# Evaluating the feasibility of using downwind methods to quantify point source oil and gas emissions using continuous monitoring fence-line sensors

Mercy Mbua*[,1], Stuart N. Riddick[1,2], Elijah Kiplimo[1], Kira B. Shonkwiler[1], Anna Hodshire[1] and Daniel Zimmerle[1]

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

[2]*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 reporting of methane ($CH_4$) emissions from point sources, such as fugitive leaks from oil and gas infrastructure, is important for-profit maximization (retaining more hydrocarbons), evaluating climate impacts, assessing $CH_4$ fees for regulatory programs, and validating $CH_4$ 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 $CH_4$ quantification methods using $CH_4$ 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 downwind methods' accuracy, this study uses fence-line simulated data collected during controlled release experiments as input for a non-standard closed-path eddy covariance (EC), the Gaussian plume inverse method (GPIM), and the backward Lagrangian stochastic (bLs) model in a range of atmospheric conditions. Generally, the closed-path EC system used in this study proved generally unreliable and largely underestimated emissions, primarily due to non-stationarity and study limitations associated with using a non-standard setup. In comparison, the Gaussian Plume Inverse Method (GPIM) consistently outperformed the EC system for both single-release and multi-release single-point emissions. The backward Lagrangian stochastic (bLs) model demonstrated the highest accuracy for single-release single-point emissions, though it exhibited greater uncertainty than GPIM under multi-release conditions. Across GPIM and bLs models, the most reliable quantification was achieved with a 15-minute averaging and a narrow 5° wind-sector range. Although EC was limited in this context, future studies should consider employing a standard EC system and further optimizing GPIM and bLs approaches—particularly complex multi-source scenarios—to enhance quantification accuracy and reduce uncertainty.

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

## 1 Introduction

Reducing methane ($CH_4$) 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) (UNEP, 2024). 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 $CH_4$ emitted from US oil and gas production are compiled by the US Environmental Protection Agency (EPA) under Subpart W. Typically, companies use a bottom-up inventory approach where emission factors ($CH_4$ 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 (US EPA, 2023) to generate emissions. This quantification approach has several shortcomings, including: 1. It separately calculates $CH_4$ 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 (Riddick et al., 2024b); 4. Difficulty in obtaining a truly representative sample from a large diverse population to generate emission factors (Allen, 2014); and possibly unreliable data reported by operators (Chan et al., 2024). Recently, mechanistic models, such as the Mechanistic Air Emissions Simulator (MAES), have been developed to address shortcomings in bottom-up $CH_4$ reporting (Mollel et al., 2025), 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 ~ 2 kg $h^{-1}$) (Johnson et al., 2021) and satellites such as Carbon Mapper (predicted 90% detection limit of about 100 kg $h^{-1}$) (Carbon Mapper - Science & Technology, 2025) can also be used to infer emissions. However, these survey methods only quantify emissions over a very short period of time (< 10 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 $CH_4$ emissions quantification has been a complex issue.

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; Riddick et al., 2022a). Detection and localization of simulated fugitive emissions using this approach have been demonstrated successfully in controlled release studies. For example, Ilonze et al. (2024) reported a 90% probability of detection for emissions between 3.9 and 18.2 kg $CH_4$ $h^{-1}$ using multi-sensor and scanning/imaging systems. However, significant uncertainty in quantification remains; their study reported emissions being misestimated by a factor of 0.2 to 42 for releases between 0.1 and 1 kg $CH_4$ $h^{-1}$, and by a factor of 0.08 to 18 for emissions above 1 kg $CH_4$ $h^{-1}$. While informative, the methods in Ilonze et al. (2024) differ in keyways from those employed here—specifically, their use of multiple sensors and a distributed monitoring configuration as opposed to the single-instrument, fence-line-based framework used in our study—limiting direct comparison of quantification accuracy. This study will evaluate the quantification accuracy of the closed-path EC, Gaussian plume inverse model (GPIM), and the backward Lagrangian

stochastic model (bLs) for oil and gas point source quantification using a single-instrument deployed at a fence line distance.

Eddy covariance is a vertical flux gradient measurement that measures $CH_4$ emissions based on the covariance between $CH_4$ concentrations measured using a fast-response analyzer (> 10 Hz) and vertical wind vector measured by a fast-response sonic anemometer (>10 Hz) (Figure 1; 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 that 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 $CH_4$ source at 0.8 m, programmed to emit when winds were coming from the source direction (± 45º), and when friction velocity ($u_*$) was above 0.13 m $s^{-1}$. 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, 2001) 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.

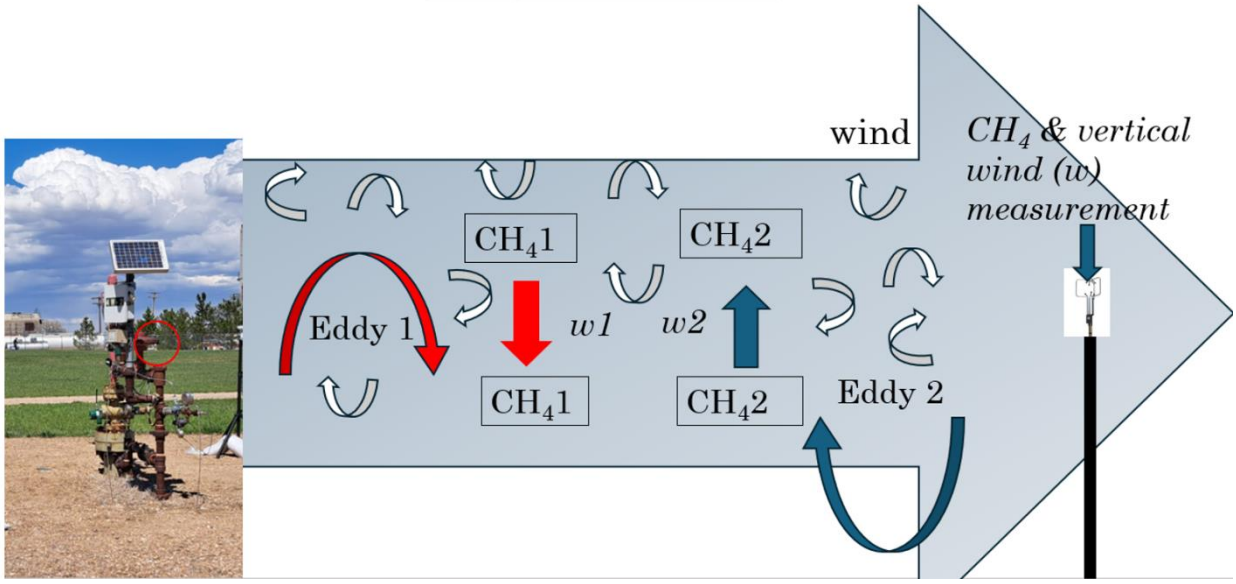


Figure 1: Illustrations of eddy covariance where $CH_4$ is methane concentration, and *w* is the vertical wind speed.

The GPIM method calculates the CH4 emission rate as a function of mole fraction at a point in space (x, y, z), downwind distance, perpendicular distance (crosswind), mean wind speed, and atmospheric stability (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% in a controlled release experiment involving 10 replicate measurements of compressed natural gas (1.5 m release height), with concentrations measured using a mobile vehicle survey. While this differs from continuous fence-line deployment, it offers insight into the inherent uncertainty of the GPIM method in field conditions. 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 are that 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).

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.

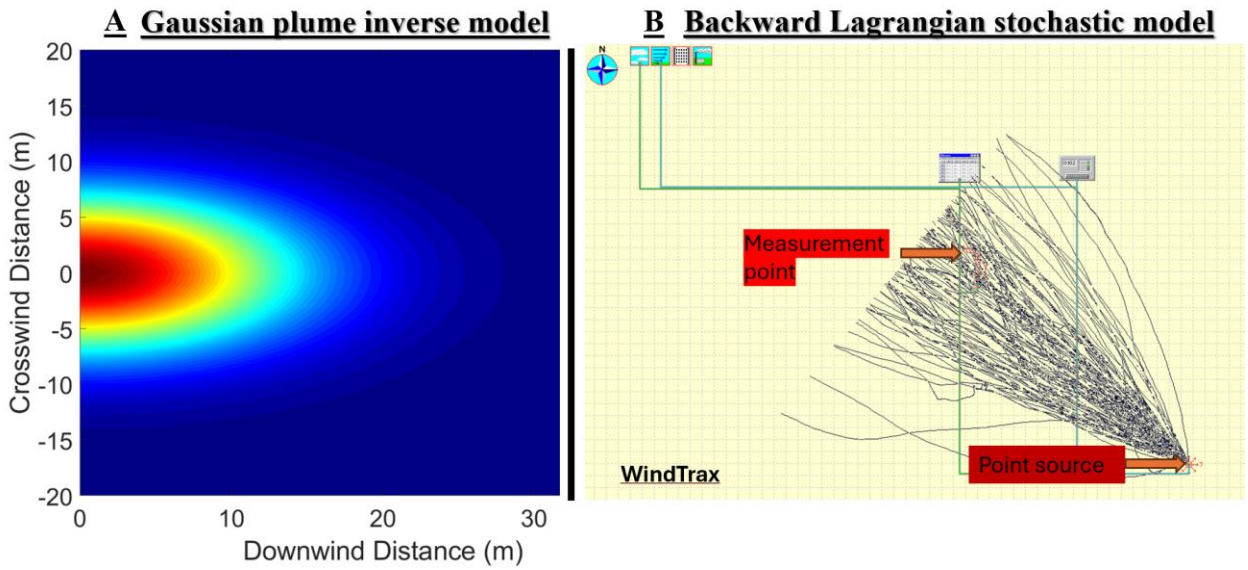


Figure 2. A: An illustration of a plume that follows a Gaussian plume inverse model where emission rate can be inferred from concentrations at different downwind distances and crosswind distances. B: An illustration of how the backward Lagrangian stochastic model traces particles to the source.

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 $CH_4$ 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. Using robust releases and environmental conditions, 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 $CH_4$ reporting. Specifically, the study will evaluate the overall quantification accuracy (linear regression slope of estimated versus actual emissions, and $R^2$) of closed-path EC, bLs model, and the GPIM method in quantifying single-release single-point and multi-release single-point emissions that simulate oil and gas emissions.

## 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, 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 | Colorado State University, 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 $s^{-1}$ and temperatures ranged between -15 and +19 °C (Supplementary Information Section 1). A stationary mast holding the instrumentation was 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 3A; Day et al., 2024; Riddick et al., 2022a). Fence-line sensors are typically placed within the oil and gas perimeter (~30 m). This study collected data for what we considered as 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 which measured micrometeorology at 10 Hz (Figure 3B). The northward, eastward, and vertical separations of the inlet tubing from the sonic anemometer were 0, 0, -10 cm, respectively.

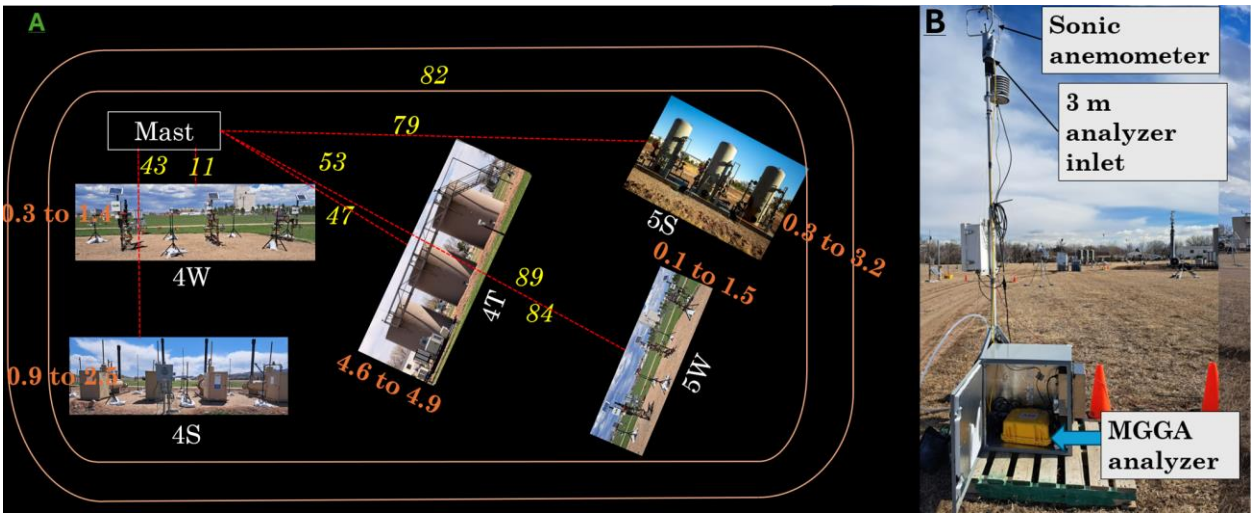


Figure 3: A: Map illustration of major pieces of equipment and the measurements points at Colorado State University's Methane Emissions Technology Evaluation Center (METEC) in Fort Collins, CO, USA. Equipment 4S denotes horizontal separators, 4W are well heads, 4T are tanks, 5S are vertical separators, and 5W are well heads. B 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 red dotted lines with yellow numbers show the average distances (meters) between emission equipment and measurement point. The orange numbers show the range of emission heights (meters) for each equipment. The analyzer was hosted in a temperature-controlled box.

### 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 (Advancing Development of Emissions Detection (ADED)). Natural gas of known $CH_4$ 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 0.01 kg $h^{-1}$ and 8.7 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 at night. The distance from the release points to the measurement points ranged between 9 and 94 m, and emission heights were between 0.1 and 4.9 m (Figure 3A). Emission points simulate the realistic size and locations of typical emissions from components such as 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 (z0) 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 z0 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 $CH_4$ data showed that actual sampling was between 4 and 12 Hz (majority of the data collected at approximately 6 Hz), even though the analyzer had been configured to sample at 10 Hz (Supplementary Information Section 2b). To account for this sampling variability, data were filtered to when sampling was equal to or greater than 8 Hz. Data sampled at frequencies above 8 Hz were down sampled to 8 Hz. The 8 Hz frequency threshold was selected to ensure uniform sampling, enough data for model evaluation as most sampling was at lower frequencies, and to preserve as much temporal resolution as possible given the system limitations. 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 $CH_4$ clock time as a reference, meteorological data from the sonic anemometer were matched to the MGGA $CH_4$ data using linear interpolation to generate concentration-meteorological 8 Hz data. While in an ideal circumstance of a fast pump and short tube length a correct timeseries matching can be achieved through establishing a clear point of maximum covariance when determining the time lag, this is difficult for our system due to a 3 lpm pump flowrate and a 3 m tubing that caused both attenuation and time lag.

The aggregated concentration-meteorological data were then merged with METEC's release data and metadata, and release event tables were 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, 15 and 30-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, ±10º, ±20º, and ±45º.

#### 2.4.1.2 Flux calculation

Turbulent fluxes were calculated using the open software EddyPro® version 7 (EddyPro 7 | Software Downloads, 2025). Acquisition frequency was set at 8 Hz, while file duration and the flux interval were set at 5, 15, and 30 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

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

In raw data processing, axis rotations for tilt correction under wind speed measurement offsets were selected. 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 enabled. Compensation for density fluctuations (Webb-Pearman-Leuning terms) (Webb et al., 1980) was disabled as the MGGA analyzer synchronously reported dry $CH_4$ and water mole fractions, cell temperature and pressure. Mauder and Foken (2004) (0-1-2 system) were used for quality checks. 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 selected. 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**

Flux data were flagged "2", low quality, based on (Mauder and Foken, 2004) (0-1-2) quality system. Cospectral analysis revealed that the EC system in this study smoothed out low-frequency eddies, as the cospectra lack the ideal shape characterized by a low-frequency rise, a peak region, and a high-frequency decay (Supplementary Information Section 2c.i). While the slope in the high-frequency region varies around the theoretical −4/3 slope, the cospectral data followed the 1:1 line, indicating consistent spectral shape across sampling periods. We also examined the relationship between $CH_4$ flux and friction velocity ($u_*$) to identify a u threshold below which flux estimates may be unreliable (Supplementary Information Section 2c. ii). However, no consistent relationship was observed across atmospheric stability classes (unstable, stable, and neutral). $CH_4$ fluxes varied widely—including both positive and negative values—across the full range of $u_*$ (~0 to 1 m $s^{-1}$), with no discernible threshold beyond which fluxes stabilized. This indicated that $CH_4$ fluxes were effectively independent of $u_*$, and thus, data from all $u_*$ values were retained. Ogive analysis was conducted to assess whether averaging durations of 5, 15, and 30 minutes were sufficient

for capturing the full turbulent flux. The resulting ogive curves deviated from the ideal asymptotic shape, particularly at the highest and lowest frequencies. Notably, the curves did not exhibit a clear plateau near the low-frequency end, where cumulative flux should approach unity. This indicates incomplete flux capture. Furthermore, the similarity in ogive shapes across different frequencies—mirroring patterns seen in the cospectra—suggests a lack of significant turbulent contributions and the influence of non-turbulent, possibly advective, processes. These results imply that the EC system may not have fully resolved the flux due to either insufficient averaging time, non-stationarity, or instrument-related limitations (Supplementary Information Section 2c.iii). As positive fluxes are generally considered emissions, and negative fluxes depositions, data were further filtered for positive fluxes which were then quantified to emission rates.

#### 2.4.1.4 Footprints Calculation

Eddy covariance footprints were calculated using the Kljun et al. (2015) and the Kormann and Meixner (2001) footprint models. For the Kljun et al. (2015), the freely online MATLAB code of the model was used, while the Kormann and Meixner (2001) was coded in MATLAB. To determine the point source footprint contribution, the study first calculated the area that contributed 90% of the vertical flux; and based on the location (x and y coordinates based on wind direction and distance from source) 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.

### 2.4.2 Gaussian Plume Inverse Method

#### 2.4.2.1 Data pre-processing

Methane concentration data from the MGGA analyzer and meteorology data from the sonic anemometer were averaged to 1 Hz and aggregated. The aggregated concentration-meteorological data were merged with METEC's release data and metadata, and release event tables were 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, 15, and 30-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, ±5º, ±10º and ±20º wind sector ranges. 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). This study focuses on single-release single-point and multi-release single-point emissions. For continuous monitoring sensors, background concentration can be determined from $CH_4$ 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 $CH_4$ in air especially during stable conditions, and capturing sensors drift in the downwind sensor. In this study, background $CH_4$ was calculated as the average of the lowest $5^{th}$ percentile, 5 minutes before each release started.

In cases where this background was greater than the mean $CH_4$ concentration in the quantifying duration, the minimum $CH_4$ concentration for that duration was used as the background. Methane enhancement was then calculated as $CH_4$ concentration minus the background.

#### 2.4.2.2 Quantification

The GPIM was evaluated under six scenarios (two equations and three different dispersion coefficients generations) using single-release single-point emissions to test when the model works best (Supplementary Information Section 2a: Equation 7 and 8). Dispersion coefficients were generated based on (1) high frequency sonic anemometer data at ~ 10 Hz, (2) EPA point-source dispersion coefficients (US EPA, 2013), and (3) 1 Hz sonic anemometer data. The scenario with the slope closest to 1, and highest $R^2$ across averaging durations, and wind sector ranges were selected and used for multi-release single-point emissions 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. 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 multi-release single-point emissions.

### 2.4.3 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-, 15- and 30-minute duration in the ±5º, ±10º, and ±20º, respectively, inputs included roughness length (z0), Monin-Obukhov length (*L-Supplementary Information Equation 3*), 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

### 3.1Eddy Covariance

#### 3.1.1 Single-Release Single-Point

For single-release single-point (SRSP) emissions, the closed-path EC underestimated emissions. Using the Kljun et al. (2015) footprint model, the slope of the estimated emissions versus actual emissions linear regression was between -0.04 and 0.54 at 5 minutes, -0.36 and -0.04 at 15 minutes, and 0.03 at 30 minutes, 45 degrees (10 and 20 degrees had insufficient data points) (Figure 4). The adjusted $R^2$ was between -0.04 and 0.12 indicating no linear relationship between the estimated and actual emission (Figure 4). Using the Kormann and Meixner (2001) footprint model, the slope was between -0.42 and 0.17 at 5 and 15 minutes, respectively, and -0.08 at 30 minutes, 45 degrees (Supplementary Information Section 3.1.1). Similarly, the adjusted $R^2$ values were between -0.07 and 0.05. These results indicate that this study's EC system using either the Kljun et al. (2015) or the Kormann and Meixner (2001)

footprint models did not reliably quantify emissions for SRSP cases. The low slopes and adjusted $R^2$ values suggest little to no linear relationship between estimated and actual emissions under the tested conditions (Figure 4; Supplementary Section 3.1.1).

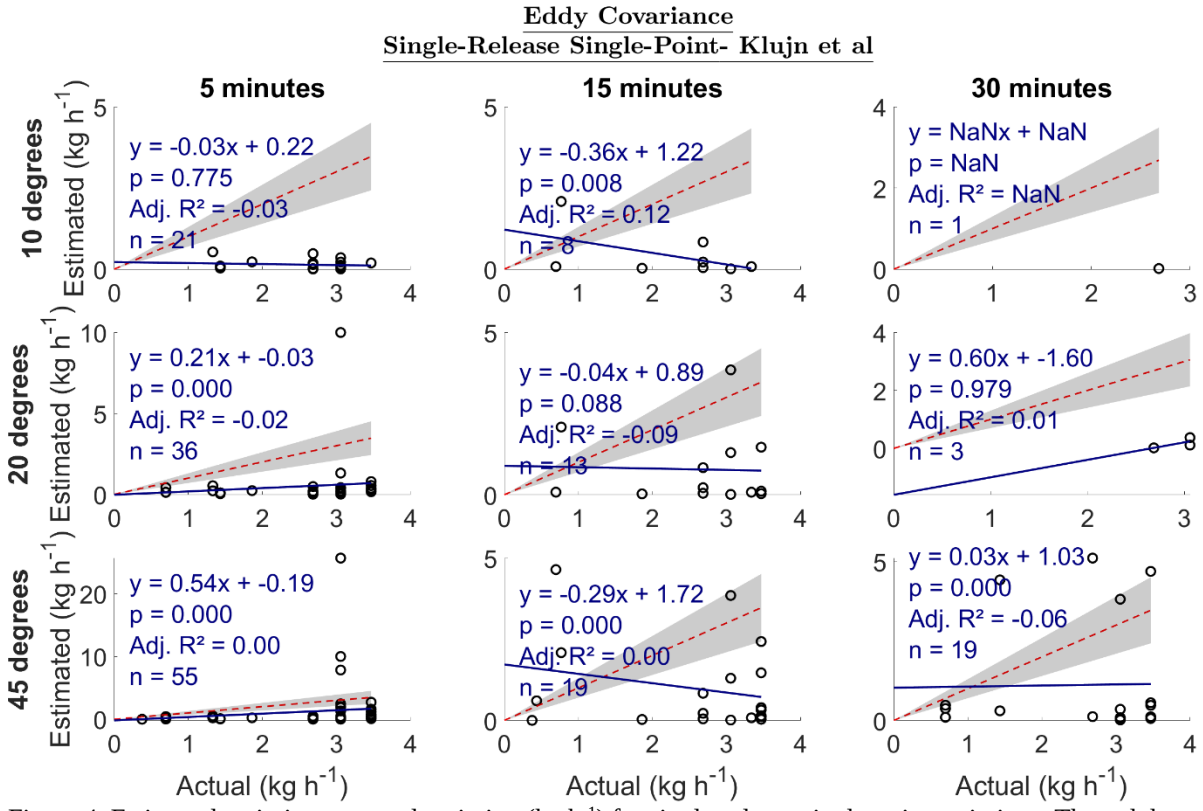


Figure 4. Estimated emission vs actual emission (kg $h^{-1}$) for single-release single-point emissions. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents ±30% of the actual emission. Adj. $R^2$ is the adjusted $R^2$. The sample size is n.

### 3.1.2 Multi-Release Single-Point

For multi-release single-point (MRSP) emissions, the closed-path EC largely underestimated emissions and did not show good agreement between estimated and actual emissions (Figure 5). Using the Kljun et al. (2015) footprint model, the slope was between -0.51 and 0.18, except for the 5 minutes 45 degrees category that had a slope of 0.61. The adjusted $R^2$ did not show a linear relationship between estimated and actual emissions with values ranging between -0.02 and 0.00 (Figure 5). Using Kormann and Meixner (2001) footprint model, the slope was between -0.03 and 1.06. The good agreement of 1.06 was at 30 minutes 45 degrees with an adjusted $R^2$ of 0.12 (Supplementary Information Section 3.1.2). The rest of the categories had an $R^2$ of between -0.02 and 0.06. These results suggest that the EC system did not reliably quantify emissions for MRSP cases under most conditions. Only one category (30 minutes, 45 degrees using the Kormann and Meixner (2001) footprint model showed moderate agreement (slope = 1.06, adjusted $R^2$ = 0.12), but even this explains only a small portion of the variability in actual emissions. Overall, the adjusted $R^2$ values across scenarios (−0.02 to 0.12) indicate a weak or no linear relationship.

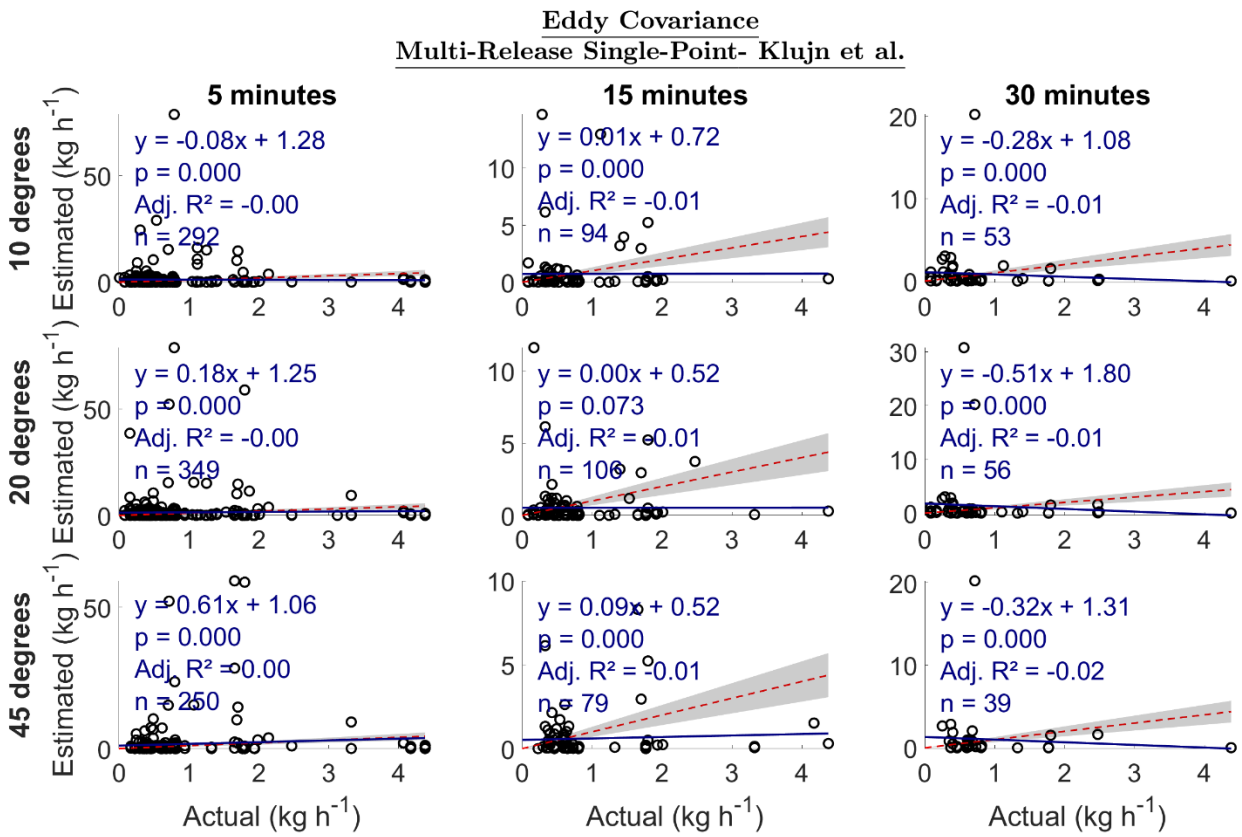


Figure 5. Estimated emission vs actual emission (kg $h^{-1}$) for multi-release single-point emissions. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents ±30% of the actual emission. Adj. $R^2$ is the adjusted $R^2$. The sample size is n.

### 3.2 Gaussian Plume Inverse Method

#### 3.2.1 Single-Release Single-Point

The GPIM sensitivity analysis comparing different equations and dispersion coefficients showed no difference in quantified emissions between Equation 7 and Equation 8. Among the dispersion coefficient sets tested, the (US EPA, 2013) coefficients resulted in the most consistent performance, with the least variability in slope and the highest overall adjusted $R^2$ values (Supplementary Information Section 3.2). For this scenario, the slope ranged from 1.65 to 3.92 (excluding the 30-minute, 5-degree case due to insufficient data), and adjusted $R^2$ values ranged from 0.40 to 0.64 (Figure 6). The 15-minute, 5-degree case had the slope closest to 1 (slope = 1.65, $R^2$ = 0.40), while the 5-minute, 5-degree case showed the strongest linear relationship overall (slope = 2.42, $R^2$ = 0.64) (Figure 6). These results suggest that while the GPIM model tends to overestimate emissions (slopes > 1), it provides relatively consistent and stronger linear agreement with actual emissions compared to the closed-path EC system tested above (Section 3.1).

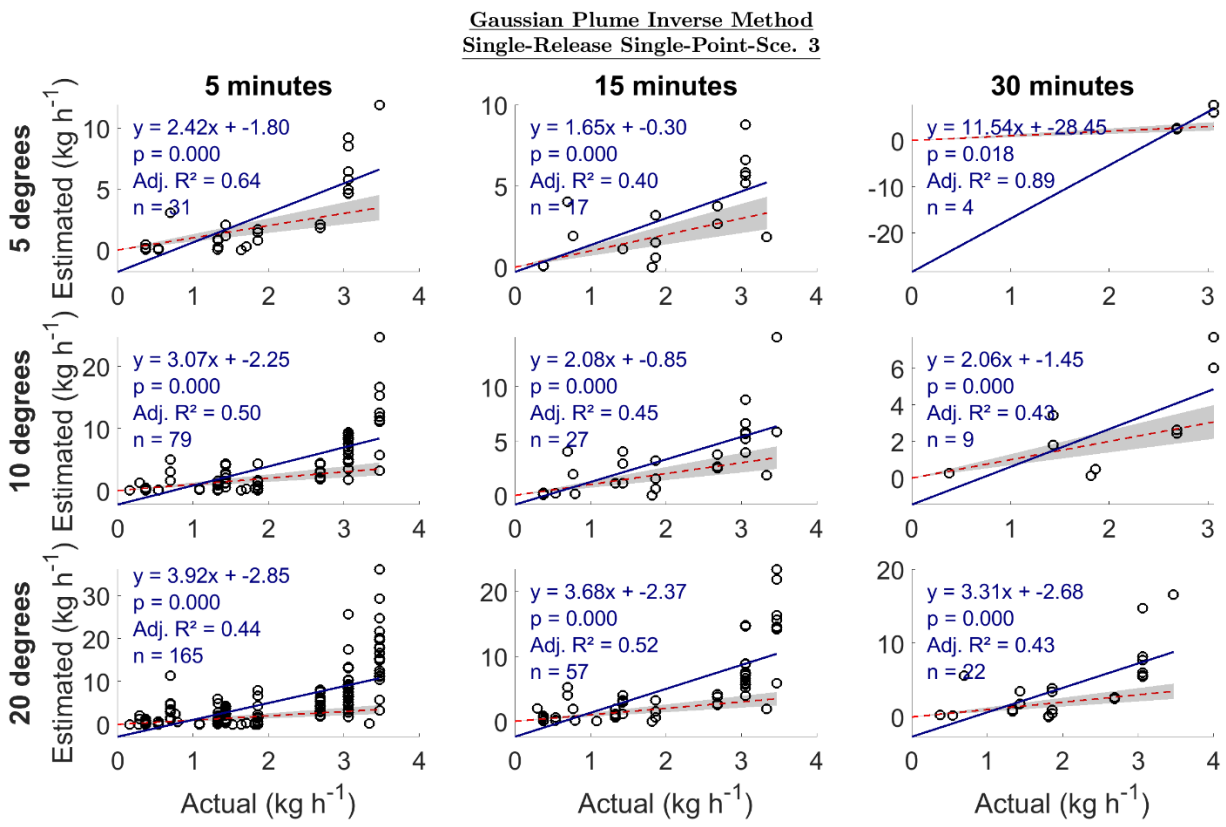


Figure 6. Estimated emission vs actual emission (kg $h^{-1}$) for single-release single-point emissions. Sce.3 refers to scenario 3 of the sensitivity analysis in Supplementary Information Section 3.2). The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents ±30% of the actual emission. Adj. $R^2$ is the adjusted $R^2$. The sample size is n.

### 3.2.2 Multi-Release Single-Point

For MRSP emissions, the GPIM model produced a wide range of slopes, from –43.05 to 1.60 (Figure 7). Excluding the 5-minute 5- and 10 degrees categories, and the 30-minute 10 degrees category, most other cases reported slopes between 0.76 and 1.22, suggesting potential quantification within ~25% of actual emissions. However, the adjusted $R^2$ values in these cases were close to zero, indicating no consistent linear relationship between estimated and actual emissions (Figure 7). This lack of correlation is likely due to the high variability in GPIM estimates, which reached up to 200 kg $h^{-1}$ for the selected categories, despite actual emissions being only around ~6 kg $h^{-1}$. These results indicate that while the GPIM model sometimes produced slope values suggesting close agreement with actual emissions, the lack of linear correlation and large overestimations highlight its limited reliability in quantifying MRSP emissions accurately.

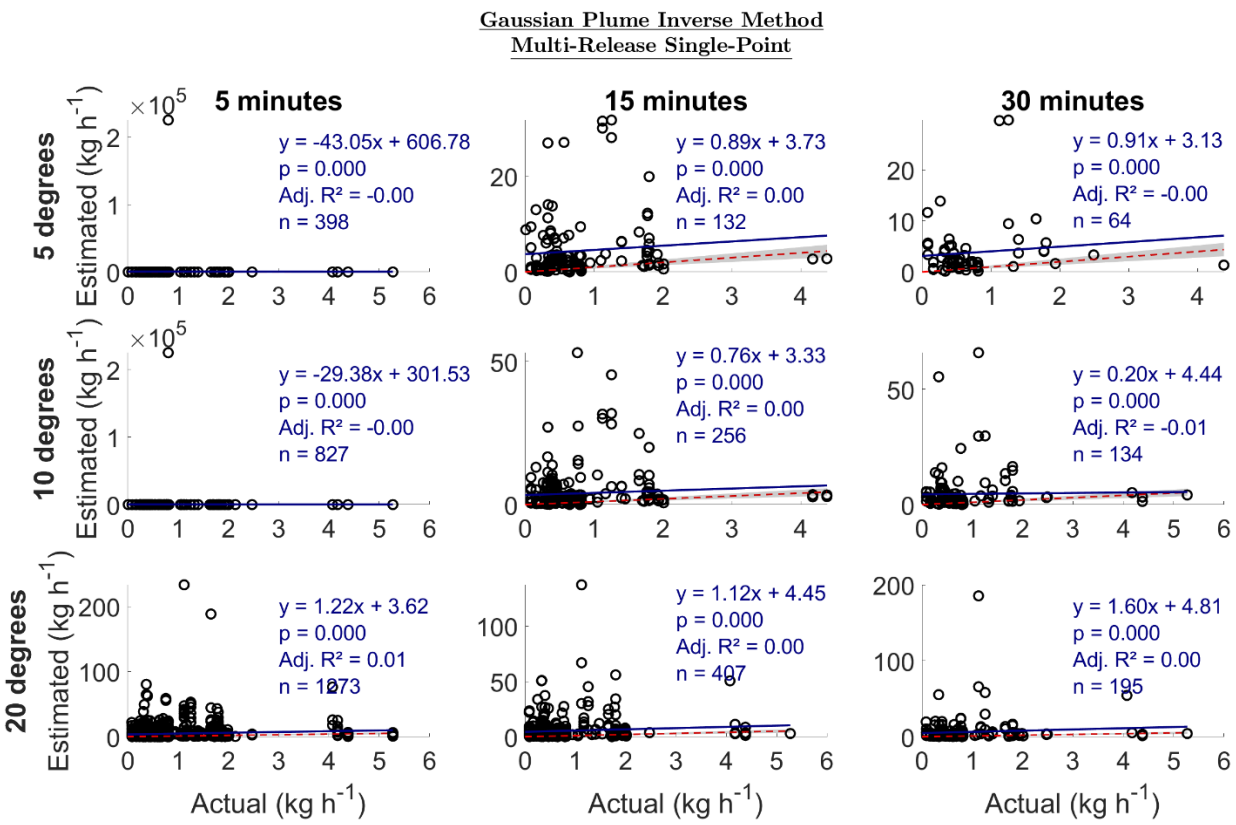


Figure 7. Estimated emission vs actual emission (kg $h^{-1}$) for multi-release single-point emissions. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents ±30% of the actual emission. Adj. $R^2$ is the adjusted $R^2$. The sample size is n.

### 3.3 Backward Lagrangian Stochastic Model

### 3.3.1 Single-Release Single-Point

For SRSP emissions, the bLS method generally produced the most accurate slopes (i.e., closest to 1) compared to the EC and GPIM methods (Figure 8). The 15-minute 5- and 10-degrees categories yielded the most accurate estimates, with slopes of 1.05 and 1.10, respectively. The adjusted $R^2$ ranged from 0.48 to 0.66 for the 5-minute averaging duration, and from 0.40 to 0.48 for the 15-minute duration. At the 30-minute averaging duration, performance improved in the 10- and 20-degrees categories, likely due to increased sample sizes (Figure 8). These results show that the bLs is more suitable for quantifying single-release single point emissions.

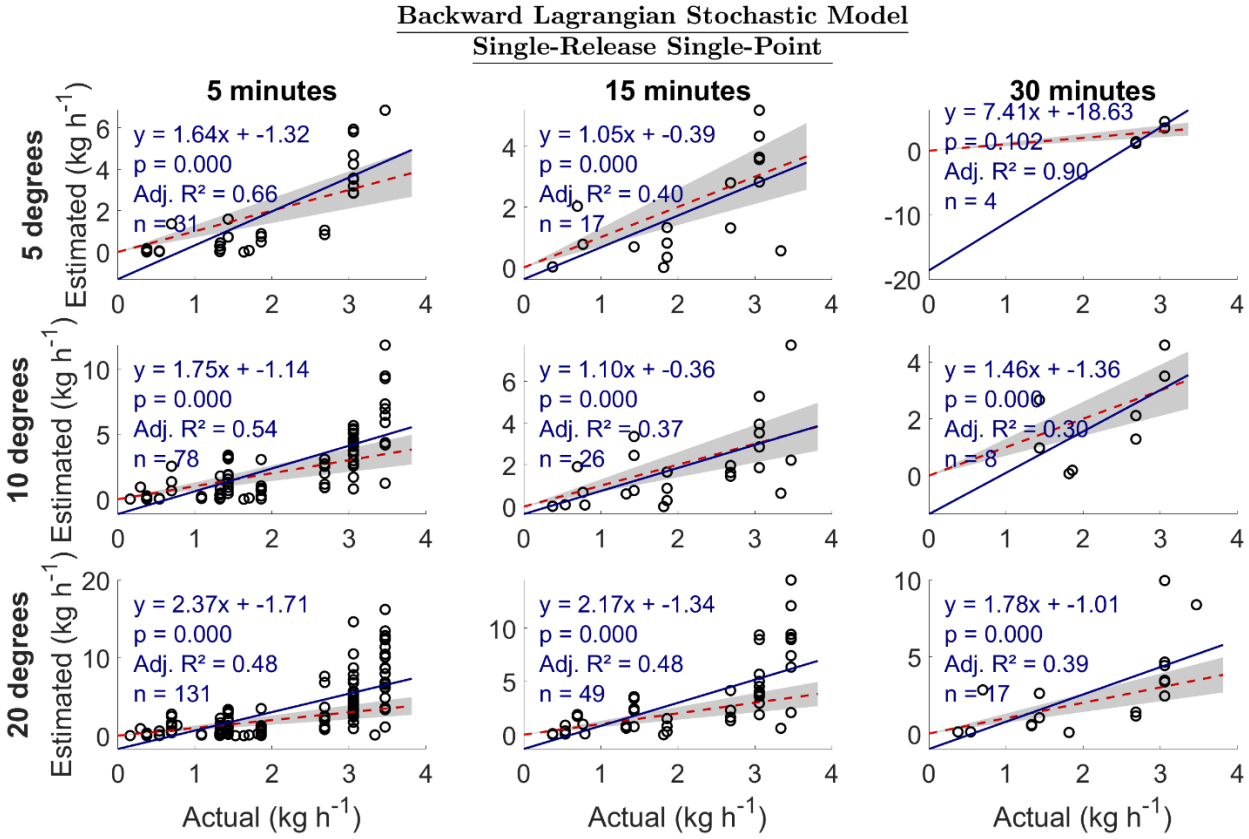


Figure 8. Estimated emission vs actual emission (kg $h^{-1}$) for single-release single-point emissions. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents ±30% of the actual emission. Adj. $R^2$ is the adjusted $R^2$. The sample size is n.

### 3.3.2 Multi-Release Single-Point

The bLs method had wide uncertainties for MSRP emissions especially in the 5 minutes, 5- and 10-degrees categories (Figure 9). In the other categories, the slopes were between -0.03 and 0.45, and $R^2$ ~0. Even though the estimated emissions spanned up to >20 kg $h^{-1}$ mostly for actual emissions of up to 6 kg $h^{-1}$, lots of points were concentrated close to 0. Compared to the GPIM MSRP results, even though both models have an $R^2$ of ~0, the GPIM had a slope closer to 1 in the 15-munites category than the bLs showing better performance. These results show that the bLs is more suitable for quantifying SRSP emissions than MSRP emissions.

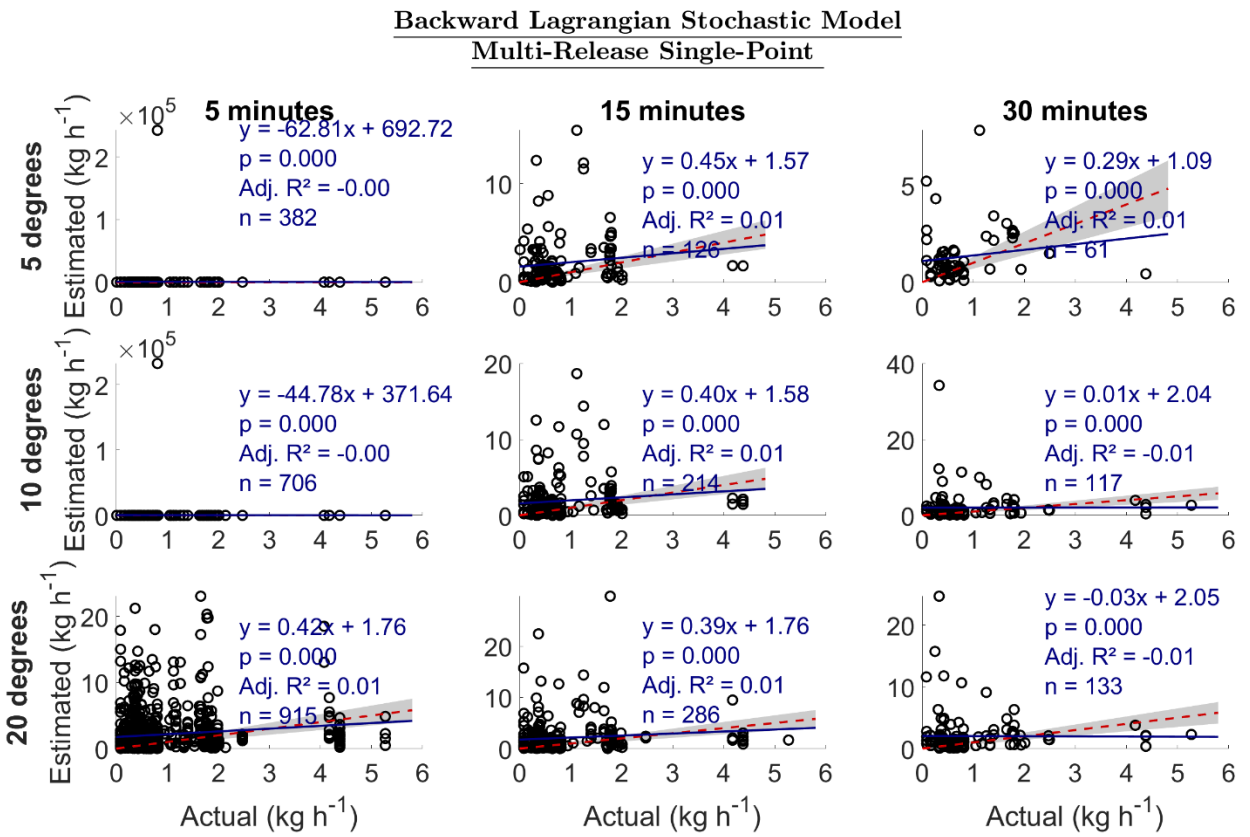


Figure 9. Estimated emission vs actual emission (kg $h^{-1}$) for multi-release single-point emissions. The red dotted line is a 1:1 line based on actual emissions i.e. points below the line are underestimated and above are overestimated emissions. The gray region represents ±30% of the actual emission. Adj. $R^2$ is the adjusted $R^2$. The sample size is n.

### 3.4 Models Comparison – Subset Data

Using a subset of the data (SRSP), filtered by 15-minute intervals within a 10-degree wind sector range where each model provided an emission estimate, the bLs model exhibited the best performance, with its linear regression closely aligned with the 1:1 line (Figure 10). The slope of the regression line for the GPIM was 1.6, indicating an overestimation, while the bLs had a slope of 0.95, suggesting high accuracy. In contrast, the EC model produced slopes of 0.08 and 0.10 when using the Kormann and Meixner (2001) and Kljun et al. (2015) footprint parameterizations, respectively, indicating significant underestimation. When emission estimates were categorized by emission point, the GPIM notably overestimated emissions at locations 4W-22 and 4W-51 (identified in Figure 3), both situated approximately 10 m from the measurement location. The EC model consistently underestimated emissions across all sites, while the bLs model provided estimates closest to the expected values. The EC model produced negative emission rates associated with negative fluxes during periods of high non-stationarity (Supplementary Information, Section 2c. iv). These deviations from stationarity reflect intermittent plume capture, where the EC system alternated between sampling emitting and non-emitting regions. Overall, these findings indicate that for source-receptor distances ranging from approximately 10 to 90 meters, the bLs model demonstrated the highest accuracy in quantifying emissions.

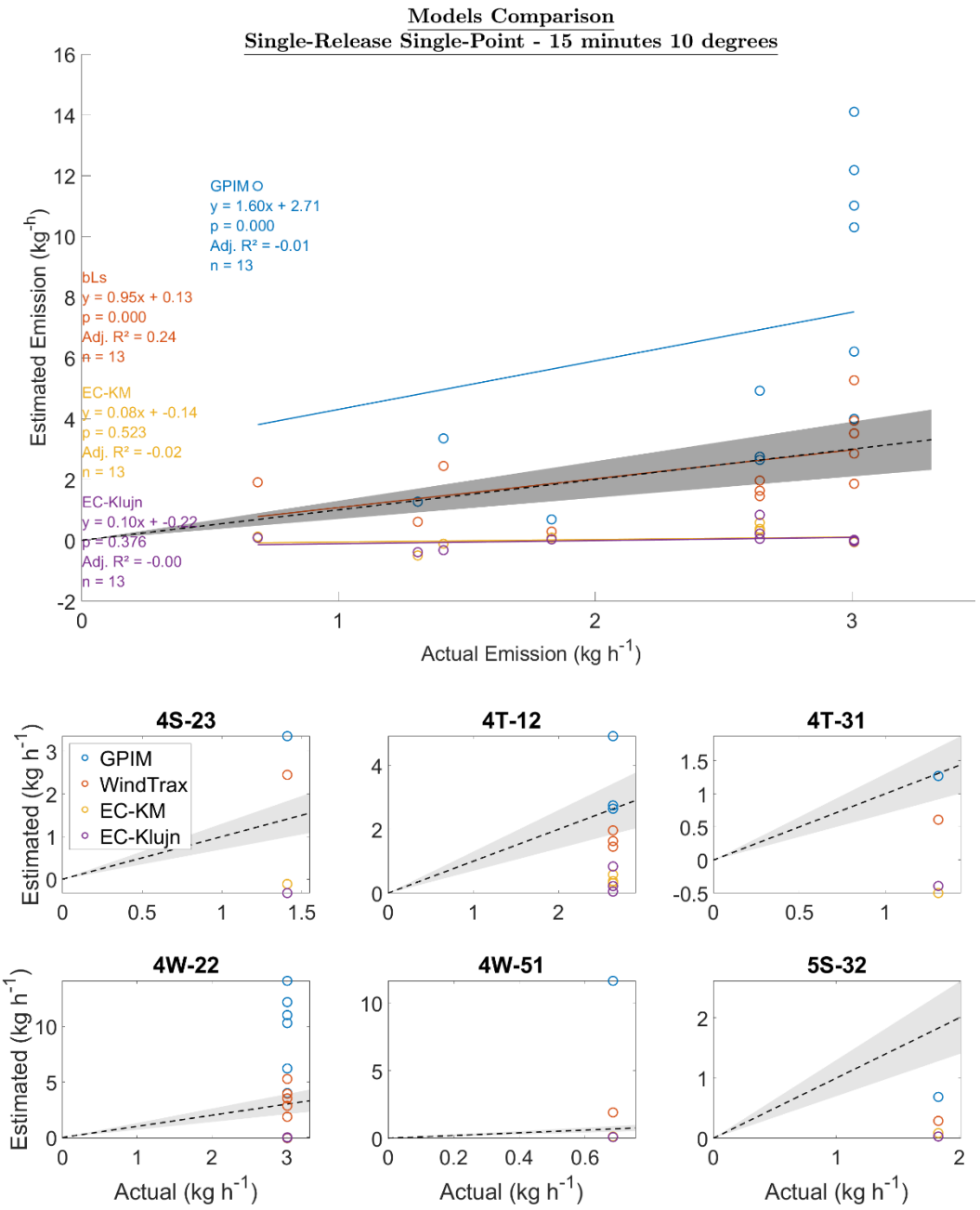


Figure 10. Top plot: Estimated emission vs actual emission (kg $h^{-1}$) for each model. GPIM is the Gaussian Plume Inverse Method, bLs is the backward Lagrangian stochastic model, EC-KM is eddy covariance with (Kormann and Meixner, 2001) footprint, and EC-Klujn is the eddy covariance estimate using the (Kljun et al., 2015) footprint. The black 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 ±30% of the actual emission. Adj. $R^2$ is the adjusted $R^2$. The sample size is n. Bottom plot: Estimated emission vs actual emission categorized by Emission Point as illustrated in Figure 3.

### 3.5 Traceability Example

To illustrate how raw data were converted into model-based emission estimates, we present one representative 15-minute interval used in Figure 10. During a controlled release at point 4W-22 (wellhead), located approximately 10.5 m from the mast, the ground-truth release rate was 3 kg $CH_4$ $h^{-1}$. Over this interval, the average $CH_4$ concentration enhancement was 8.3 ppm above the background (determined using the 5th percentile method, see Section 2.4.2.1). The cross wiwind direction was 153° (0.3 m crosswind distance), with an average wind speed of 5.8 m $s^{-1}$. The same interval was processed through the three modeling frameworks:

- The bLs model (WindTrax), using measured concentration, geometry, and meteorological data, estimated 3.5 kg $h^{-1}$.
- The GPIM model, using Equation 7 and 8 (Supplementary Information) and dispersion coefficients, estimated 10.3 kg $h^{-1}$.
- The EC method (using the (Kormann and Meixner, 2001) footprint) estimated –0.004 kg $h^{-1}$ due to a negative flux under high non-stationarity conditions.

This example illustrates how the bLs model reproduced the true emission most closely, GPIM overestimated, and EC underestimated the emission. More examples of data presented in Figure 10 are available under supplementary data "MATLAB Code & Software Configuration – Validation.xlsx".

## 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 varying emissions duration, amongst others. The ability to precisely quantify emissions using data collected by a point sensor, downwind of a source is directly influenced by plume dynamics. The $CH_4$ 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, we evaluated the ability of downwind methods—including a non-standard closed-path EC system, the GPIM, and the bLs model—to quantify emissions from single-release and multi-release point sources. While the field measurements took place under naturally varying meteorological conditions, these were not explicitly stratified or analyzed as experimental factors. Additionally, although on-site infrastructure such as storage tanks were present, their distance from the sampling instruments (~50 m) likely rendered their aerodynamic influence negligible. As such, the analysis focuses on quantification of performance under realistic but uncontrolled field scenarios, without attributing model behavior to specific atmospheric or obstacle-related conditions.

### 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 underestimated emissions with a linear regression slope for estimated emissions versus actual emissions of between -0.42 and 0.54 using the Kljun et al. (2015) and the Kormann and Meixner (2001) footprint models, and adjusted $R^2$ was ~0 (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 (~1.5 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 kg $h^{-1}$. Notably, the results for the non-standard EC system tested this

study may not be representative of EC performance in oil and gas as ogive and cospectra analysis indicated that the flux may not have been fully resolved due to non-stationarity, and instrument-related limitations.

Our results were derived from data filtered to include only periods with sampling frequencies ≥8 Hz, which significantly reduced the number of usable emission measurements. Although the instrument was configured to sample at 10 Hz, it did not consistently achieve this rate. This discrepancy may be attributed to instrument-related factors such as the 0.4-second gas flow response time, which could delay analysis of the drawn air sample in the cavity, or the use of a 3 lpm pump with 3 meters of tubing, which reduced the effective turnover rate. The dataset used for eddy covariance evaluation was predominantly flagged as low quality (flag 2) according to the (Mauder and Foken, 2004) quality control test, which classifies flux data based on steady-state conditions and the presence of well-developed turbulence (flags 0 = high, 1 = intermediate, 2 = low quality). Many of the low-quality flags were likely driven by wide deviation in w/$CH_4$ stationarity reflecting intermittent plume capture, where the EC system alternated between sampling emitting and non-emitting regions. The EC model produced negative emission rates associated with negative fluxes during periods of high non-stationarity (Supplementary Information, Section 2c. iv).

Despite high non-stationarity that resulted in low data quality issues resulting in EC inaccuracies, this study acknowledges our design limitations. Our study did not have a reliable method for aligning the asynchronous $CH_4$ and sonic anemometer data streams, which likely introduced substantial timing errors and contributed to uncertainty in the flux calculations. The intake for the closed-path system was positioned approximately 10 cm below the sonic anemometer to protect the inlet tubing from debris and precipitation by mounting it on an aluminum shield facing downward. We recognize that even this small vertical separation can introduce additional errors in flux measurements when using short towers. This design choice was a compromise to ensure instrument protection while maintaining data collection in field conditions. We acknowledge that the system used in this study was not designed or configured for standard eddy covariance analysis, and that this limitation impacts the interpretation of our results in the context of EC-based flux quantification.

In this study, continuous monitoring was conducted using a single sensor with an inlet deployed at a fence-line distance. This system requires instrumentation capable of measuring a wide concentration range, as emissions from oil and gas sites can vary between 0 and 250 ppm (Supplementary Information Section 1). While continuous monitoring systems, comprising multiple sensors, can offer enhanced spatial coverage and source localization, they also introduce higher costs. The limitations and findings reported here therefore apply specifically to this single-sensor fence-line continuous monitoring approach and may not be representative of all continuous monitoring frameworks. This study acknowledges the limitations of the eddy covariance (EC) setup used, particularly that the ABB MGGA GLA131 Series analyzer is not designed specifically for EC applications. As a result, the conclusions drawn from the EC data are constrained. The study recommends further EC testing with instruments specifically designed for EC, ideally featuring a wide measurement range (0 to ~500 ppm), faster pump speeds, shorter tubing, synchronized data logging, sampling frequencies above 10 Hz, and rugged designs suitable for field deployment. Additionally, the study recognizes that environmental factors—such as obstructions, intermittent emissions, and variable wind directions causing plume meandering—can degrade EC data quality and complicate its application in oil and gas field studies.

### 4.2 Gaussian Plume Inverse Method

The GPIM method quantified emissions within a slope of 1.65 to 3.92 and adjusted $R^2$ of between 0.4 and 0.64 with highest performance at 15-minutes 5-degrees wind sector (slope = 1.65, $R^2$ =0.4), and 5 minutes 5-degrees (slope = 2.42, $R^2$ = 0.64) for SRSP emissions (Section 3.2). For MSRP emissions, the GPIM showed wide uncertainties even though the slopes for other categories excluding 5-minutes 5-and 10-degrees, and 30-minutes 10-degrees categories, were between 0.74 and 1.60, with $R^2$ ~0 (Section 3.2). The $R^2$ close to zero showed that there was no linear relationship between the estimated and actual emissions for MSRP conditions. Overall, the GPIM performed well under 15-minutes averaging duration, and 5-degree wind-sector ranges in both SRSP and MSRP categories. The MSRP emission profiles tested in this study were complex challenging the GPIM application as the 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, even when the sensor is nominally downwind of a single source based on the average wind direction, quantification can be complicated by interference from neighboring sources. However, it is important to emphasize that such complexity is not a fundamental limitation of quantification itself, but rather a function of the experimental design and study objectives. For example, plume interference can often be minimized through strategic localization and optimization using multiple sensors—an approach that differs from the single-instrument setup used in this study. This study's design involves defining plumes based on wind sector ranges, as opposed to using multiple sensors to localize sources, and therefore does not replicate how various continuous monitoring solutions typically operate. The GPIM has previously been reported to quantify emissions within 40.7 and 60% error for a single point-source using controlled release experiments (Riddick et al., 2022b). However, GPIM correct quantification has been suggested to be better for longer distances where the plume is well mixed as seen in Figure 10. 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 was the most accurate in quantifying emissions for the SRSP release profiles but had wider uncertainties than the GPIM for MRSP scenarios (Section 3.1). For SRSP emissions, the slopes closest to 1 were during the 15-minute 5-degrees (slope = 1.05, $R^2$ = 0.4) and 10-degrees (slope = 1.10, $R^2$ =0.37). The best $R^2$ was in the 5-minute 5-degrees category (slope = 1.64, $R^2$ = 0.66). However, for MSRP emissions, the slopes were between -0.03 and 0.45 in the best categories with $R^2$ of ~0. 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 was generally more accurate than the EC and GPIM. However, for MRSP emissions, quantification accuracy was low. This discrepancy may be due to design-related challenges—specifically, interference from neighboring sources and the lack of distinct plume separation in complex flow conditions. Although the measurement point was nominally downwind of a single source, the real-world plume structure may not align with model assumptions. Additionally, the bLs implementation in WindTrax is designed for single-source scenarios and applying it in multi-source environments without adaptation can lead to inaccuracies. The GPIM and bLs methods are sensitive to background correction, which in this study was complicated by temporal overlap between release events and residual $CH_4$ accumulation, particularly

under stable atmospheric conditions. Although this is a controlled-release study, residual methane from prior emissions and the presence of multiple plumes can affect the $CH_4$ concentration during a candidate event, challenging the assumptions used to define background and isolate a single-source plume using wind-sector-based criteria. These findings highlight the importance of aligning modeling assumptions with the experimental context rather than pointing to a fundamental limitation of the method itself.

### 4.4 Implications

In recent years, there has been growing interest and need for accurate $CH_4$ 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 $CH_4$ 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 this study's design, cases involving multiple emissions with more than one release point located upwind posed challenges for the specific Gaussian and backward Lagrangian stochastic (bLs) model implementations, which were applied assuming a single active source at a time. While these models can be extended to handle multi-source scenarios, the assumptions used here limit their ability to distinguish individual contributions when plumes overlap. As a result, interference from neighboring emissions introduced ambiguity in model-observation alignment, particularly under complex wind conditions. Closed-path eddy covariance was generally unreliable in this study due to non-stationarity and limitations associated with using a non-standard EC system. In contrast, the Gaussian Plume Inverse Method (GPIM) outperformed the non-standard EC system for both single-release and multi-release single-point emissions. The backward Lagrangian stochastic (bLs) method was the most accurate for single-release single-point emissions but was less accurate than the GPIM under multi-release conditions. For both GPIM and bLs, 15-minute averaging with a narrow wind-sector (5°) yielded the best performance. While EC results in this study were limited by system constraints, future work is recommended using standard EC instruments and further optimizing GPIM and bLs models—particularly for complex multi-release scenarios—to improve accuracy and reduce uncertainties.

## Author contributions

M.M.: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Writing original draft, Review, and Editing. S.N.R.: Conceptualization, Methodology, Review, Editing, Project Administration, Supervision, and Funding Acquisition. E.K.: Investigation, Review, and Editing. K.N.S.: Review and Editing. A.H.: Review and Editing. D.Z.: 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 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

**References**

Allen, D. T.: Methane emissions from natural gas production and use: reconciling bottom-up and top-down measurements, Curr. Opin. Chem. Eng., 5, 78–83, https://doi.org/10.1016/j.coche.2014.05.004, 2014.

Bell, C., Ilonze, C., Duggan, A., and Zimmerle, D.: 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, 2023.

Brown, J. A., Harrison, M. R., Rufael, T., Roman-White, S. A., Ross, G. B., George, F. C., and Zimmerle, D.: 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, 2023.

Carbon Mapper - Science & Technology: https://carbonmapper.org/work/science, last access: 18 February 2025.

Casal, J.: Chapter 6 Atmospheric dispersion of toxic or flammable clouds, in: Industrial Safety Series, vol. 8, Elsevier, 195–248, https://doi.org/10.1016/S0921-9110(08)80008-0, 2008.

Chan, E., Vogel, F., Smyth, S., Barrigar, O., Ishizawa, M., Kim, J., Neish, M., Chan, D., and Worthy, D. E. J.: Hybrid bottom-up and top-down framework resolves discrepancies in Canada's oil and gas methane inventories, Commun. Earth Environ., 5, 566, https://doi.org/10.1038/s43247-024-01728-6, 2024.

Advancing Development of Emissions Detection (ADED): https://metec.colostate.edu/aded/.

Conrad, B. M., Tyner, D. R., and Johnson, M. R.: 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, 2023.

Crenna, B.: An introduction to WindTrax, 2006.

Day, R. E., Emerson, E., Bell, C., and Zimmerle, D.: Point Sensor Networks Struggle to Detect and Quantify Short Controlled Releases at Oil and Gas Sites, https://doi.org/10.26434/chemrxiv-2024-m0cww-v3, 20 February 2024.

Dumortier, P., Aubinet, M., Lebeau, F., Naiken, A., and Heinesch, B.: 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, 2019.

Foken, Th. and Wichura, B.: 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, 1996.

Foster-Wittig, T. A., Thoma, E. D., and Albertson, J. D.: 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, 2015.

Horst, T. W. and Lenschow, D. H.: Attenuation of Scalar Fluxes Measured with Spatially-displaced Sensors, Bound.-Layer Meteorol., 130, 275–300, https://doi.org/10.1007/s10546-008-9348-0, 2009.

Hsieh, C.-I., Katul, G., and Chi, T.: 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, 2000.

Hutchinson, M., Oh, H., and Chen, W.-H.: 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, 2017.

Ilonze, C., Emerson, E., Duggan, A., and Zimmerle, D.: Assessing the progress of the performance of continuous monitoring solutions under single-blind controlled testing protocol, https://doi.org/10.26434/chemrxiv-2023-8bfgm-v2, 8 February 2024.

Johnson, M. R., Tyner, D. R., and Szekeres, A. J.: 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, 2021.

Kljun, N., Calanca, P., Rotach, M. W., and Schmid, H. P.: 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, 2015.

Kormann, R. and Meixner, F. X.: An Analytical Footprint Model For Non-Neutral Stratification, Bound.-Layer Meteorol., 99, 207–224, https://doi.org/10.1023/A:1018991015119, 2001.

EddyPro 7 | Software Downloads: https://www.licor.com/support/EddyPro/software.html, last access: 30 January 2025.

Mauder, M. and Foken, T.: Documentation and Instruction Manual of the Eddy Covariance Software Package TK2, 2004.

METEC | Colorado State University: https://metec.colostate.edu/, last access: 18 February 2025.

Mollel, W., Zimmerle, D., Santos, A., and Hodshire, A.: Using Prototypical Oil and Gas Sites to Model Methane Emissions in Colorado's Denver-Julesburg Basin Using a Mechanistic Emission Estimation Tool, ACS EST Air, 2, 723–735, https://doi.org/10.1021/acsestair.4c00168, 2025.

Moncrieff, J., Clement, R., Finnigan, J., and Meyers, T.: Averaging, Detrending, and Filtering of Eddy Covariance Time Series, in: Handbook of Micrometeorology, vol. 29, edited by: Lee, X., Massman, W., and Law, B., Kluwer Academic Publishers, Dordrecht, 7–31, https://doi.org/10.1007/1-4020-2265-4_2, 2005.

Morin, T. H.: Advances in the Eddy Covariance Approach to $CH_4$ Monitoring Over Two and a Half Decades, J. Geophys. Res. Biogeosciences, 124, 453–460, https://doi.org/10.1029/2018JG004796, 2019.

Polonik, P., Chan, W. S., Billesbach, D. P., Burba, G., Li, J., Nottrott, A., Bogoev, I., Conrad, B., and Biraud, S. C.: 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, 2019.

Rey-Sanchez, C., Arias-Ortiz, A., Kasak, K., Chu, H., Szutu, D., Verfaillie, J., and Baldocchi, D.: Detecting Hot Spots of Methane Flux Using Footprint-Weighted Flux Maps, J. Geophys. Res. Biogeosciences, 127, e2022JG006977, https://doi.org/10.1029/2022JG006977, 2022.

Riddick, S. N. and Mauzerall, D. L.: 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, 2023.

Riddick, S. N., Ancona, R., Cheptonui, F., Bell, C. S., Duggan, A., Bennett, K. E., and Zimmerle, D. J.: 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, 2022a.

Riddick, S. N., Cheptonui, F., Yuan, K., Mbua, M., Day, R., Vaughn, T. L., Duggan, A., Bennett, K. E., and Zimmerle, D. J.: 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, 2022b.

Riddick, S. N., Mbua, M., Anand, A., Kiplimo, E., Santos, A., Upreti, A., and Zimmerle, D. J.: Estimating Total Methane Emissions from the Denver-Julesburg Basin Using Bottom-Up Approaches, Gases, 4, 236–252, https://doi.org/10.3390/gases4030014, 2024a.

Riddick, S. N., Mbua, M., Santos, A., Hartzell, W., and Zimmerle, D. J.: 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, 2024b.

Tagliaferri, F., Invernizzi, M., Capra, F., and Sironi, S.: 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, 2023.

UNEP: Homepage | The Oil & Gas Methane Partnership 2.0, 2024.

US EPA: 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, 2013.

US EPA: Natural Gas and Petroleum Systems in the GHG Inventory: Additional Information on the 1990-2021 GHG Inventory (published April 2023), 2023.

Vickers, D. and Mahrt, L.: 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, 1997.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Q. J. R. Meteorol. Soc., 106, 85–100, https://doi.org/10.1002/qj.49710644707, 1980.

WindTrax 2.0: thunder beach scientific, n.d.

Zimmerle, D., Dileep, S., and Quinn, C.: 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, 2024.