <https://doi.org/10.1021/acs.est.2c09235>
- Brown, J.A., Harrison, M.R., Rufael, T., Roman White, S.A., Ross, G.B., George, F.C., Zimmerle, D., 2023. Informing Methane Emissions Inventories Using Facility Aerial Measurements at Midstream Natural Gas Facilities. *Environ. Sci. Technol.* 57, 14539–14547. <https://doi.org/10.1021/acs.est.3c01321>
- Carbon Mapper, 2024. Methane, CO2 Detection Satellite | Greenhouse Gas | Carbon Mapper [WWW Document]. Carbon Mapper. URL <https://carbonmapper.org/> (accessed 2.26.24).
- Casal, J., 2008. Chapter 6 Atmospheric dispersion of toxic or flammable clouds, in: *Industrial Safety Series*. Elsevier, pp. 195–248. [https://doi.org/10.1016/S0921-9110\(08\)80008-0](https://doi.org/10.1016/S0921-9110(08)80008-0)
- Caulton, D.R., Li, Q., Bou Zeid, E., Fitts, J.P., Golston, L.M., Pan, D., Lu, J., Lane, H.M., Buchholz, B., Guo, X., McSpirt, J., Wendt, L., Zondlo, M.A., 2018. Quantifying uncertainties from mobile laboratory derived emissions of well pads using inverse Gaussian methods. *Atmospheric Chem. Phys.* 18, 15145–15168. <https://doi.org/10.5194/acp-18-15145-2018>
- 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., Miller, B.R., 2014. Toward a better understanding and quantification of methane emissions from shale gas development. *Proc. Natl. Acad. Sci.* 111, 6237–6242. <https://doi.org/10.1073/pnas.1316546111>
- CCAC, 2024. Climate and Clean Air Coalition | Homepage | Climate & Clean Air Coalition [WWW Document]. URL <https://www.ccacoalition.org/> (accessed 8.15.24).
- Chung, T., 2021. Global Assessment: Urgent Steps Must Be Taken to Reduce Methane Emissions This Decade | UNFCCC [WWW Document]. URL <https://unfccc.int/news/global-assessment-urgent-steps-must-be-taken-to-reduce-methane-emissions-this-decade> (accessed 10.23.23).
- CO2EFFICIENT, 2022. METHANE QUANTIFICATION: TOWARD DIFFERENTIATED GAS [WWW Document]. CO2E. URL <https://co2efficient.com/methane> (accessed 2.11.24).
- Colorado State University, 2021. Mechanistic Emission Estimation Tool (MEET) [WWW Document]. Energy Inst. URL <https://energy.colostate.edu/mechanistic-air-emissions-simulator/> (accessed 2.26.24).
- Conrad, B.M., Tyner, D.R., Johnson, M.R., 2023. Robust probabilities of detection and quantification uncertainty for aerial methane detection: Examples for three airborne technologies. *Remote Sens. Environ.* 288, 113499. <https://doi.org/10.1016/j.rse.2023.113499>
- Grenna, B., 2006. An introduction to WindTrax.
- Day, R.E., Emerson, E., Bell, C., Zimmerle, D., 2024. Point Sensor Networks Struggle to Detect and Quantify Short Controlled Releases at Oil and Gas Sites. *Sensors* 24, 2419. <https://doi.org/10.3390/s24082419>
- Denmead, O.T., 2008. 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>
- Edie, R., Robertson, A.M., Field, R.A., Soltis, J., Snare, D.A., Zimmerle, D., Bell, C.S., Vaughn, T.L., Murphy, S.M., 2020. Constraining the accuracy of flux estimates using OTM-33A. *Atmospheric Meas. Tech.* 13, 341–353. <https://doi.org/10.5194/amt-13-341-2020>
- Foster Wittig, T.A., Thoma, E.D., Albertson, J.D., 2015. Estimation of point source fugitive emission rates from a single sensor time series: A conditionally-sampled Gaussian plume reconstruction. *Atmos. Environ.* 115, 101–109. <https://doi.org/10.1016/j.atmosenv.2015.05.042>
- Heimbürger, A.M.F., Harvey, R.M., Shepson, P.B., Stirm, B.H., Gore, C., Turnbull, J., Cambaliza, M.O.L., Salmon, O.E., Kerlo, A. E.M., Lavoie, T.N., Davis, K.J., Lauvaux, T., Karion, A., Sweeney, C., Brewer, W.A., Hardesty, R.M., Gurney, K.R., 2017. Assessing the optimized precision of the aircraft mass balance method for measurement of urban greenhouse gas emission rates through averaging. *Elem. Sci. Anthr.* 5, 26. <https://doi.org/10.1525/elementa.134>
- Hutchinson, M., Oh, H., Chen, W. H., 2017. A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Inf. Fusion* 36, 130–148. <https://doi.org/10.1016/j.inffus.2016.11.010>

- Honze, C., Emerson, E., Duggan, A., Zimmerle, D., 2024. Assessing the Progress of the Performance of Continuous Monitoring Solutions under a Single Blind Controlled Testing Protocol. *Environ. Sci. Technol.* 58, 10941–10955. <https://doi.org/10.1021/acs.est.3c08511>
- Jia, M., Daniels, W., Hammerling, D., 2023. Comparison of the Gaussian plume and puff atmospheric dispersion models on oil and gas facilities (preprint). *Chemistry*. <https://doi.org/10.26434/chemrxiv-2023-hc95q>
- Kahl, J.D.W., Chapman, H.L., 2018. Atmospheric stability characterization using the Pasquill method: A critical evaluation. *Atmos. Environ.* 187, 196–209. <https://doi.org/10.1016/j.atmosenv.2018.05.058>
- Kamp, J.N., Häni, C., Nyord, T., Feilberg, A., Sørensen, L.L., 2020. The Aerodynamic Gradient Method: Implications of Non-Simultaneous Measurements at Alternating Heights. *Atmosphere* 11, 1067. <https://doi.org/10.3390/atmos11101067>
- Kljun, N., Calanca, P., Rotach, M.W., Schmid, H.P., 2015. A simple two-dimensional parameterisation for Flux Footprint Prediction (FFP). *Geosci. Model Dev.* 8, 3695–3713. <https://doi.org/10.5194/gmd-8-3695-2015>
- METEC | Colorado State University [WWW Document], 2024. URL <https://metec.colostate.edu/> (accessed 8.26.24).
- Morin, T.H., 2019. Advances in the Eddy Covariance Approach to CH<sub>4</sub> Monitoring Over Two and a Half Decades. *J. Geophys. Res. Biogeosciences* 124, 453–460. <https://doi.org/10.1029/2018JG004796>
- Nemitz, E., Mammarella, I., Ibrom, A., Aurela, M., Burba, G.G., Dengel, S., Gielen, B., Grelle, A., Heinesch, B., Herbst, M., Hörtnagl, L., Klemetsson, L., Lindroth, A., Lohila, A., McDermitt, D.K., Meier, P., Merbold, L., Nelson, D., Nicolini, G., Nilsson, M.B., Peltola, O., Rinne, J., Zahniser, M., 2018. Standardisation of eddy covariance flux measurements of methane and nitrous oxide. *Int. Agrophysics* 32, 517–549. <https://doi.org/10.1515/intag-2017-0042>
- Prueger, J.H., Kustas, W.P., 2015. Aerodynamic Methods for Estimating Turbulent Fluxes, in: Hatfield, J.L., Baker, J.M. (Eds.), *Agronomy Monographs*. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, WI, USA, pp. 407–436. <https://doi.org/10.2134/agronmonogr47.c18>
- Querino, C.A.S., Smeets, C.J.P.P., Vigano, I., Holzinger, R., Moura, V., Gatti, L.V., Martinewski, A., Manzi, A.O., De Araújo, A.C., Röckmann, T., 2011. Methane flux, vertical gradient and mixing ratio measurements in a tropical forest. *Atmospheric Chem. Phys.* 11, 7943–7953. <https://doi.org/10.5194/acp-11-7943-2011>
- Rey-Sanchez, C., Arias-Ortiz, A., Kasak, K., Chu, H., Szutu, D., Verfaillie, J., Baldocchi, D., 2022. Detecting Hot Spots of Methane Flux Using Footprint Weighted Flux Maps. *J. Geophys. Res. Biogeosciences* 127, e2022JG006977. <https://doi.org/10.1029/2022JG006977>
- Riddick, S.N., Ancona, R., Cheptonui, F., Bell, C.S., Duggan, A., Bennett, K.E., Zimmerle, D.J., 2022a. A cautionary report of calculating methane emissions using low-cost fence-line sensors. *Elem. Sci. Anthr.* 10, 00021. <https://doi.org/10.1525/elementa.2022.00021>
- Riddick, S.N., Cheptonui, F., Yuan, K., Mbua, M., Day, R., Vaughn, T.L., Duggan, A., Bennett, K.E., Zimmerle, D.J., 2022b. Estimating Regional Methane Emission Factors from Energy and Agricultural Sector Sources Using a Portable Measurement System: Case Study of the Denver–Julesburg Basin. *Sensors* 22, 7410. <https://doi.org/10.3390/s22197410>
- Riddick, S.N., Mauzerall, D.L., 2023. Likely substantial underestimation of reported methane emissions from United Kingdom upstream oil and gas activities. *Energy Environ. Sci.* 16, 295–304. <https://doi.org/10.1039/D2EE03072A>
- Riddick, S.N., Mauzerall, D.L., Celia, M., Allen, G., Pitt, J., Kang, M., Riddick, J.C., 2020. The calibration and deployment of a low-cost methane sensor. *Atmos. Environ.* 230, 117440. <https://doi.org/10.1016/j.atmosenv.2020.117440>
- Riddick, S.N., Mbua, M., Anand, A., Kiplimo, E., Santos, A., Upreti, A., Zimmerle, D.J., 2024a. Estimating Total Methane Emissions from the Denver–Julesburg Basin Using Bottom-Up Approaches. *Gases* 4, 236–252. <https://doi.org/10.3390/gases4030014>
- Riddick, S.N., Mbua, M., Santos, A., Hartzell, W., Zimmerle, D.J., 2024b. Potential Underestimate in Reported Bottom-up Methane Emissions from Oil and Gas Operations in the Delaware Basin. *Atmosphere* 15, 202. <https://doi.org/10.3390/atmos15020202>
- R.M. Young Company, 2023. . Ultrason. Anemometer Model 81000. URL <https://www.youngusa.com/wp-content/uploads/2008/01/81000-9028I29.pdf>
- Shaw, J.T., Shah, A., Yong, H., Allen, G., 2021. Methods for quantifying methane emissions using unmanned aerial vehicles: a review. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* 379, 20200450. <https://doi.org/10.1098/rsta.2020.0450>
- Sonderfeld, H., Bösch, H., Jeanjean, A.P.R., Riddick, S.N., Allen, G., Ars, S., Davies, S., Harris, N., Humpage, N., Leigh, R., Pitt, J., 2017. CH<sub>4</sub> emission estimates from an active landfill site

inferred from a combined approach of CFD modelling and in situ FTIR measurements. *Atmospheric Meas. Tech.* 10, 3931–3946. <https://doi.org/10.5194/amt-10-3931-2017>

Stull, R.B., 1988. *An Introduction to Boundary Layer Meteorology*, 1988 edition. ed, Atmospheric and Oceanographic Sciences Library. Springer Nature, Dordrecht. <https://doi.org/10.1007/978-94-009-3027-8>

UN Environment Programme, 2024. UNEP—UN Environment Programme [WWW Document]. URL <https://www.unep.org/node> (accessed 8.15.24).

United Nations Climate Change, 2015. Adoption of the Paris Agreement. Proposal by the President. | UNFCCC [WWW Document]. URL <https://unfccc.int/documents/9064> (accessed 8.15.24).

US EPA, 2013. Standard Operating Procedure for Analysis of US EPA Geospatial Measurement of Air Pollution Remote Emission Quantification by Direct Assessment (GMAP-REQDA) Method Data for Methane Emission Rate Quantification using the Point Source Gaussian Method [WWW Document]. URL [https://www.epa.gov/sites/default/files/2020-08/documents/otm\\_33a\\_appendix\\_f1\\_psg\\_analysis\\_sop.pdf](https://www.epa.gov/sites/default/files/2020-08/documents/otm_33a_appendix_f1_psg_analysis_sop.pdf) (accessed 5.6.24).

US EPA, OAR, 2023. Natural Gas and Petroleum Systems in the GHG Inventory: Additional Information on the 1990–2021 GHG Inventory (published April 2023) [WWW Document]. URL <https://www.epa.gov/ghgemissions/natural-gas-and-petroleum-systems-ghg-inventory-additional-information-1990-2021-ghg> (accessed 5.6.24).

US EPA, OA, 2023. Methane Emissions Reduction Program [WWW Document]. URL <https://www.epa.gov/inflation-reduction-act/methane-emissions-reduction-program> (accessed 1.16.24).

US EPA, O., 2016. Understanding Global Warming Potentials [WWW Document]. URL <https://www.epa.gov/ghgemissions/understanding-global-warming-potentials> (accessed 3.18.24).

Vogel, E., Davis, K.J., Wu, K., Miles, N.L., Richardson, S.J., Gurney, K.R., Monteiro, V., Roest, G.S., Kenion, H.C.R., Horne, J.P., 2024. Using eddy covariance to measure the effects of COVID-19 restrictions on CO<sub>2</sub> emissions in a neighborhood of Indianapolis, IN. *Carbon Manag.* 15, 2365900. <https://doi.org/10.1080/17583004.2024.2365900>

Xu, L., Lin, X., Amen, J., Welding, K., McDermitt, D., 2014. Impact of changes in barometric pressure on landfill methane emission. *Glob. Biogeochem. Cycles* 28, 679–695. <https://doi.org/10.1002/2013GB004571>

Zimmerle, D., Dileep, S., Quinn, C., 2024. Unaddressed Uncertainties When Scaling Regional Aircraft Emission Surveys to Basin Emission Estimates. *Environ. Sci. Technol.* 58, 6575–6585. <https://doi.org/10.1021/acs.est.3c08972>

Zinke, J., Nilsson, E.D., Markuszewski, P., Zieger, P., Mårtensson, E.M., Rutgersson, A., Nilsson, E., Salter, M.E., 2024. Sea spray emissions from the Baltic Sea: comparison of aerosol eddy covariance fluxes and chamber-simulated sea spray emissions. *Atmospheric Chem. Phys.* 24, 1895–1918. <https://doi.org/10.5194/acp-24-1895-2024>