*atmosphere*

MDPI

*Article*

# Comparative Performance of Gaussian Plume and Backward Lagrangian Stochastic Models for Near-Field Methane Emission Estimation Using a Single Controlled Release Experiment

**Aashish Upreti [1,2,*], Kira B. Shonkwiler [1], Stuart N. Riddick [1,3,*] and Daniel J. Zimmerle [1]**

[1] Methane Emission Technology Evaluation Center (METEC), Energy Institute, Colorado State University, Fort Collins, CO 80524, USA; kira.shonkwiler@colostate.edu (K.B.S.); dan.zimmerle@colostate.edu (D.J.Z.)
[2] Department of Mechanical Engineering, Colorado State University, Fort Collins, CO 80521, USA
[3] Department of Science, Engineering and Aviation, University of the Highlands and Islands Perth, Crieff Road, Perth PH1 2NX, UK
* Correspondence: aashish.upreti@colostate.edu (A.U.); stuart.riddick.perth@uhi.ac.uk (S.N.R.)

## Abstract

Methane ($CH_4$) is a major component of natural gas and a potent greenhouse gas. Increasing atmospheric methane concentrations are attributed to emissive anthropogenic activities by an average of 13 ppb per yr since 2020 and are linked to a changing global climate. Mitigating $CH_4$ emissions from oil and gas production sites has recently become a target to reduce overall greenhouse gas emissions; however, monitoring the efficacy of mitigation strategies depends on accurate quantification of $CH_4$ emissions at the facility-level. Near-field quantification of methane ($CH_4$) emissions from oil and gas (O&G) facilities remains challenging due to the effects of atmospheric variability and sensor configuration on atmospheric dispersion models. This study evaluates the performance of two atmospheric dispersion models, the Gaussian plume (GP) and backward Lagrangian stochastic (bLS), by comparing calculated $CH_4$ emissions to controlled single-point emissions between 0.4 and 5.2 kg $CH_4$ $h^{-1}$. Emissions were calculated by both models using 121 individual sets of measurements comprising five-minute averaged downwind methane mixing ratios and matching meteorological data. The comparison shows that the bLS approach achieved a higher proportion of emission estimates within a factor of two (FAC2) of the known emission rates compared to the GP approach. The emissions calculated by the bLS model also had a lower multiplicative error and reduced bias relative to GP. Other error-based metrics further confirmed the bLS model performed better, as it yielded lower RMSE and MAE than GP. Statistical analysis of the emission data shows that the lateral and vertical alignment of the source and the sensor plays a critical role in emission estimations, as measurements made closer to the plume centerline and at a distance between 40 and 80 m downwind yielded the best FAC2 agreement. High wind meander degraded the ability of both approaches to generate representative emissions, particularly with the GP approach, as it violates the modeling approach's assumption of steady-state emissions. Data suggest emissions calculated by the bLS model are comprehensively in better agreement, but the computational demands of the modeling approach and integration into fenceline systems limit real-time applicability. While these results provide insight into model performance under controlled near-field conditions, their applicability to more complex or heterogeneous oil and gas production environments (e.g., the regions Marcellus or Unita Basins) remains limited and uncertain.

Check for updates

Academic Editor: James Lee

Received: 26 February 2026
Revised: 12 April 2026
Accepted: 14 April 2026
Published: 20 April 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

**Keywords:** methane emission quantification; oil and gas facilities; gaussian plume; near-field dispersion modeling; sensor placement

## 1. Introduction

Methane ($CH_4$) is a major component of natural gas (NG) and a potent greenhouse gas (GHG), with a global warming potential ($GWP_{100}$) 27.9 times larger than $CO_2$ over a 100-year time period [1]. Anthropogenic activities like oil and gas (O&G) extraction, agriculture, biomass burning, and waste disposal account for between 60 and 65% of the total $CH_4$ emission globally, while natural emitters like termites, freshwater ecosystems, and wetlands contribute to the remaining fraction [2]. Atmospheric $CH_4$ concentration has increased by an average of 13 ppb $yr^{-1}$ since 2020 [3,4], contributing to changes in the global climate over the past decade. As a controllable emission source, mitigating $CH_4$ emissions from O&G production sites has been targeted for reducing overall greenhouse gas emissions [5,6].

Monitoring the efficacy of mitigation efforts depends on the accurate quantification of facility-level emissions. Yet quantifying facility-level emission estimate remains highly uncertain due to the effects of atmospheric conditions and aerodynamically complex equipment present at O&G extracting sites [7–12]. O&G production facilities comprise thousands of components (connectors, valves, and flanges) on multiple pieces of equipment wellheads, separators, and water/condensate storage tanks, often with several to dozens of potential leak points per facility [13,14]. The potential for emissions from different heights and position on multiple components add additional complexity to quantification, therefore further challenge quantification using measurements [15,16].

Emissions on O&G production facilities typically can be emitted from vented, fugitive and combustion-related sources [14]. Gas is typically vented as the result of a process action (e.g., from pneumatic controllers on the wellhead or tank flashing) or from maintenance activities (such as the blowdown of pressurized equipment or liquid uploading). Fugitive emissions are unintended leaks and are typically emitted from flanges, valves or seals as equipment on the facility ages or becomes damaged [14,17]. Combustion-related emissions arise from burning of natural gas in turbines, engines, heaters, and flares [14,18]. Emissions from some vented and most fugitive sources (including vents, pneumatic controllers, valves, flanges, and tank hatches) are typically treated as point sources, whereas other vented and combustion emissions may behave as spatially distributed (or effective area sources) depending on the measurement scale [19–21]. Point sources are usually easier to detect as they have higher methane concentrations at the emission point but are not necessarily larger than area sources as they emit over a smaller area.

Methane emission quantification using large-scale, remote sensing approaches, like aerial surveys, satellite observations, and tall tower inversion methods are suitable for constraining basin-wide $CH_4$ fluxes but their applicability in making facility-scale emission estimations are limited by both spatial resolution and temporal coverage [22,23]. These methods typically rely on spatial integration of methane enhancements over footprints spanning meters to several kilometers, which prohibits the detection and attribution at the equipment or facility level [22,23]. Temporally, their measurements are either very long (tall-tower multiple-day averages; [24]) or very short (satellite and aerial surveys with <1 min observation), and infrequent sampling limits sensitivity to intermittent or short-term emission events [25].

To address this, near-field techniques that use near-field measurement and dispersion models are used to generate facility-level emission estimates. However, the accuracy of

the emissions calculated using these dispersion approaches is limited due to meteorological variability and modeling assumptions, which are rarely valid at complex O&G sites [10,26,27].

Among commonly used dispersion modeling methods applied to make $CH_4$ emission estimates across a range of source types are the Gaussian plume (GP) and backward Lagrangian stochastic (bLS) approaches [27–30]. The GP method models the spread of the pollutant plume from a point source following a Gaussian profile under steady-state conditions with constant meteorological parameters [31–33]. Under its standard formulation, the GP approach models atmospheric transport as the dispersion and advection of a continuous point source by the mean wind field, with turbulent dispersion in the vertical and crosswind directions, based on the analytical advection–diffusion solution [34,35]. This approach's main advantages are mathematical simplicity and computational efficiency. In the O&G sector, the GP approach has been widely applied to quantify the $CH_4$ emissions from individual pieces of production equipment to facility-scale using mobile and fixed survey and continuous monitoring strategies [29,32,35]. The GP-based approach has been applied to both onshore and offshore production facilities, where near-field conditions and steady-state assumptions are presumed to be reasonable [26,29,31,32]. Typical estimates quantified using GP methods in O&G applications range from small (<1 kg $CH_4$ $h^{-1}$) equipment-level sources to larger facility-level emissions (order of 10 kg $CH_4$ $h^{-1}$) [29,31–33]. The uncertainty associated with emissions calculated using the GP approach is mostly driven by the steady state assumptions made and simplified source representations, particularly when there are intermittent emissions, multiple sources and complex aerodynamic conditions caused by structures at the facility [26,31,33]. The insensitivity of the model to changes in wind speed and direction adds additional uncertainty to the GP based estimates [11,32]. Evaluations of the GP method in both field studies and controlled experiments indicate an uncertainty of the order ±50% to ±100%, with larger errors observed within highly variable meteorological conditions or when plume intermittency/geometry is not adequately represented [29,31–33].

The bLS approach estimates emissions by simulating turbulent air particles backward in time from receptor to potential source using a stochastic representation of atmospheric turbulence [27,36]. Observed concentration enhancements are used to calculate surface emission fluxes or emissions by quantifying the fraction of particle trajectories that originate from emitting surface, without assuming Gaussian concentration profile [36,37]. This method assumes homogenous surface conditions and constant turbulence over the averaging period [27,36]. The accuracy of emission estimates improves with larger numbers of modeled particle trajectories and better resolved wind statistics [9,27,36,37]. Evaluation of $CH_4$ quantification in the O&G sector indicates that bLS provides robust estimates when turbulence is well-characterized [29,38]. The bLS approach can better represent aerodynamic conditions between the source and receptor and atmospheric stability than the GP approach. However, these advantages are accompanied by a higher computational requirement and greater sensitivity to turbulence characterization [13,32]. Studies have shown that emission estimates calculated using a bLS model typically have uncertainties between ±30 and ±50%, with uncertainties rising toward a factor of two or greater when turbulence is poorly resolved [28,29,38].

Although bLS and GP approaches are widely used for $CH_4$ emission quantification at O&G facilities, most prior evaluations of dispersion model performance relied on aggregated facility-level emissions, multi-source configurations and extended temporal averaging [11,29,39]. These studies are valuable for estimating facility-level but have limited success in isolating the effect of short-time-scale plume intermittency, atmospheric variability, and source-receptor geometry on the performance of the model

approach in near-field conditions. Consequently, the relative characteristics and performance of bLS and GP models to characterize single-emission point sources, representative of individual equipment-level emissions, are currently not well-constrained for complex O&G environment.

This study aims to directly address this knowledge gap using data from controlled, single-point source $CH_4$ release experiments to evaluate the GP and bLS approaches' performance under near-field O&G conditions. Using a constant, controlled emission source, we aim to systematically identify the impact of short time-scale meteorological variability and plume-geometry on the accuracy of calculated emissions. The experimental framework enables a direct comparison of bLS and GP performance under identical measurement and atmospheric conditions. In contrast to prior studies that relied on long-term averaging and source aggregation to evaluate dispersion model performance [29,33,36,38], this study isolates key uncertainty drivers, model-specific sensitivities, and failure modes in the GP and bLS approaches. As a result, the work may provide insight into the conditions under which GP and bLS yield reliable emission estimates and clarifies their relevance for $CH_4$ quantification in near-field O&G settings.

## 2. Materials and Methods

### *2.1. Site Description and Experiment Design*

All experiments reported in this study were performed at the Methane Emission Technology Evaluation Center (METEC) test facility, operated by Colorado State University's Energy Institute in Fort Collins, Colorado, USA. At the time when the measurements for this study were made, METEC was an experimental site that simulated NG emissions from five O&G production pads using controlled release emission points at likely/known fugitive and vented sites on tanks, separators, and well-heads (Figure 1). The NG used for the controlled releases was stored in compressed gas cylinders, and the emission rates controlled using precision orifices, pressure regulators, and solenoid valves [29].

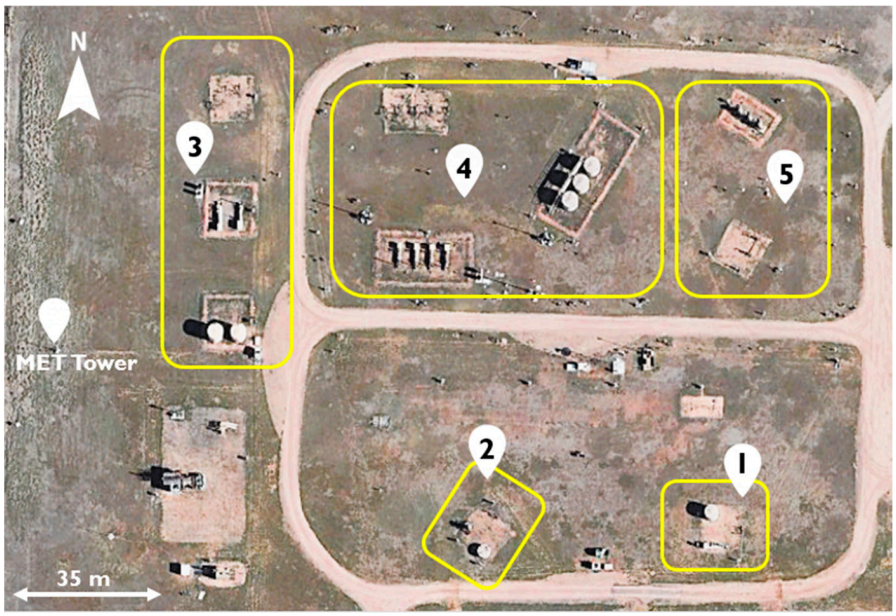



**Figure 1.** Aerial view of METEC's experimental layout in Ft. Collins, Colorado. Each labeled "pad" consists of oil and gas equipment fitted with controlled release points at numerous places in locations where real-world leaks are known to occur. In this study, experiments were conducted on all pads. Equipment types for releases in these experiments were wellheads (on pad 4), separators (on pads 1–5), and tanks (on pad 4) at varying rates and release heights/locations. METEC has expanded its capabilities in the time since experiments were performed and since this photo was taken.

As part of the larger ADED (Advanced Development of Emissions Detection) field campaign conducted at METEC [40], an experimental schedule was generated by METEC staff to vary the location, duration and emission rate of the NG releases across the METEC site between the 19 March and 20 September. The location and duration were shared with us before each measurement. For each measurement, a portable methane analyzer was positioned downwind of the known release point and left in place for at least 5 min. The location of each measurement point was determined using a real-time kinematic (RTK) GPS unit (SparkFun RTK Facet L-Band, Boulder, CO, USA). At any given time during each measurement period, only a single controlled emission point was active, ensuring that all observations correspond to a single-point source. For each release, the first measurement was conducted downwind to the source at a varying distance selected to ensure plume interception under prevailing wind direction. The measurement position was selected dynamically to ensure plume interception and representative mixing ratio measurement. After each 5 min measurement, the methane analyzer was moved approximately 10 m downwind from the emission point; this was repeated between four and seven times, typically extending measurements towards the facility fenceline depending on wind direction and site constraints.

### *2.2. Meteorological Measurement*

Meteorological data were measured at 1 Hz using the R.M. Young 81000 3D sonic anemometer (R.M. Young, Traverse City, MI, USA) permanently mounted at METEC, (Figure S1), a LI-200R-SMV (LI-COR Biosciences, Lincoln, NE, USA), and a Met One 597 (Met One Instruments, Grant Pass, OR, USA). The sonic anemometer was positioned 6.7 m above ground with a maximum horizontal distance between the sonic anemometer and a controlled release of 147 m. The Met One 597 measured at 6.2 m above the ground and was located at the sonic anemometer tower. The placement of Met One (6.2 m) and sonic anemometer (6.7 m) on a single meteorological tower at METEC reflect its fixed installation for site-wide observations. The LI-200R-SMV was located 270 m from nearest source.

The sonic anemometer provided wind speed (accuracy of $\pm 1\%$) and wind direction data (accuracy of $\pm 2°$). The Met One 597 provided ambient air temperature (accuracy of $\pm 2$ °C), RH (accuracy of $\pm 2\%$), and barometric pressure (P; hPA). Solar radiation was measured using the LI-200R-SMV model pyranometer located 250 m away from METEC over similar surface characteristics (i.e., roughness length, displacement height, topography). The solar radiation delta-T (SRDT) method was used for the PGSC (Pasquill–Gifford Stability Classes) classification (Tables S1 and S2) [41,42]. PGSC define atmospheric stability categories derived from solar radiation and wind speed and are used to parameterize dispersion coefficients.

### *2.3. Methane Mixing Ratio Measurement*

Downwind methane mixing ratio data were measured using an ABB Micro-portable Greenhouse Gas Analyzer (MGGA; ABB, Zurich, Switzerland) [43]. The analyzer was connected to an inlet 1 m above the ground using 3.175 mm ID tubing (Figure S2). A sensor inlet height of 1 m above the ground is consistent with the U.S. Environmental Protection Agency's Other Test Method (OTM)-33A near-field surface sampling guidelines [44]. The sample cell flow rate of the MGGA was $2.6 \pm 0.2$ min$^{-1}$, measurement accuracy of the order of 1% of the reading, and a resolution of 0.9 ppb (at 1 s) and 0.3 ppb at 3 s [43].

Before each downwind measurement, the background methane mixing ratio was measured for 15 min and averaged. Due to in-built sampling rates, the MGGA recorded the mixing ratio data at two sampling rates: 5 Hz (4.967 Hz to 6.597 Hz) for experiment 2, and 1 Hz (1.001 Hz to 1.004 Hz) for all others.

### *2.4. Controlled Release System*

The controlled release experiment was performed between 19 March and 20 September, 2024. The gas composition of the controlled release NG was determined using a gas chromatograph (7890 GC, Agilent Technologies, Santa Clara, CA, USA) with nominally 85 mol% $CH_4$. The NG release rate ranged between 0.53 and 6.03 kg $h^{-1}$ (0.45 to 5.13 kg $CH_4$ $h^{-1}$) which was comparable to moderate to large fugitive leaks in upstream O&G operations [13,44]. Field studies at NG production and transmission facilities have shown that the most equipment-level methane leaks are generally under 1 kg $CH_4$ $h^{-1}$ [14], while component-level leaks in the storage and transmission sector typically range between 0.05 and 5 kg $CH_4$ $h^{-1}$ [45]. The selected release rates of NG fall in the range that represent emissions comparable to moderate fugitive leak events in upstream O&G operations. The controlled releases during our experiment were from separators, tanks, and wellheads at heights between 0.4 and 4.7 m above ground.

After each experiment, the methane mixing ratio and the meteorological and controlled-release data were collated, time-synched and merged into a single data file. The number of measurements in each experiment is presented in Table S3. The 5 min averages of each parameter were then determined for each measurement. Even though previous studies used >10 min averaging periods in the past [29,36,38], a 5 min averaging period was chosen due to practical constraints associated with conducting single controlled-release experiments during the ongoing ADED campaign. In addition, the frequent shift in wind direction at the METEC site limits the steady plume condition where longer averaging periods would have required multiple relocations of the gas analyzer, leading to additional variability in source-sensor geometry. Estimates derived from short averaging durations can increase uncertainty and bias, as longer averaging periods better satisfy steady-state assumptions and reduce atmospheric and instrumental noise. However, under controlled release conditions with relatively high frequency measurements and stable meteorology, a 5 min averaging period can reasonably approximate steady behavior of plume. At the METEC site, high frequency $CH_4$ measurements, under steady controlled emissions and moderate wind conditions allow turbulent fluctuation to converge toward a representative mean concentration. Nevertheless, the observed wind direction variability in the dataset indicates that steady-state conditions were not always satisfied; thus, the 5 min averaging period should be interpreted as a practical and experimental specific choice rather than universally sufficient representation of plume behavior. Therefore, although the 5-min averaging interval provides suitable inputs for near-field quantification in METEC, the results should be interpreted within the scope of short duration measurements, with caution applied when generalizing to longer averaging period.

### *2.5. Modeling*

#### 2.5.1. Gaussian Plume (GP) Approach

The GP approach is used to model the dispersion of pollutants from a continuously emitting point source. The model assumes three-dimensional steady-state dispersion in the downwind direction following a normal (Gaussian) distribution in the vertical ($z$, m) and crosswind/lateral ($y$, m) directions from the centerline of the plume ($x$, m) [31,32,34,35]. The horizontal ($\sigma_y$, m) and vertical dispersion coefficients ($\sigma_z$, m) were obtained from distance downwind and PGSC (Tables S2-i and S2-ii).

The emission rate ($Q$, g $s^{-1}$) from a point source of height ($H$, m) can be calculated using the mixing ratio ($C$, μ g $m^{-3}$) at any given point downwind in a 3-dimensional space ($x$, $y$, $z$) and wind speed ($U$, m $s^{-1}$) (Equation (1)). In this study, we used the physical height as the effective source height, as the plume rise effects were negligible, and the plume

rise for controlled low-height $CH_4$ release. Equation (1) was applied in its inverse form, substituting the mixing ratio and solving algebraically for unknown emission rate $Q$.

$$Q = C\ (x,\ y,\ z)\ 2\pi U\sigma_y\sigma_z\ e^{\frac{y^2}{2\sigma_y^2}}\left[\left(e^{-\frac{(z-H)^2}{2\sigma_z^2}}\right)+\left(e^{-\frac{(z+H)^2}{2\sigma_z^2}}\right)\right]^{-1} \quad (1)$$

The Gaussian plume approach makes several assumptions about the behavior of the emission source and plume, and these include: 1. the source emits gas at constant rate; 2. the sensor is typically positioned sufficiently downwind (often >100 m) to better satisfy steady-state assumption; 3. the terrain between the source and the sensor is homogeneous; 4. meteorological conditions are horizontally uniform and steady over the averaging period; 5. dispersion in crosswind and vertical directions follows a Gaussian distribution; and 6. effects of complex surface roughness and flow recirculation are negligible [46]. One major shortcoming of this approach is that emissions can always be calculated even if any of the assumptions are violated [26,46]. Although some underlying assumptions of the Gaussian plume (GP) approach (e.g., steady-state conditions) may be violated, the model is still applied in near-field conditions and is not a strict requirement for application. Therefore, the use of GP in this work reflects its application under assumption-challenging conditions (e.g., measurement at <100 m) rather than ideal deployment scenarios.

The model is mathematically simple and efficient in terms of computation and implementation. However, it is not clear how useful this approach is for calculating emissions in the real-world, as the model is based on a steady-state assumption of emission with constant meteorological conditions over homogenous terrain [47]. Therefore, it is unlikely that the GP approach can be used for conditions at oil and gas sites which include flow obstructions and are often located in heterogenous surroundings [38,46]. Accordingly, results from GP approach in this study has to be interpreted in the context of near-field, application specific context, rather than as an indicator of intrinsic model performance under fully compliant assumption.

#### 2.5.2. Backward Lagrangian Stochastic (bLS) Approach

The backward Lagrangian stochastic (bLS) approach tracks the trajectories of released particles in the atmosphere [27,36–38] inversely from the detector to the emission source using stochastic processes that simulate the turbulent nature of the dispersion [27]. The bLS model used in this study was WindTrax (version 2.0.9.7) developed by Thunder Beach Scientific (Halifax, NS, Canada) [37]. Data input to the model comprised 5 min averages of wind speed, wind direction, air temperature, barometric pressure, PGSC, $CH_4$ mixing ratios, background mixing ratios, roughness length, meteorological measurement location/heights, source location/height, and measurement location/height [Table S4].

The initial number of particles used in the inversion for each 5 min measurement was 500,000. In the advent of a quantification error, the number of particles increased in steps of 1 million from 1 million to 20 million until the error was resolved. Despite increasing particle count, 18 samples did not resolve the errors using 20 million particles. The average particle count used for the simulations was 6.4 million.

The two methods differ substantially in terms of computational requirements. The GP method uses a direct analytical formulation, providing estimates with negligible computational time, whereas the bLS approach requires stochastic particle simulations using WindTrax, where computational time increases with the increase in particle count. For instance, the simulation requiring particles above 14 million required approximately 24–36 h to complete. Also, the WindTrax is freely available, but limitation is associated with computational time rather than software accessibility.

### *2.6. Performance Analysis*

The calculated emissions were compared against the known controlled-release values to evaluate models' performance in varying meteorological conditions and plume geometries. Metrics used to assess model performance for this study were: 1. calculated emissions within a factor of two of the known emission rate (FAC2) [Equation (S1)]; 2. the geometric mean (GM) bias calculated by using geometric mean of the ratio of estimated to known emission [Equation (S2)]; 3. the median factor of error (MFoE) calculated by taking exponential of median absolute logarithmic error between estimated and known emission [Equation (S3)] [48,49]; 4. mean absolute error (MAE) calculated as the mean of absolute difference between estimated and known emission rate [Equation (S4)]; 5. the root mean square error (RMSE) calculated as a square root of mean of the squared differences between estimated and known emission [Equation (S5)]; 6. mean absolute bias (MB) calculated as the mean of the signed difference between estimated and known emission [Equation (S6)]; 7. fractional bias (FB) calculated as symmetrically normalized mean difference between estimated and known emissions [Equation (S7)]; and 8. normalized mean square error (NMSE calculated as normalized mean squared difference between estimated and known emissions [Equation (S8)]).

The FAC2 measures the proportion of emissions calculated by the model that are within a factor of two of the known emission rates, i.e., +100% and −50%. This was used because factor of two criterion is widely adopted in model evaluation in atmospheric dispersion as a practical benchmark for acceptable agreement subject to measurement uncertainty and meteorological variability [48,49]. The GM identifies over/underestimation of model prediction, returning a value of one for an unbiased model, a value less than one for a model that underestimates the emission, and values greater than one for a model that overestimates emissions. The MFoE calculates a value representing the multiplicative spread of error (always $\geq$1). MFoE values close to 1 indicate a model that calculates emissions with a small deviation from the known value. These three indicators, when taken together, are used to understand the veracity of the model's calculated emissions in varying environmental conditions.

The MAE measures the average absolute magnitude of error between the estimated and true emission rates, regardless of direction. The RMSE predicts the square root of the average of the squared errors, which is highly sensitive to outliers. Unlike MAE and RMSE, MB shows the bias direction of the error. The positive error in MB means that the model is overestimating, and the negative means that it is underestimating. These three indicators, when taken together, are used to understand magnitude, variability and directional bias of the model estimation errors. In addition, FB provides symmetrically normalized measure of systematic bias between estimated and known emissions which identifies the consistency of over or underprediction. The NMSE quantifies overall deviation between estimated and known emissions, combining both bias components and variance into a normalized metric.

### *2.7. Statistical Analysis*

Numerical performance metrics alone are insufficient to explain the variation in the model's predictive capacity and emphasize physical interpretation [48,49]. To better understand the model's performance, statistical analysis was conducted to evaluate the effect of other factors like release rates, signal strength, wind meander, and plume geometry. In each case, the effect of each factor was evaluated using the FAC2 metric to identify optimal conditions for each environmental or geometric factor.

Signal strength: Signal strength is defined as the 5 min average mixing ratio reported by the methane analyzer. The definition of signal strength used in this study is the absolute measured concentration rather than enhancement. Signal strength is therefore used as an

indicator of plume presence and measurement quality. Signal strength bins were grouped into four evenly populated quartiles: ≤2.3 ppm, 2.3–2.9 ppm, 2.9–4.0 ppm, and >4.0 ppm.

Release rates: Release rate bins were created to determine if a model performs consistently across varying source strengths. Release rate bins (kg $h^{-1}$) were defined as: low for emission rates less than 1.8 kg $h^{-1}$; medium between 1.8 and 3.4 kg $h^{-1}$; and high between 3.2 and 6.0 kg $h^{-1}$.

Atmospheric stability: Atmospheric stability was categorized using PGSC and derived from 5 min averages of solar radiation and wind speed (Tables S1 and S2). All measurements at METEC were unstable daylight conditions, specifically classes A (extremely unstable), B (moderately unstable), and C (slightly unstable).

Wind meander: Wind meander captures the directional variability of the wind direction field [18,30]. In this study, we define wind meander as the standard deviation of wind direction during the 5 min averaging interval and indicates whether dispersion is governed by ambient airflow or turbulence (mechanical or convective). Strong variation in wind direction can disrupt the steady state assumption of GP, thereby reducing the reliability of model estimates [36,49]. Wind meander data categorized calculated emissions into seven 5° bins (±0–5°, ±5–10°, ±10–15°, ±15–20°, ±20–25°, ±25–30° and >±30°.

Plume centeredness: The plume centeredness score is the ratio of crosswind distance (y) to the horizontal dispersion coefficient ($\sigma_y$), a measure of lateral plume spread [32]. This score indicates the relative distance of a measurement point to the centerline of the downwind plume. Larger values indicate the measurement position was near the plume edge, with smaller values corresponding to the plume axis [50]. The centeredness score was categorized into four bins: 0–0.5; 0.5–1; 1–2; and >2.

Plume height: The height match score is determined by using the ratio of the absolute difference between the source and sensor height to the plume's vertical dispersion coefficient, $\sigma_z$, and used to indicate the extent of vertical overlap between the sensor and plume heights. In this analysis, four height match score bins (0–0.5, 0.5–1, and 1–2) are used to indicate the relative vertical distance between the analyzer inlet and the plume centerline and are expressed as multiples of plume height standard deviation ($\sigma_z$).

Plume centerline distance: The plume centerline distance represents the downwind distance between the source and the analyzer inlet measured along the plume's trajectory. Dispersion and dilution increase with downwind distance. At smaller downwind distances the plume is likely to be more concentrated, while farther downwind, mixing is enhanced and signal strength is smaller [34,50]. To evaluate this, the plume centerline distances were binned into three categories: 0–40, 40–80, and >80 m.

## 3. Results

### *3.1. Measurement Data*

#### 3.1.1. Methane Mixing Ratio

Between March and September 2024, 121 5 min methane mixing ratio measurements were made downwind of single-point source releases. These data comprise 5 min average $CH_4$ mixing ratios measured at 1 m above ground level. The observed average $CH_4$ mixing ratios ranged from a minimum of 1.8 ppm to a maximum of 21.1 ppm. The median of mixing ratio across the measurement set was 2.8 ppm, and the interquartile range spanning 2.3 to 4 ppm.

#### 3.1.2. Meteorology

The 5 min wind speeds during the measurements ranged between 0.9 and 7.9 m $s^{-1}$, with a median of 3.1 m $s^{-1}$ and an average of 3.3 ± 1.4 m $s^{-1}$. The data was filtered for wind speeds <1 m $s^{-1}$, as low speeds do not provide adequate transport conditions

for a well-defined plume and are not typically included in downwind dispersion analysis [36,38,51]. The 5 min averaged wind direction data included sampling winds from nearly all directions (between 18 and 342°). The measurements were conducted in air temperatures between 3.7 and 33.4 °C. The average 5 min relative humidity during measurements varied between 10% and 59% (Table S5) across all measurements, reflective of the dry atmospheric conditions of Northern Colorado.

### *3.2. Emissions*

#### 3.2.1. Calculated Emissions

Emissions for each downwind measurement were calculated using both the GP and bLS approaches. Emissions calculated using the GP approach range from $4.2 \times 10^{-27}$ to 14.7 kg $CH_4$ $h^{-1}$, with median value of 0.92 kg $CH_4$ $h^{-1}$ and mean of 2.03 kg $CH_4$ $h^{-1}$. Corresponding bLS estimates range from 0.043 to 22.52 kg $CH_4$ $h^{-1}$ (in Supplementary Materials SI.xlsx), with median value of 1.42 kg $CH_4$ $h^{-1}$ and mean of 2.7 kg $CH_4$ $h^{-1}$.

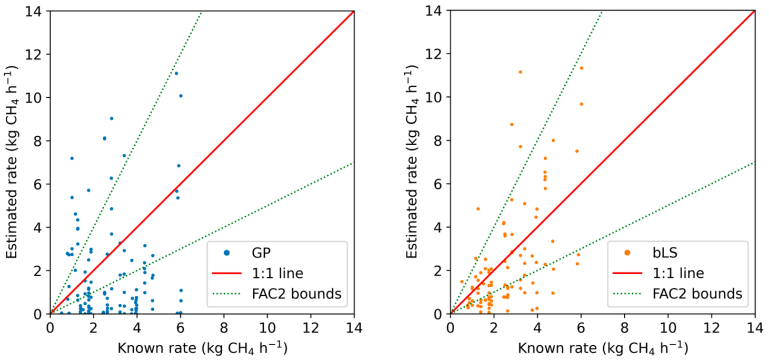
#### 3.2.2. GP and bLS Performance Analysis

Overall, 25% and 52% of emissions calculated using the GP approach and the bLS model were within the FAC2 bounds, respectively (Table 1). This percentage was calculated from the total emissions within the FAC2 bounds and the number of samples, with a total number of samples for GP and bLS of 121 and 103, respectively. The geometric means (GM) of the GP approach are 0.08 and bLS is 0.63. The median factor of error (MFoE) of the GP and bLS models are 3.5 and 1.9, respectively. The MAE and RMSE were lower for bLS calculated emissions (1.5 and 2.6 kg $h^{-1}$, respectively) compared to GP predictions (2.4 and 3 kg $h^{-1}$, respectively). The MB of the bLS model has a smaller residual bias (0.04 kg $h^{-1}$) than the GP approach ($-0.8$ kg $h^{-1}$). Also, both models exhibited comparable normalized error magnitudes (NMSE = 0.89 for GP and 0.85 for bLS), but GP showed stronger underprediction (FB = $-0.74$) compared to bLS (FB = $-0.35$). It should be noted that 18 bLS simulations failed to converge despite their higher particle counts resulting in smaller sample size ($n$ = 103) compared to GP ($n$ = 121). Those 18 cases represents that the bLS approach was unable to establish sufficient particle–source connectivity to produce reliable estimate. This difference should be considered when evaluating the comparative results. Only converged cases of bLS performance are reported. Unresolved cases indicate limitations in model applicability and should be considered when interpreting comparisons. Notably, the reported bLS performance is evaluated only for converged cases, while GP comprises all measured cases. Thus, the comparison represents the model performance under conditions where bLS converged and do not include overall model applicability.

**Table 1.** Comparison of the Gaussian plume (GP) and backward Lagrangian stochastic (bLS) model performance metrics. Metrics included calculated emissions within a factor of two of the known emission rate (FAC2), the geometric mean (GM), the median factor of error (MFoE), the mean absolute error (MAE), the root mean square error (RMSE), and the absolute bias (MB).

| Metric | GP | bLS |
|---|---|---|
| FAC2 (%) | 24.8 | 51.5 |
| GM | 0.08 | 0.63 |
| MFoE | 3.5 | 1.9 |
| MAE (kg $h^{-1}$) | 2.4 | 1.5 |
| RMSE (kg $h^{-1}$) | 3.0 | 2.6 |
| MB (kg $h^{-1}$) | $-0.8$ | 0.04 |
| FB | $-0.74$ | $-0.35$ |
| NMSE | 0.89 | 0.85 |

When the emission rate calculated using the GP approach is plotted against known emission rates (Figure 2a), the low $R^2$ value (0.05) reflects the weak linear relationship between estimated and known emissions, while the gradient (m = 0.45) shows the calculated emissions are largely underestimated. The line of best fit between emissions calculated using the bLS approach to the known emission rates has a gradient of 1.26 and $R^2$ of 0.27 (Figure 2b).




(**a**) (**b**)

**Figure 2.** Calculated emissions versus the known controlled emission rate (see in Supplementary Materials SI.xlsx) using: (**a**) the Gaussian plume (GP) approach (*n* = 121); and (**b**) the backward Lagrangian stochastic (bLS) model (*n* =103).

### *3.3. Statistical Analyses*

For both modeling approaches, the highest number of calculated emissions between the FAC2 bounds was for a signal strength greater than 4 ppm, although the bLS modeled data was in better agreement (30% for GP and 69.2% for bLS). For the GP approach, the FAC2 percentage suggested a better agreement when the signal strength was greater than 2.9 ppm (Figure 3; Table S6). For the bLS modeled data, the FAC2 percentage is higher in all signal strength bins compared to GP, ranging from 27% at the lowest signal strength to 69% in the highest bin.

A.

| Variable | Bin number --------> | | | | | | |
|---|---|---|---|---|---|---|---|
| Signal strength | 1 | 2 | 3 | 4 | | | |
| Release rate | 1 | 2 | 3 | | | | |
| Stability | 1 | 2 | 3 | | | | |
| Meander | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Centeredness | 1 | 2 | 3 | 4 | | | |
| Height match | 1 | 2 | 3 | | | | |
| Distance | 1 | 2 | 3 | | | | |

B.

| Variable | Bin number --------> | | | | | | |
|---|---|---|---|---|---|---|---|
| Signal strength | 1 | 2 | 3 | 4 | | | |
| Release rate | 1 | 2 | 3 | | | | |
| Stability | 1 | 2 | 3 | | | | |
| Meander | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Centeredness | 1 | 2 | 3 | 4 | | | |
| Height match | 1 | 2 | 3 | | | | |
| Distance | 1 | 2 | 3 | | | | |

**Figure 3.** Relative difference between average FAC2 percentage (average for GP approach = 24.8%; average for bLs model = 51.5%) and FAC2 data filtered by bin for the signal strength, release rate, stability, meander, centeredness, height match and downwind distance. Red boxes indicate a FAC2 difference less than −5%, orange boxes indicate a FAC2 difference between 0 and −5%, light green indicate a FAC2 difference between 0 and +5%, and dark green indicate a FAC2 difference greater than +5%. (**A**) is for GP and (**B**) is for bLS.

## 4. Discussion

### *4.1. Comparative Performance of GP and bLS*

In this study, emissions calculated from measurements downwind of controlled-single source release using a bLS approach, when convergence was achieved, were generally more representative than emissions calculated using a GP approach. Overall, the percentage of emissions between the FAC2 bounds (±factor of 2) using the GP approach was 25% compared to 52% for emissions calculated using the bLS model (Table 1). Similarly, the multiplicative predictive error was higher for the GP data (3.5) compared to the bLS data (1.9) as was the geometric mean bias (bLS data underpredicted by 37% while the GP data were an underprediction of 92%). For this dataset, both modeling approaches underpredict emissions, but emissions calculated using the bLS model were notably closer to known controlled release rates. These comparisons reflect only converged bLS cases and exclude unresolved simulations, which represents an important limitation in practical model applicability. Therefore, although bLS performed under converged conditions, unresolved cases remain important in determining its real-world applicability. Estimates derived from GP approach showed occasional spurious high values that did not correspond to known emissions, suggesting sensitivity to assumption violations (e.g., steady-state conditions). This may highlight the important limitation of GP approach in near-field application. Error-based metrics, e.g., MAE, RMSE, and MB, also indicate the better agreement between the bLS emission data and known release rates. The MAE and RMSE were lower for bLS emission estimates (1.5 and 2.6 kg $h^{-1}$, respectively) compared to GP predictions (2.4 and 3 kg $h^{-1}$, respectively). The MB of bLS data has a small residual bias (0.04 kg $h^{-1}$), while the underprediction of the GP data is much larger (−0.8 kg $h^{-1}$). The behavior is further reflected by FB, with GP showing stronger underprediction (FB = −0.74) relative to bLS (FB = −0.35). While both approaches showed comparable overall error magnitude (NMSE ≈ 0.85–0.89), a slightly lower NMSE for bLS suggests a slight improved predictive performance under the conditions evaluated in this study.

We suggest the difference in the performance by the two modeling approaches is due to underlying fundamental physical assumptions. The GP approach is based on the steady-state transport of a plume under uniform wind conditions and homogenous terrain; these assumptions are frequently violated in O&G sites. As a result, GP modeled emissions are strongly influenced by deviation from idealized plume geometry and meteorological stability. The bLS simulates individual random particles driven by turbulent atmospheric variability, allowing it to capture the varying nature of particle trajectories within a plume. This stochastic nature of the bLS model adapts to variation in wind conditions, which leads to better near-field estimations. Unlike bLS, the GP approach depends on parameterized dispersion coefficients to represent average plume dispersion, restricting the ability to capture and simulate turbulence variations directly. As a result, bLS achieves better agreement with true emission rates, as it can simulate the behavior of the plume more realistically than the GP approach (Figures 2 and 4).

Overall, the bLS model had greater predictive reliability under conditions where model convergence was achieved, with approximately twice as many calculated emissions falling between the FAC2 bounds. While the GP approach was faster and more convenient for calculating estimations, emissions were comprehensively underestimated due to wind variability and complexity in plume geometry. Conversely, the bLS model, though computationally more intensive, more accurately estimated the emissions under unstable near-field O&G conditions when convergence was achieved.

Practically, the performance advantages of the bLS approach must be interpreted in the context of its operational constraints. While bLS yields more accurate near-field estimates, it requires a high particle count, stochastic trajectory simulations to internally

represent turbulent transport, and final inversions, making it computationally demanding and impractical for real-time monitoring applications. In contrast to bLS, the GP method, although prone to underpredicting emission estimates, remains practical for real-time application due to its ease of implementation and computational simplicity.

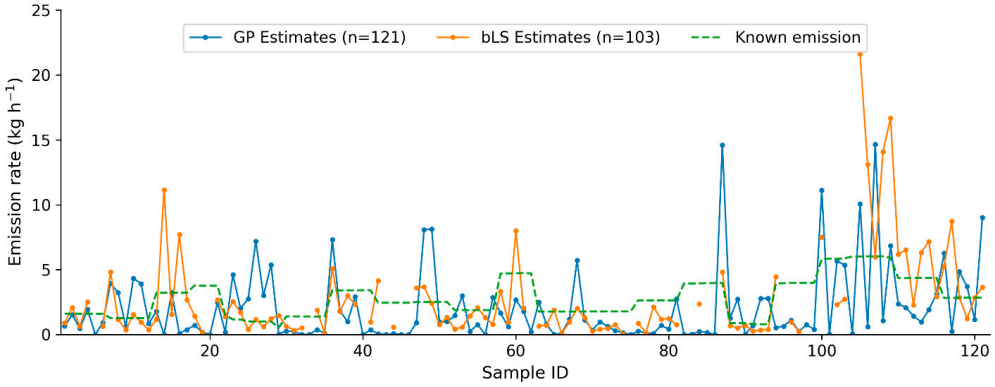



**Figure 4.** Comparison of emissions estimates from GP (blue) and bLS (orange), and corresponding known emission rates (green dash) by sample ID. Gaps in the bLS series reflect samples for which WindTrax estimates were unavailable; all estimates are aligned by sample ID.

### *4.2. Statistical Analyses*

The statistical analysis was used to identify the geometric and environmental conditions associated with improved agreement with known release rates and to provide physical interpretation for the performance differences in GP and bLS approaches. Rather than relying solely on aggregate performance metrics, model agreement with known rates was assessed using FAC2 criterion across binned subsets of signal strength, release rate, plume centeredness, height match score, wind meander and atmospheric stability.

The bLS model exhibits higher FAC2 percentage in all signal strength bins compared to GP, with percentages ranging from 26.9% at the lowest signal strength to 69.2% in the highest bin. This suggests bLS is more effective with higher signal strength. We would expect a higher FAC2 percentage with increasing signal strength, but the pattern of the results is not entirely consistent. Weaker signals typically reflect that the plume overlaps the sensor partially, limiting the detection signal, leading to inaccuracy in emission estimation [36,38]. Collectively, the result indicates that sufficient signal strength is essential for a reliable estimation, emphasizing the importance of sensor placement strategies that maximize mixing ratio enhancement above background levels.

The release rate binning describes the contrasting responses of model to source strength. Results reinforce the assertion that the performance of the GP model does not systematically improve with increasing emission magnitude. Also, the varying FAC2 percentages across release rate bins for bLS approach (45% for low, 40% for medium, and 80% for high; Table S7) does not follow a monotonic trend with increasing emissions. Although an increased FAC2 observed at the highest bin may suggest that stronger emissions enhance performance, this relationship should be interpreted with caution. The findings imply that a higher release rate alone does not necessarily improve GP-based quantification. From a measurement perspective, sensor placement determines whether increased release rates translate into improved model performance, especially for bLS, since only configurations that effectively intercept the plume allow higher emission magnitudes to result in stronger measurable signals. Therefore, the observed improvement at the highest release rate likely results from better plume-sensor alignment, rather than a direct causal influence of emission magnitude.

As per the prior studies, atmospheric stability strongly influences both models' output, as it is incorporated into the formulation [34,36,50]. The bLS model comprehensively outperforms the GP model in each PGSC category based on FAC2 percentage. But, in the outcomes from both the models, there is limited variation in the FAC2 percentage between stability classes. Therefore, variation in stability has not systematically translated into improved performance.

Wind meander score indicates whether dispersion is the degree of directional variability of plume over the averaging period and serve as a proxy of plume alignment stability. Strong variation in wind direction can disrupt the steady state assumption of GP, thereby reducing the reliability of model estimates [49,50]. The bLS model showed greater resilience as wind variability increases. The FAC2 percentage for bLS exceeded 50% in four meander bins in the bLS analysis. This increased robustness reflects the particle-based formulation of bLS, in which the wind direction variability is explicitly represented through stochastic transport. The analysis reflects that wind meander is a critical factor in dispersion modeling, indicating that the assumption of steady plume behavior over a 5 min averaging period is not consistently valid across all conditions in this dataset. It indicates that the chosen averaging duration is a practical compromise for the experiment, not as a strict, applicable approach.

When the plume centeredness score is small (0–0.5), a small deviation in wind direction may cause large swings in GP estimations [50], especially for shorter averaging intervals that may miss short-term deviations in the signal, much like the 5 min measurements used in this study. The bLS approach reflected greater robustness to centeredness but is less sensitive towards varying bins. Overall, the result underscores the importance of lateral source-sensor alignment for better estimations using GP and bLS.

The FAC2 percentage showed a differing pattern with the increase in the height match score (Table S11) for bLS and GP approaches, exhibiting contrasting values across bins. However, this trend is deceiving because the interpretation of this pattern is limited due to uneven sample sizes across bins. There are only four and one samples in the score range of 1–2 for GP and bLS, respectively, which are not sufficient for reliable statistical inference. Therefore, no robust statistical conclusion can be drawn regarding model performance with increasing height match score. Instead, in the 0–0.5 match score range, where maximum samples are available (98 and 90 for GP and bLS, respectively), the bLS had a better FAC2 percentage.

Dispersion and dilution increase with plume centerline distance, which can influence the model's predictability. At a shorter distance, the plume is more concentrated, while with increasing centerline distance, mixing is enhanced, and signal strength is also smaller [34,50]. Disregarding the bins with smaller sample sizes (>80 m; only seven out of 121 and 103 for GP and bLS, respectively), the highest FAC2 percent for both modeling approaches occur at distance of between 40 and 80 m downwind of the source. Within the conditions of this study, measurements in the 40 to 80 m downwind range were associated with improved agreement between calculated and known emissions. This observation in the study is site-specific and should be interpreted in the context of the experimental setup in a specific site including predominantly unstable daytime conditions and limited range of controlled emissions.

Overall, the statistical analyses show that near-field methane quantification performance by GP and bLS approach is driven by combined effects of signal strength, sampling geometry, and plume interception rather than by individual meteorological factors alone. For both modeling approaches, reduced agreement was associated with poor plume alignment and low signal strength, suggesting that estimates based on weak plume interception may need to be filtered from quantitative analysis. The observed variation in model perfor-

mance with downwind distance within the METEC facility indicates that measurements in the 40 to 80 m range were associated with improved agreement in this dataset; however, this reflects site-specific conditions, and interpretation is further limited by the small sample size in bins beyond 80 m.

### *4.3. Limitation and Scope*

This study was constrained by several limitations that pre-defined the scope of the experiments and gives bounds to which the results should be interpreted. First, the study is based on only one experimental site (METEC) representative of mock upstream O&G in dry windy conditions in Northern Colorado and featured a facility with single controlled releases. As such, the study is unlikely to fully represent the dispersion complexities of different aerodynamic and emissive environments. Second, all the measurements for this study were performed under unstable daytime atmospheric conditions (PGSC classes A–C) which limits the representation of stable or neutral conditions which are common on nighttime operations or during synoptic transitions like frontal passage.

Third, emission estimates calculated by the models were derived from relatively short-duration measurements (5 min). Although, this choice may enhance operational relevance for leak detection and quantification, it increases sensitivity to plume source-sensor geometry and wind meander. The use of a longer averaging interval may result in different performance behavior, particularly for GP in which the formulation relies on steady-state assumptions.

Fourth, the bLS model is influenced by stochastic sampling limitations and computational constraints. Under more turbulent atmospheric conditions, bLS generally offers more reliable estimates compared to GP, provided that sufficient particle counts and accurate meteorological inputs are used. However, calculating emissions using the bLS model takes longer than the GP approach due to its computational demands. Although particle counts were progressively increased to improve numerical stability, some samples failed to provide the estimates even at maximum particle number employed (defined as 20 million in this study).

Finally, this study does not attempt optimize model formulations or develop new dispersion algorithms. Instead, both the models were implemented using established and commonly accepted configurations to evaluate model performance in near-field conditions using the same underlying dataset. Future studies could expand the analysis presented here to include complex terrain, longer averaging durations, multiple simultaneous sources, and stable/neutral atmospheric conditions. While it is likely that the results presented here are suitable for use in leak detection technology in the Denver–Julesburg Basin, it is currently unknown if it will be useful in more vertiginous or heavily vegetated oil and gas producing regions of the Marcellus or Uinta Basins.

## 5. Conclusions

This study evaluated the performance of backward Lagrangian stochastic (bLS) and Gaussian plume (GP) dispersion approaches by comparing calculated $CH_4$ emissions to controlled release single-point emissions in near-field conditions at a simulated O&G production facility. The emission estimates were calculated using 5 min average downwind mixing ratio measurements taken at the METEC facility and were compared against the true emission rates across a range of emission strengths, meteorological and stability conditions, and plume geometries. The bLS approach showed better predictive performance from converged simulations compared to those calculated using the GP approach. The data calculated by the bLS model also had a lower multiplicative error and reduced bias relative

to GP. Other error-based metrics further confirmed the bLS model performed better, as it yielded lower RMSE and MAE than GP.

Statistical analyses of the emission data show that the lateral and vertical alignment of the source and the sensor plays a critical role in emission estimations as measurements made closer to the plume centerline and at a distance between 40 and 80 m downwind yielded the best FAC2 agreement. However, these findings are specific to the experimental conditions of this study and reflect trends within this dataset (such as limited measurement samples beyond 80 m) rather than generalizable performance characteristics. High wind meander degraded the ability of both approaches to generate representative emissions particularly with the GP approach, as it potentially violates the modeling approach's assumption of steady-state emissions. Data suggest that the emissions calculated by the bLS model are comprehensively in better agreement, but the computational demands of the modeling approach and integration into fenceline systems limit real-time applicability.

It is important to note that the conclusions in this study are limited by the experimental design, including single-source release, single test facility, short duration measurement, and unstable daytime atmospheric conditions. These factors restrict the generalizability of the results, especially for stable nighttime conditions, complex multi-source facilities, and areas with different terrain and topographical complexes, and therefore the observed relationships between sensor distance and model performance should be interpreted strictly within the scope of the METEC experimental setup.

Overall, the study clarifies the conditions under which the GP and bLS modeling approaches produce reliable near-field $CH_4$ emission estimates within the scope of experimental setup in METEC. These findings provide preliminary, site specific recommendations for measurement design and interpretation. Within the dataset of this study, the bLs approach achieved estimates within a factor of two 52% of the time within the converged simulations, and performance may improve, as calculations are made using high-concentration data of this study-specific facility. Data suggest that emissions calculated by the GP approach are likely an underestimate and that filtering for signal strength or instrument positioning will improve the estimates by a small extent. Finally, the findings in this study should be interpreted within the limitation of the experimental design and are not intended to represent and should not be generalized for applicability across varying facilities or atmospheric conditions.

**Supplementary Materials:** The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/atmos17040417/s1, SI.xlsx: Dataset containing controlled methane release experiment data used for model evaluation, including known emission rates, estimated emission rates, release heights, centerline distances, meteorological parameters (e.g., wind speed, stability class) for Gaussian plume (GP) and backward Lagrangian stochastic (bLS) apaproaches, and associated performance metrics such as FAC2, MG, MFoE, MAE, RSME, MAB. Figure S1. Meteorological tower in the METEC site. The anemometer is kept at 6.7 m above ground level, and the temperature and barometric pressure probe at 6.2 m. Table S1. Key to Solar Radiation Delta-T (SRDT) method for estimating PGSC. Table S2-i. Parameters used to calculate $\sigma_y$, where x is in kilometers. Table S2-ii. Solar Radiation Delta-T (SRDT) method for estimating PGSC Parameters used to calculate $\sigma_z$, where x is in kilometers. Table S3. Date and time of experiments. The summary under #Samples includes total measurements taken on a particular experiment, total measurements under a single release, and multiple releases. WS Criterion represents the total measurement in the corresponding experiment in which wind speed > 1 m $s^{-1}$ (* Wind speed < 1 m $s^{-1}$; note this only occurred for a single release in experiment 21). Figure S2. Example of a field deployment of methane concentration sampling. The tubing inlet connected to the gas analyzer (MGGA) and mounted to the tripod at 1 m above the ground level. Table S4. Model domain setup in WindTrax. Table S5. Key to Solar Radiation Delta-T (SRDT) method for estimating PGSC Summary of meteorological conditions for

single release measurements. Variables listed: mean and standard deviation of wind speed (U), air temperature (Tair), relative humidity (RH), mean of barometric pressure P (kPa), and sector (prevailing wind direction). The standard deviation of P is not mentioned because P never deviated more than 0.3 hPa. * Number of samples ($n$ = 121) reflects the final total filtered for single releases (low winds: U < 1 m $s^{-1}$ excluded). Table S6. Quartiles classification based on signal strength (in ppm) for models' performance analysis. FAC2 hits are the number of samples within a factor of two. $n_{bin}$ is the total number of estimates within the specified range/bin, and FAC2 hit rate (%) is the percentage of FAC2 hits on specific bins. Table S7. estimating Classification based of low, medium, and high release rates (kg $CH_4$ $h^{-1}$). FAC2 hits are the number of samples within a factor of two. $n_{bin}$ is the total number of estimates within the specified range/bin, and FAC2 hit rate (%) is the percentage of FAC2 hits on specific bins. Table S8. Number of samples under various PGSC and corresponding FAC2. Sampled data had the PGSC of classes A, B, and C, which were characterized using the SRDT's method. FAC2 hits are the number of samples within a factor of two. $n_{bin}$ is the total number of estimates within the specified range, and FAC2 hit rate (%) is the percentage of FAC2 hits on specific bins. Table S9. Wind meander classification based on its standard deviation for the models' performance analysis. FAC2 hits are the number of samples within a factor of two, and FAC2 hit rate (%) is the percentage of FAC2 hits on specific bins. Table S10. Plume centeredness score based on the ratio of crosswind offset (y) to lateral plume spread ($\sigma_y$) for the GP performance analysis. FAC2 hits are the number of samples within a factor of two, and FAC2 hit rate (%) is the percentage of FAC2 hits on specific bins. Table S11. Height match score classification based on the ratio of vertical plume spread ($\sigma_z$) to source height (z). FAC2 hits are the number of samples within a factor of two. $n_{bin}$ is the total number of estimates within the specified range/bin, and FAC2 hit rate (%) is the percentage of FAC2 hits on specific bins. Table S12. Plume centerline distance classification based on the downwind distance between source and sensor along plume axis for the model performance analysis. FAC2 hits are the number of samples within a factor of two, and FAC2 hit rate (%) is the percentage of FAC2 hits on specific bins.

**Author Contributions:** A.U.: conceptualization, investigation, methodology, field experiment, data collection writing—original draft preparation, review, editing; S.N.R.: funding acquisition, conceptualization, investigation, methodology, supervision, writing—original draft preparation, review and editing; K.B.S.: supervision, writing, review and editing; D.J.Z.: funding acquisition, project administration, supervision. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study is based upon work supported by the Department of Energy under award number(s) DE-FE0032276.

**Disclaimer:** This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author(s).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References


1. Forster, P.; Storelvmo, T.; Armour, K.; Collins, W.; Dufresne, J.-L.; Frame, D.; Lunt, D.; Mauritsen, T.; Palmer, M.; Watanabe, M.; et al. Chapter 7: The Earth's energy budget, climate feedbacks, and climate sensitivity. In *Climate Change 2021—The Physical Science Basis: Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2021; pp. 923–1054. Available online: https://www.ipcc.ch/report/ar6/wg1/chapter/chapter-7/ (accessed on 24 October 2025).
2. Saunois, M.; Martinez, A.; Poulter, B.; Zhang, Z.; Raymond, P.A.; Regnier, P.; Canadell, J.G.; Jackson, R.B.; Patra, P.K.; Bousquet, P.; et al. Global methane budget 2000–2020. *Earth Syst. Sci. Data* **2025**, *17*, 1873–1958. [CrossRef]
3. Stein, T. No Sign of Greenhouse Gases Increases Slowing in 2023. NOAA Research 2023. Available online: https://research.noaa.gov/no-sign-of-greenhouse-gases-increases-slowing-in-2023/ (accessed on 24 October 2025).
4. National Oceanic and Atmospheric Administration. Current Greenhouse Gas Levels. NOAA Climate.gov 2025. Available online: https://www.climate.gov/ghg/current-levels (accessed on 24 October 2025).
5. Omara, M.; Zavala-Araiza, D.; Lyon, D.R.; Hmiel, B.; Roberts, K.A.; Hamburg, S.P. Methane emissions from US low production oil and natural gas well sites. *Nat. Commun.* **2022**, *13*, 2085. [CrossRef]
6. Guo, J.; Zhang, J.; Shao, J.; Chen, T.; Bai, K.; Sun, Y.; Li, N.; Wu, J.; Li, R.; Li, J.; et al. A merged continental planetary boundary layer height dataset based on high-resolution radiosonde measurements, ERA5 reanalysis, and GLDAS. *Earth Syst. Sci. Data* **2024**, *16*, 1–14. [CrossRef]
7. National Academies of Sciences, Engineering, and Medicine. Methane emission measurement and monitoring methods. In *Improving Characterization of Anthropogenic Methane Emissions in the United States*; National Academies Press: Washington, DC, USA, 2018. Available online: https://www.ncbi.nlm.nih.gov/books/NBK519293/ (accessed on 3 October 2025).
8. Peng, S. Challenges and opportunities in the global methane cycle. *iScience* **2023**, *26*, 106878. [CrossRef] [PubMed]
9. Storm, I.M.L.D.; Hellwing, A.L.F.; Nielsen, N.I.; Madsen, J. Methods for measuring and estimating methane emission from ruminants. *Animals* **2012**, *2*, 160–183. [CrossRef]
10. Alvarez, R.A.; Zavala-Araiza, D.; Lyon, D.R.; Allen, D.T.; Barkley, Z.R.; Brandt, A.R.; Davis, K.J.; Herndon, S.C.; Jacob, D.J.; Karion, A.; et al. Assessment of methane emissions from the U.S. oil and gas supply chain. *Science* **2018**, *361*, 186–188. [CrossRef]
11. Lyon, D.R.; Alvarez, R.A.; Zavala-Araiza, D.; Brandt, A.R.; Jackson, R.B.; Hamburg, S.P. Aerial surveys of elevated hydrocarbon emissions from oil and gas production sites. *Environ. Sci. Technol.* **2016**, *50*, 4877–4886. [CrossRef]
12. Mbua, M.; Riddick, S.N.; Kiplimo, E.; Shonkwiler, K.B.; Hodshire, A.; Zimmerle, D. Evaluating the feasibility of using downwind methods to quantify point source oil and gas emissions using continuously monitoring fence-line sensors. *Atmos. Meas. Tech.* **2025**, *18*, 5687–5703. [CrossRef]
13. Zavala-Araiza, D.; Lyon, D.; Alvarez, R.A.; Palacios, V.; Harriss, R.; Lan, X.; Talbot, R.; Hamburg, S.P. Toward a functional definition of methane super-emitters: Application to natural gas production sites. *Environ. Sci. Technol.* **2015**, *49*, 8167–8174. [CrossRef] [PubMed]
14. Allen, D.T.; Torres, V.M.; Thomas, J.; Sullivan, D.W.; Harrison, M.; Hendler, A.; Herndon, S.C.; Kolb, C.E.; Fraser, M.P.; Hill, A.D.; et al. Measurements of methane emissions at natural gas production sites in the United States. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17768–17773. [CrossRef]
15. Pandey, S.; Worden, J.; Cusworth, D.H.; Varon, D.J.; Thill, M.D.; Jacob, D.J.; Bowman, K.W. Relating multi-scale plume detection and area estimates of methane emissions: A theoretical and empirical analysis. *Environ. Sci. Technol.* **2025**, *59*, 7931–7947. [CrossRef]
16. Worden, J.; Green, P.; Eldering, A.; Sherwin, E. *Common Practices for Quantifying Methane Emissions from Plumes Detected by Remote Sensing*; NIST Interagency Report NIST.IR.8575; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2025. [CrossRef]
17. Clarke, L.; Wei, Y.-M.; De La Vega Navarro, A.; Garg, A.; Hahmann, A.N.; Khennas, S.; Azevedo, I.M.L.; Löschel, A.; Singh, A.K.; Strbac, G.; et al. Energy systems. In *Climate Change 2022: Mitigation of Climate Change*; Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2022; pp. 613–746. [CrossRef]
18. U.S. Environmental Protection Agency. Subpart W—Petroleum and Natural Gas Systems. EPA Greenhouse Gas Reporting Program 2025. Available online: https://www.epa.gov/ghgreporting/subpart-w-petroleum-and-natural-gas-systems (accessed on 11 February 2026).
19. Rutherford, J.S.; Sherwin, E.D.; Ravikumar, A.P.; Heath, G.A.; Englander, J.; Cooley, D.; Lyon, D.; Omara, M.; Langfitt, Q.; Brandt, A.R. Closing the methane gap in US oil and natural gas production emissions inventories. *Nat. Commun.* **2021**, *12*, 4715. [CrossRef]
20. Williams, J.P.; El Hachem, K.; Kang, M. Controlled-release testing of the static chamber methodology for direct measurements of methane emissions. *Atmos. Meas. Tech.* **2023**, *16*, 3421–3435. [CrossRef]
21. Ravikumar, A.P.; Brandt, A.R. Designing better methane mitigation policies: The challenge of distributed small sources in the natural gas sector. *Environ. Res. Lett.* **2017**, *12*, 044023. [CrossRef]

22. Jacob, D.J.; Varon, D.J.; Cusworth, D.H.; Dennison, P.E.; Frankenberg, C.; Gautam, R.; Guanter, L.; Kelley, J.; McKeever, J.; Ott, L.E.; et al. Quantifying methane emissions from the global scale down to point sources using satellite observations of atmospheric methane. *Atmos. Chem. Phys.* **2022**, *22*, 9617–9646. [CrossRef]
23. Sherwin, E.D.; El Abbadi, S.H.; Burdeau, P.M.; Zhang, Z.; Chen, Z.; Rutherford, J.S.; Chen, Y.; Brandt, A.R. Single-blind test of nine methane-sensing satellite systems. *Atmos. Meas. Tech.* **2024**, *17*, 765–782. [CrossRef]
24. Barkley, Z.; Davis, K.; Miles, N.; Richardson, S.; Deng, A.; Hmiel, B.; Lyon, D.; Lauvaux, T. Quantification of oil and gas methane emissions in the Delaware and Marcellus basins using a network of continuous tower-based measurements. *Atmos. Chem. Phys.* **2023**, *23*, 6127–6144. [CrossRef]
25. Rocky Mountain Institute. Clean Energy 101: Methane-Detecting Satellites. 2023. Available online: https://rmi.org/clean-energy-101-methane-detecting-satellites/ (accessed on 11 February 2026).
26. Anand, A.; Riddick, S.; Shonkwiler, K.B.; Upreti, A.; Moy, M.; Kiplimo, E.; Mbua, M.; Zimmerle, D.J. Computational fluid dynamics-based modeling of methane flows around oil and gas equipment. *Atmosphere* **2025**, *16*, 811. [CrossRef]
27. Flesch, T.K.; Wilson, J.D.; Yee, E. Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. *J. Appl. Meteorol.* **1995**, *34*, 1320–1332. [CrossRef]
28. McGinn, S.M.; Beauchemin, K.A.; Flesch, T.K.; Coates, T. Performance of a dispersion model to estimate methane loss from cattle in pens. *J. Environ. Qual.* **2009**, *38*, 1796–1802. [CrossRef] [PubMed]
29. Riddick, S.N.; Ancona, R.; Mbua, M.; Bell, C.S.; Duggan, A.; Vaughn, T.L.; Bennett, K. A quantitative comparison of methods used to measure smaller methane emissions typically observed from superannuated oil and gas infrastructure. *Atmos. Meas. Tech.* **2022**, *15*, 6285–6296. [CrossRef]
30. Riddick, S.N. Quantifying methane emission rates using downwind measurements. *Encyclopedia* **2025**, *5*, 57. [CrossRef]
31. Ball, D.; Ismail, U.; Eichenlaub, N.; Metzger, N.; Lashgari, A. Performance evaluation of multi-source methane emission quantification models using fixed-point continuous monitoring systems. *EGUsphere* **2025**, *preprint*. [CrossRef]
32. Blackmore, D.C.; Hickey, J.-P.; Wigle, A.; Osadetz, K.; Daun, K.J. A Bayesian technique for quantifying methane emissions using vehicle-mounted sensors with a Gaussian plume model. *Atmos. Environ.* **2025**, *344*, 121002. [CrossRef]
33. Jia, M.; Daniels, W.S.; Hammerling, D.M. Comparison of the Gaussian plume and puff atmospheric dispersion models for methane modeling on oil and gas sites. *ChemRxiv* **2023**, *preprint*. [CrossRef]
34. Stockie, J.M. The mathematics of atmospheric dispersion modeling. *SIAM Rev.* **2011**, *53*, 349–372. [CrossRef]
35. United States Environmental Protection Agency. Air Quality Dispersion Modeling—Alternative Models. Environmental Protection Agency 2025. Available online: https://www.epa.gov/scram/air-quality-dispersion-modeling-alternative-models (accessed on 7 October 2025).
36. Flesch, T.K.; Wilson, J.D.; Harper, L.A.; Crenna, B.P.; Sharpe, R.R. Deducing ground-to-air emissions from observed trace gas concentrations: A field trial. *J. Appl. Meteorol.* **2004**, *43*, 487–502. [CrossRef]
37. Crenna, B. *An Introduction to WindTrax*. 2006. Available online: http://thunderbeachscientific.com/ (accessed on 24 October 2025).
38. Flesch, T.K.; McGinn, S.M.; Chen, D.; Wilson, J.D.; Desjardins, R.L. Data filtering for inverse dispersion emission calculations. *Agric. For. Meteorol.* **2014**, *198–199*, 1–6. [CrossRef]
39. Cusworth, D.H.; Duren, R.M.; Thorpe, A.K.; Olson-Duvall, W.; Heckler, J.; Chapman, J.W.; Eastwood, M.L.; Helmlinger, M.C.; Green, R.O.; Asner, G.P.; et al. Intermittency of large methane emitters in the Permian Basin. *Environ. Sci. Technol. Lett.* **2021**, *8*, 567–573. [CrossRef]
40. Day, R.E.; Emerson, E.; Bell, C.; Zimmerle, D. Point sensor networks struggle to detect and quantify short controlled releases at oil and gas sites. *Sensors* **2024**, *24*, 2419. [CrossRef]
41. U.S. Environmental Protection Agency. *Meteorological Monitoring Guidance for Regulatory Modeling Applications*; EPA-454/R-99-005; Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency: Research Triangle Park, NC, USA, 2000. Available online: https://www.epa.gov/sites/default/files/2020-10/documents/mmgrma_0.pdf (accessed on 7 October 2025).
42. U.S. Environmental Protection Agency. *User's Guide for the Industrial Source Complex (ISC3) Dispersion Models, Volume II: Description of Model Algorithms; EPA-454/B-95-003*; U.S. Environmental Protection Agency: Washington, DC, USA, 1995. Available online: https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=0000324F.PDF (accessed on 7 October 2025).
43. ABB Inc. *GLA131 Series Micro-Portable Laser Greenhouse Gas Analyzer—Technical Datasheet, Rev. E.*; ABB Measurement & Analytics: Zurich, Switzerland, 2024. Available online: https://library.e.abb.com/public/38c584d5512742aeae816d8dd62c00b3/ABB+GLA133_series+Datasheet_EN_Rev.E.pdf (accessed on 24 October 2025).
44. United States Environmental Protection Agency. *OTM 33A—Point Source Gaussian (PSG) Analysis Standard Operating Procedure, Revision 0*; Office of Research and Development, National Risk Management Research Laboratory: Research Triangle Park, NC, USA, 2013. Available online: https://www.epa.gov/sites/default/files/2020-08/documents/otm_33a_appendix_f1_psg_analysis_sop.pdf (accessed on 24 October 2025).

45. Zimmerle, D.J.; Vaughn, T.L.; Bell, C.M.; Bennett, M.L.; Duggan, G.A.; Harrison, B.K.; Schneider, J.; Robertson, A.L.; Heath, M.K.; Lyon, D.R.; et al. Methane emissions from the natural gas transmission and storage system in the United States. *Environ. Sci. Technol.* **2015**, *49*, 9374–9383. [CrossRef] [PubMed]
46. Riddick, S.N.; Mbua, M.; Laughery, C.; Zimmerle, D.J. A review of offshore methane quantification methodologies. *Atmosphere* **2025**, *16*, 626. [CrossRef]
47. Seinfeld, J.H.; Pandis, S.N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2016.
48. Chang, J.C.; Hanna, S.R. Air quality model performance evaluation. *Meteorol. Atmos. Phys.* **2004**, *87*, 167–196. [CrossRef]
49. Hanna, S.R.; Chang, J.C.; Strimaitis, D.G. Hazardous gas model evaluation with field observations. *Atmos. Environ. A* **1993**, *27*, 2265–2285. [CrossRef]
50. Hanna, S.R.; Briggs, G.A.; Hosker, R.P., Jr. *Handbook on Atmospheric Diffusion*; DOE/TIC-11223; U.S. Department of Energy: Washington, DC, USA, 1982. [CrossRef]
51. South Coast Air Quality Management District. Meteorological Data for AERMOD. South Coast AQMD 2025. Available online: https://www.aqmd.gov/home/air-quality/meteorological-data/data-for-aermod (accessed on 7 October 2025).


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.