Assessing the Performance of Emerging and Existing Continuous Monitoring Solutions under a Single-blind Controlled Testing Protocol

Fancy Cheptonui[1,2], Ethan Emerson[2], Chiemezie Ilonze[2], Rachel Day[2], Ezra Levin[2], Daniel Fleischmann[2], Ryan Brouwer[2], and Daniel J. Zimmerle[2, *]

[1]*Department of Systems Engineering, Colorado State University, Fort Collins, CO, 80523, USA*

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

*Corresponding Author: Daniel Zimmerle (dan.zimmerle@colostate.edu)

Abstract

Continuous Monitoring (CM) solutions have been identified as a method to address methane emissions at oil and gas (O&G) facilities, as these solutions can facilitate faster emission detection and repair compared to traditional survey methods. This study tested 13 CM solutions over 12 weeks using single-blind controlled testing. Controlled release rates ranged from 0.08 to 6.75 kg CH_4 hr^{-1} and lasted 18 minutes to 8 hours. Six solutions demonstrated 90 % method detection limits (DL90s) within the range of controlled releases, from 0.5 [0.3, 0.6] kg CH_4 hr^{-1} to 6.7 [5.9, 8.0] kg CH_4 hr^{-1}. Of the six solutions, four had False Positive (FP) rates between 7.8% and 14.0% (less than 15%), and four had False Negative rates (FN) of 8.0% to 34.1% (less than 50%). Compared with Ilonze et al., these results show retested solutions balancing method sensitivity with FP and FN rates. All scanning/imaging solutions had high localization (≥ 40%) precision and accuracy to the equipment unit. Eleven of the 13 solutions were tested for quantification. Single quantification estimates exhibited high relative quantification errors, outside the -90 % to 1000% range. The mean quantification error ranged from 33 [0.9, 66] %, 95% CI to 1326 [1003, 1648] %, 95 % CI for small emissions, between 0.1 and 1 kg CH_4 hr^{-1}, and from 3 [-20, 26] %, 95% CI to 3578 [-2832, 9988] %, 95% CI for larger emissions, greater than 1 kg CH_4 hr^{-1}. There were 3 and

9 solutions in the small and large emissions categories, respectively, that had less than 100% relative quantification error. Relative to previous studies, errors in quantification estimates decreased, as did FN and FP rates, with improved DL90s for two of the four retested solutions. These findings highlight that continuous, rigorous testing enhances solution performance, with notable improvements observed across multiple testing programs using the same test protocol.

Keywords

Methane, Detection, Detection Limit, Quantification, Continuous Monitoring, Probability of Detection

Introduction

Dry natural Gas (NG) production has gradually increased over the past ten years (EIA 2024). The production increased by 4%, 5 billion cubic feet per day (Bcf/d), in 2023 compared to the production in 2022 (EIA 2024). Natural gas consists of between 70 % and 90 % methane (CH_4), one of the major greenhouse gases in the atmosphere. Although CH_4 stays in the atmosphere for a shorter period than CO_2, it has a global warming potential that is 27 - 30 times higher than CO_2 in a 100-year timespan (ORD US EPA 2017; OAR US EPA 2016; IPCC 2021). Methane is emitted from natural sources, such as wetlands, and anthropogenic sources, such as energy production and agricultural activities (NASA, 2024).

Globally, emissions from the energy sector accounted for approximately 130 million tonnes of CH_4 in 2023 (IEA 2024). In the US, operations in the oil and gas (O & G) industry are the second-largest industrial source of emissions (EPA, 2024a). Federal and state regulations require operators to implement the leak detection and repair (LDAR) program (CDPHE; EPA, 2024). The program aims to reduce intentional (vented) and unintentional (fugitive) methane emissions (Zimmerle, Dileep, and Quinn 2024) from the O&G industry in the US and Canada.

Current regulatory-approved methods for LDAR programs, such as the EPA method 21 and OGI camera surveys, can provide precise component- or equipment unit-level measurements (EPA,2024b). However, they are expensive and labor-intensive when scaled over thousands of O&G facilities. Also, studies have shown that source, facility-level, and annually assessed emissions exhibit spatial and temporal variability (Brandt et al. 2014; Lavoie et al. 2017; Vaughn et al. 2018; Huang et al. 2024). These variabilities make it challenging for surveys to estimate the long-term size of emissions using snapshot measurements from a facility. Further, studies show that super-emitters skew the emissions distribution (Collins et al. 2022; Huang et al. 2024; Schuit et al. 2023; Zavala-Araiza et al. 2015). Super emitters emit at rates greater than 100 kg hr^{-1} (EPA, 2024b). Detecting such emitters using traditional methods is challenging because they are usually intermittent.

Regulations (EPA, 2024b; CDPHE, 2024; European Parliament, 2024; EPA, 2023) and research efforts have motivated the development and testing of new and existing Leak Detection and Quantification (LDAQ) technologies (Bell et al., 2020, 2023; Chen et al., 2024; Ilonze et al., 2024; Liu et al., 2024; MONITOR| Arpa-e, 2014; Ravikumar et al., 2019; Siebenaler et al., 2016; Titchener et al., 2022). LDAQ technologies, such as Continuous Monitors (CM), promise to address some emission measurement concerns by cost-effectively providing temporally resolved monitoring of O&G facilities. These technologies are scanning, imaging, or point sensors installed at O & G facilities to monitor NG or CH_4 emissions for a long-term period. The major advantages of CM solutions include the capability to detect large emissions, sustained operational capacity, and provision of temporal data. Despite these advantages, CM solutions cannot be used for regulatory LDAR programs unless they demonstrate equal or better emissions reduction than existing LDAR survey methods (EPA, 2024c).

The EPA's final rule stipulates the performance CM solutions must meet to be approved as an alternative approach for regulatory LDAR programs (EPA, 2024b). According to subpart OOOOb 40 CFR part 60 by the EPA, the detection threshold for CM systems must be at least 0.4 kg hr^{-1} of CH_4 over facility-specific background emissions (EPA, 2024b). Operators already use CM solutions for internal emissions accounting and voluntary methane emissions reduction programs like OGMP 2.0 (OGMP 2024). However, more rigorous testing is needed to improve confidence in output data (such as the 90% method detection limit (DL90), quantification accuracy, detection, and localization) and track solutions' performance trends over time.

Colorado State University's (CSU) Methane Emissions Technology Evaluation Centre (METEC) developed a standardized, controlled testing protocol for the Advancing Development of Emission Detection (ADED) project. The project was developed to assess the performance of CM and survey solutions (Bell and Zimmerle 2022). Starting in 2022, the protocol has been utilized for an extended CM test from February to April each year. The 2022 testing is documented in Bell et al., 2023, the 2023 testing in Ilonze et al. (2024), and the 2024 testing in this paper.

The 2022 test showed that CM solutions exhibited quantification errors over ± 60% for single emission estimates. One of the primary metrics assessed for all the CM solutions is the DL90, which represents the emission rate at which a deployed solution detects emissions 90% of the time (Bell et al. 2023). In 2023 testing, retested solutions showed decreased DL90 and FP rates (the percentage of detection reports from a solution that incorrectly attributes a controlled release). However, there were consistently wide quantification errors for single emission estimates. The 2023 testing also indicates that solutions with the lowest DL90 had the lowest FP and FN rates (the percentage of controlled releases not identified by a solution). This contrasts with the expectation that raising detection limits to lower FP and FN rates is necessary.

The current study (2024 testing) tested 13 CM solutions, including four solutions that participated in the 2023 testing (Ilonze et al., 2024), and 2 of the four solutions and 1 of the thirteen were also tested in 2022 (Bell et al. 2023). The solutions were tested at CSU's METEC facility in Colorado, US. Solutions were categorized into point sensor networks (PSNs) and scanning/imaging solutions. The PSN solutions consist of point sensors deployed at/on the facility's equipment to detect the presence of hydrocarbons in the atmosphere (Day et al. 2024). Scanning/imaging solutions deploy scanning lasers and short/mid-infrared cameras that generate 2D images of gas plumes at the deployed locations (Bell et al., 2023; Ilonze et al., 2024). Algorithms are used with meteorological data to produce quantification and localization information. Performance metrics such as the probability of detection (POD), quantification accuracy, localization accuracy, and precision of a solution are presented in this study and compared to the performance of solutions from 2023 (Ilonze et al., 2024) and 2022 (Bell et al. 2023).

Methodology

Testing Facility

The solutions were tested between February and April 2024 at CSU's METEC facility. Long-duration testing was conducted depending on the facility's availability. Additionally, testing is aimed at various conditions, as seen in winter and spring. The facility is an 8-acre (3 ha) outdoor facility for testing and improving CH_4 emissions LDAQ solutions. The facility consists of 7 wellpads, complete with retired O & G equipment, including, but not limited to, wellheads, separators, and tanks (SI Figure S-1). For this study, wellpads 4 and 5 were used for controlled releases of compressed natural gas. Wellpad 4 consists of 5 wellheads, four horizontal separators, and three tanks, while Wellpad 5 consists of 3 wellheads and three vertical separators (SI Figure

S-3). The tanks, wellheads, and separators have unique emission points (EPs: with unique tag identifiers) and flow meters that control the gas flow.

ADED Continuous Monitoring Protocol R1.0

Overview: The first revision of ADED's controlled testing protocol (Bell & Zimmerle, 2022) was developed in September 2020 by METEC. The protocol received contributions from stakeholders in the O & G industry, LDAQ technology developers, regulatory bodies, non-governmental agencies, and state and federal environmental agencies. The protocol employs single-blind, controlled testing to evaluate the performance of solutions in a rigorous and reproducible manner. This testing allows for inter-solution comparison and year-over-year performance evaluation. In a single-blind test, the test center knows the emission rates and locations, while the performers (solution providers) are unaware of this information. Testing is conducted for extended periods, typically months, to assess the performance of a solution over a wide range of emission rates with varied environmental and meteorological conditions.

According to the protocol, a solution consists of one or more natural gas (NG) or CH_4 sensors, data analytics, and a data management system. The analytics convert raw sensor readings to quantification and localization data and infer the 1.) presence or absence of an emission, 2.) the source emission rate, and 3.) location of an emission. The data management system reports detections, localizations, and applicable quantification data.

Experimental Process

Compressed NG consisting of 84.2% CH_4, 13.3% C_2H_6, 2.0% C_3H_8, and a trace amount of C_4+ was released from storage gas cylinders. Emission points released gas at rates ranging from 0.08 to 6.75 kg CH_4 hr^{-1}. Flow meters assigned to each equipment group controlled gas flow through

the EPs on Pads 4 and 5. Testing involved a series of experiments with the number of EPs ranging from 1 to 5 per experiment (SI Figure S9). The EPs were located on different equipment units for experiments with multiple EPs. The duration of each experiment ranged from 18 minutes to 8 hours. The durations were selected to account for the duration of vented emissions and those that may last for extended periods without being repaired.

Experiments ran 24 hours a day for 5 to 7 days per week. The controlled releases started and ended simultaneously for experiments with multiple EPs. Multiple controlled releases are designed to test the capability of a solution to attribute an emission to a particular location. Individual experiments were separated from the preceding experiment by a no-emission period lasting from 30 min to 4 hours (SI Figure S-9). The no-emission period enabled solutions to potentially identify the start and end of experiments when the concentration levels were reduced to the original background concentration. Additionally, the no-emission periods were selected based on the time it takes for the concentration in the atmosphere to return to the background; depending on prevailing atmospheric conditions, this time varies.

The test facility assigned an experiment ID to an experiment and recorded the start and end time, the duration, the metered flow rate, the uncertainty of the metered flow rate, the gas composition (as measured using gas chromatography), and the meteorological conditions. Regularly, the team processed the detection reports of various solutions to determine the emission rates and durations that had not been covered while testing (small, large, and longer-duration releases).

Reporting: Performers submitted detection reports to METEC's email reporting system for evaluation using the protocol's performance metrics (Bell & Zimmerle, 2022). Detection reports notify the test center of identified individual emission events, the start time, the location, and the emission rate of a source. Each report should be attributed to a single controlled release in an

experiment with multiple controlled releases for proper detection classification. Mandatory information in the detection report includes the *DetectionReportID*, *DetectionReportDateTime*, *EmissionStartDateTime*, *EmissionSourceID*, *Gas* (the species monitored by the solution), and *Equipment Unit* (SI Section S5). The protocol allows performers to update a detection by sending new reports referencing the same *EmissionSourceID* from a previous report. This provision enables performers to change data, such as the estimated emission rate, the timing of a controlled release, and the GPS coordinates, to reflect the best solution estimates. The update was permitted before the test center sent the performance results to the performers.

In addition to the detection reports, performers submit offline reports whenever their solution is not operating during testing. The offline reports state the offline start and end times, allowing the exclusion of offline data during analysis. Additionally, the test center records all maintenance events at the site, such as gas house filling and fixing faulty equipment. All the detection reports sent during maintenance events are excluded from the analysis.

Before classification, specific data were removed from the analysis, including 1) data when the performer reported their system was offline, 2) data that occurred outside the testing window, 3) data that the performer stated the equipment unit as being off facility, i.e., outside the test facility's fence line, and 4) data that occurred during maintenance events as shown in the maintenance log sheet. As stipulated in the protocol (Bell & Zimmerle, 2022), the classification process was as follows:

1.) All the controlled releases conducted within the test facility were sorted by equipment unit identifier and then by the flow rate, if reported by the performer, in descending order.
2.) From the first step above, the remaining controlled releases were sorted by equipment group identifier and then by flow rate, if reported by a performer, in descending order.

3.) Finally, the remaining controlled releases were sorted by flow rate, if reported by a performer, in descending order. After this step, all the controlled releases and unpaired detections were classified as FN and FP detections, respectively.

In steps 1, 2, and 3 above, all the detections identifying the sorted equipment unit, group, or facility as the emission source were selected, respectively. The selected detections whose emission start time was between the controlled release start and end time were filtered and sorted by flow rate, if reported by a performer, in descending order. Finally, the uppermost detection was paired with a controlled release as a TP detection with correct unit, group, or facility-level localization precision for steps 1, 2, and 3, respectively. After each step above, a matched pair was removed from further classification.

Performance Metrics: Controlled releases within the testing period were paired with detection reports to evaluate the performance of the solutions (SI Section S8). The detection reports were classified as true positives (TP) or false positives (FP). According to the protocol, a TP occurs when a performer's detection report is attributed to the correct and coinciding controlled release. A false positive occurs when a detection report incorrectly attributes a controlled release within the test facility. During the second (Ilonze *et al.*, 2024) and current protocol implementation, FPs were further classified into FP-no controlled release and FP-excess. An FP-no controlled release occurs when a detection report identifies an emission with no controlled release at the test facility. An FP-Excess occurs when a detection report identifies a controlled release already paired with a different detection report. A controlled release is classified as a false negative (FN) when there is no detection report to pair with it.

The current study used the primary performance metrics, including the probability of detection (POD), localization precision and accuracy, quantification accuracy, and detection time, to assess

a solution's performance. POD is a fraction of controlled releases under a set of independent variables that were classified as detections. Independent variables include the controlled emission rate and prevailing meteorological conditions. The localization accuracy at various levels (equipment unit, equipment group, and facility levels) is the ratio of the TP detections to the sum of the TP and FP reports. Localization precision of each TP detection was assigned to three levels of precision (i.e., equipment unit, equipment group, and facility levels).

A precise equipment unit localization occurs when a performer correctly attributes a detection to an equipment unit where the controlled release occurred. For instance, if a controlled release occurred from the first wellhead unit on pad 4 (i.e., 4W-1), and the report identifies 4W-1, this TP detection will have a localization precision to the unit level. However, if the controlled release occurred on the third wellhead unit in pad 4 (4W-3), the TP detection will have a localization precision to the equipment group. For the solutions that reported the source emission rate, the quantification accuracy was assessed by computing the relative error between the controlled release rate and the reported emission rate. Additionally, quantification accuracy was assessed by analyzing the percentage of single estimates within a factor of 3. The detection time is the time difference between the start of a controlled release and when a performer submitted the first detection report.

ADED Participating Solutions: Spring 2024

Participants were recruited through an open invitation on the METEC facility's website and by contacting solution developers identified during protocol development. Ten performers deployed their solutions at wellpads 4 and 5 before the start of controlled testing. Two performers deployed more than one solution for testing, summing up to 13 CM solutions. Some solutions monitored a subset of equipment groups, while others deployed their solutions in all the equipment groups in

wellpads 4 and 5 (SI Figure S-3). After deployment, performers operated their solutions remotely except when their hardware malfunctioned; in this case, a performer's team would fix their hardware during 'no release' periods.

In alphabetical order, the participating performers were Blue-Rock, ChampionX, Earthview, Honeywell, Oiler-Equation, Project-Canary, Sensirion, Sensit, Shepherd-Safety, and SLB. Table 1 shows all the solutions deployed during testing, identified with a unique solution ID. Three solutions tested in the current study were scanning/imaging solutions, while 10 were PSNs. As shown in Table 1, four solutions were tested in 2023 (Ilonze *et al.*, 2024) relative to the current study (solutions N, F, D, and P). Solutions C, D, and F were also tested in the study by Bell *et al.*, (2023). All the solutions were deployed across Pads 4 and 5 except solution U, which was only deployed on the tanks in Pad 4.

Table 1.
Participating solutions in the current study (2024), from 2023 (Ilonze et al., 2024), and 2022 (Bell et al. 2023) testing, in alphabetical order of the solution ID.

Solution ID	Sensor Type	Sensor Count			Detection	Quantification
		2024 (this study)	2023 (Ilonze et al)	2022 (Bell et al)		
C	PSN	12	-	6	✓	✓
D	PSN	8	8	8	✓	✓
F	PSN	9	10	8	✓	✓
I	Scanning/Imaging	1	-	-	✓	✓
M	Scanning/Imaging	7	-	-	✓	✓
N	PSN	50	18	-	✓	✓
P	PSN	14	6	-	✓	×
R	PSN	11	-	-	✓	✓
S	PSN	12	-	-	✓	✓
T	PSN	7	-	-	✓	✓
U	PSN	3	-	-	✓	×
V	Scanning/Imaging	1	-	-	✓	✓
W	PSN	35	-	-	✓	✓

The table shows two different types of solutions that were tested: the PSNs and scanning/imaging solutions. The current study tested 13 solutions, with 4 of the 13 tested in (Ilonze et al., 2024), 2 of the 4, and 1 of the 13 tested in (Bell et al. 2023).
✓ means the performer reported the data necessary for detection or quantification.
× means the performer did not report the data necessary for detection or quantification.

Data Analysis

Classified data was divided into various quantiles, ranging from 30 to 50 points per bin. The number of points was determined using quantile-based discretization, which produced the highest *goodness of fit* (R^2) for a solution. This approach ensured that each bin used when evaluating the POD curve had an equal sample size of data classified as TP or FN—hence, approximately equal statistical significance. In quantile-based discretization, balanced data bins were generated for various test conditions. A detection fraction, a ratio of the number of TPs to the sum of the number

of TPs and FNs (nTPs /(nTPs + nFNs)), was derived from each bin. The detection fractions from all the bins were fit to a POD curve using power functions; the intercept was set to zero for these curves. The POD curve was calculated as a function of the source emission rate, duration, and wind speed.

Multiple independent variables could impact the probability of detecting emissions. Therefore, a multivariable logistics regression model was fitted to the detection data to assess the impact of independent variables (wind speed, sensor location relative to the source, wind direction, release duration, temperature, and emission rate) on the probability of a solution detecting emissions (SI Section 8.2). Information on the analysis can be found in the SI sections S7, S10.1, and S10.2.

We applied a 5-minute buffer time on the maintenance end time to account for any residual emissions in the atmosphere and to avoid assigning an FP detection to such reports. Additionally, we applied 20-minute buffer times to the start and end times of the experiments to maintain consistency with previous studies and to track the performance of solutions that have been retested effectively.

Protocol Constraints

The performance of the tested solutions is only representative of the test conditions; therefore, the results cannot be extrapolated to untested conditions. Tested conditions represent the site's structural configuration, prevailing meteorological conditions, duration of controlled releases, and the range of emission rates spanned during testing. Untested conditions include emissions longer than 8 hrs, wind speed greater than 13m s-1, and, generally, atmospheric stabilities during summer. For instance, some CM solutions rely on the accuracy of measurements from emissions that run for extended periods. Since the controlled releases in this study ran for up to a maximum of 8 hrs, the results may not cover the full capabilities of such solutions. Therefore, these untested

conditions may limit the work's applicability in this study. Testing at the site configuration resembled a simplified upstream production facility. Therefore, results cannot be extrapolated to more complex production facilities.

Emission rates ranging from 0.08 to 6.75 kg CH_4 hr^{-1} were tested; however, for most production facilities, there can be larger emission rates not tested in the study due to safety reasons. The protocol implementation requires testing the solution for component emission detection and quantification. However, some solutions operate as full facility monitors and may provide facility-level emission rates rather than detection at the component level. The testing in this study involved steady-state emission rates; however, this may not be the case in production facilities.

The detection reports in this study are human-influenced. The protocol requires the performers to infer the presence/absence of emissions based on sensor readings rather than sending the sensor readings directly to the test center. As this is the third implementation, we acknowledge that repeated testing by some solutions may result in well-trained solution algorithms, and as such, detection, localization, and quantification results may have improved substantially for such solutions.

Results

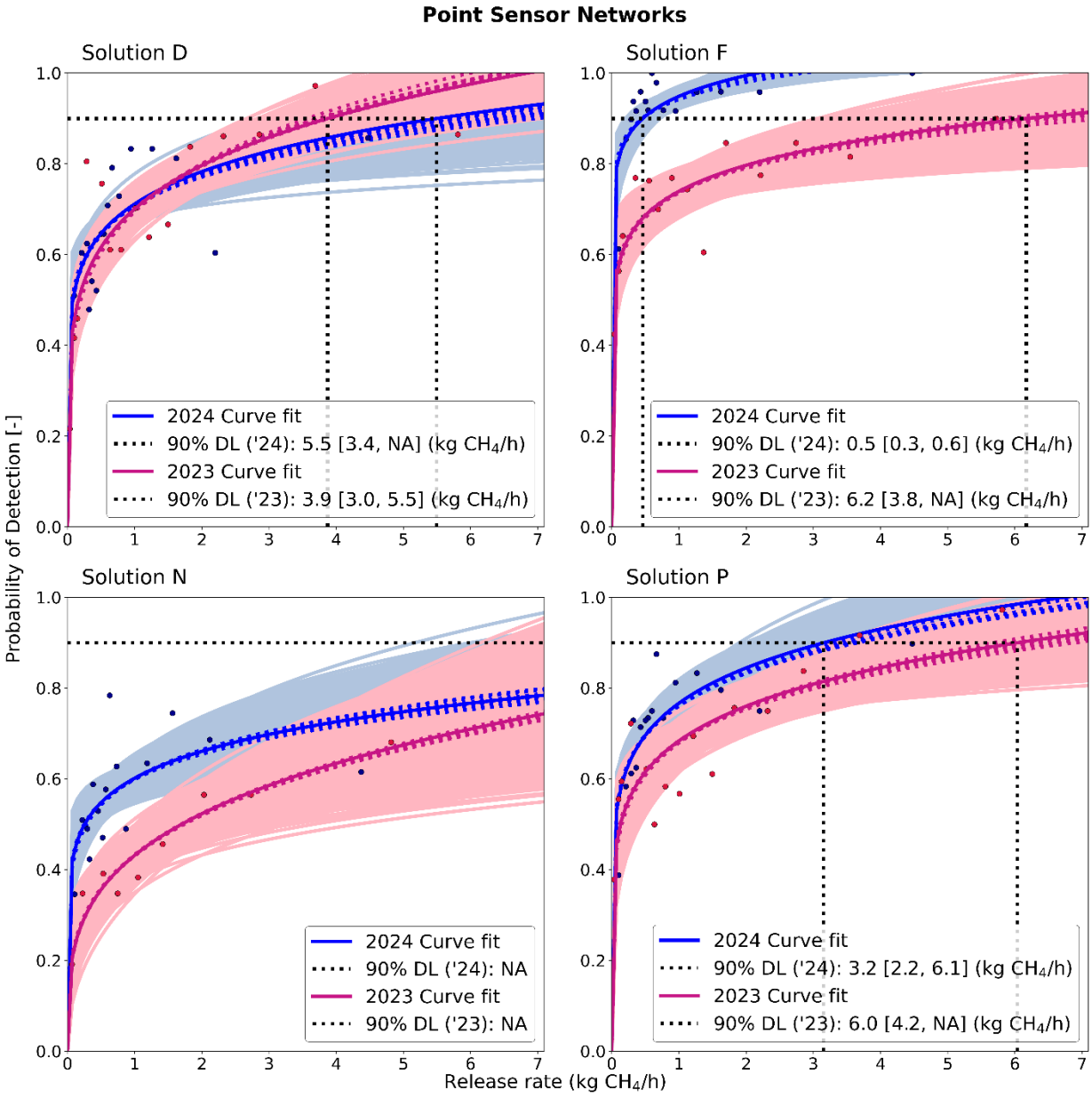
Overview of the Results

Results of a multivariable logistic regression model showed that the emission rate significantly ($p<0.05$) affected the probability of all solutions to detect emissions except for solution T (SI Table S13). Therefore, POD curves were generated as functions of the emission rate. Other variables tested for their impact on the POD include the release duration, ambient temperature, wind speed, wind direction, and distance to the closest sensor (Tables S4-S16). The DL90 represents the emission rate at which, results indicate, the solution could detect emissions in 90% of cases,

regardless of the prevailing meteorological conditions (Bell et al., 2023; Ilonze et al., 2024). In the figures, the DL90 was labeled as NA if it was outside the tested release rate range (0.08 – 6.75 kg CH_4 hr^{-1}).

Figure 1 compares the POD curves for solutions retested in this study relative to Ilonze *et al.*, (2024). Figures 2 and 3 show the POD curves for the newly tested solutions. In Figure 1, the red (data points from 2023 testing) and blue (data points from the current testing) markers represent the fraction of controlled releases in a bin classified as TPs. The faded curves in all the plots were derived from bootstrapping the data points to obtain the uncertainty of the POD curve (SI Section S10.1).

Figure 1.



The probability of detection against the release rate for PSNs tested in 2023 and 2024. The 90 % DL for both years is based on the controlled release rate range for the current testing (0.08 to 6.75 kg CH_4 hr^{-1}). The DL90s of three of four solutions (Solutions D, F, and P) are within the current study's range of tested release rates, while 1 is outside the tested release rate range (Solution N).

Probability of Detection: In this section, we discuss the solutions' PODs in two categories: PODs within and those outside (> 6.75 kg hr^{-1}) the range of the tested emission rate. Additionally, we perform an inter-solution comparison of the PODs (PSN versus scanning/imaging solutions).

Generally, figures 1, 2, and 3 show that 6 out of 13 solutions had DL90s ranging from 0.5 to 6.7 kg CH_4 hr^{-1}. For the 6 solutions, results indicate that these solutions balance method sensitivity and low FP and FN rates (as in (Ilonze et al., 2024)). Table 2 shows that 5 of the six solutions had the lowest FP rates (less than 20%: 8.0 to 18.9%), and 4 had the lowest FN rates (less than 50%: 8.0-34.1%) (SI Section S8.1).

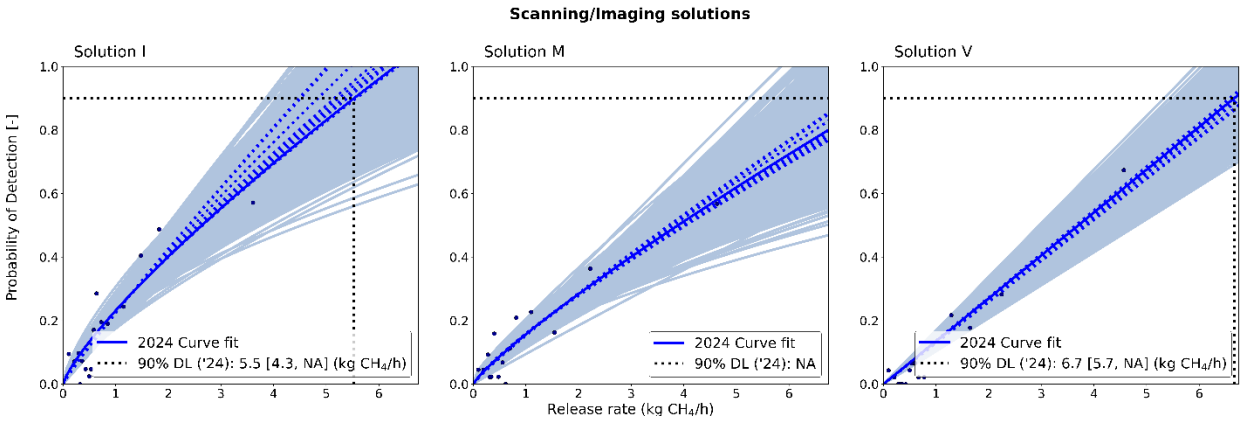
Seven of 13 solutions with DL90s outside the controlled release rate range struggle with high FP and FN rates (Table 2). Four of the seven solutions had the highest FP (>20%) and FN (>50%) rates. Most of the tested variables affected the PODs of the 7 solutions, i.e., the distance to the closest sensor, wind speed, duration, wind direction, and temperature (SI Section S8.2). As for these solutions (N, C, U, M, R, T, and W), their DL90s should be used cautiously as they fall outside the range of tested release rates, 0.08 – 6.75 kg CH_4 hr^{-1}.

Although all solutions had less than 50% FP rates, the percentage of FPs further classified as excess reports for 10 of 13 solutions was high (≥50%). These solutions sent multiple detection reports identifying already paired controlled releases. Multiple emissions identification could be costly, as it would send misleading alerts for repair. Solutions N and W, which have the lowest FP rates overall (4 % and 5 %, respectively), deployed the largest number of sensors (50 and 35, respectively).

As for the FNs, 6 of 13 solutions had FN rates greater than 50%. All scanning/imaging solutions had the highest FN rates (> 80%). A high FN rate indicates the likelihood of missing emissions, which could delay repairs. For the scanning/imaging solutions, the scan time of a solution impacts

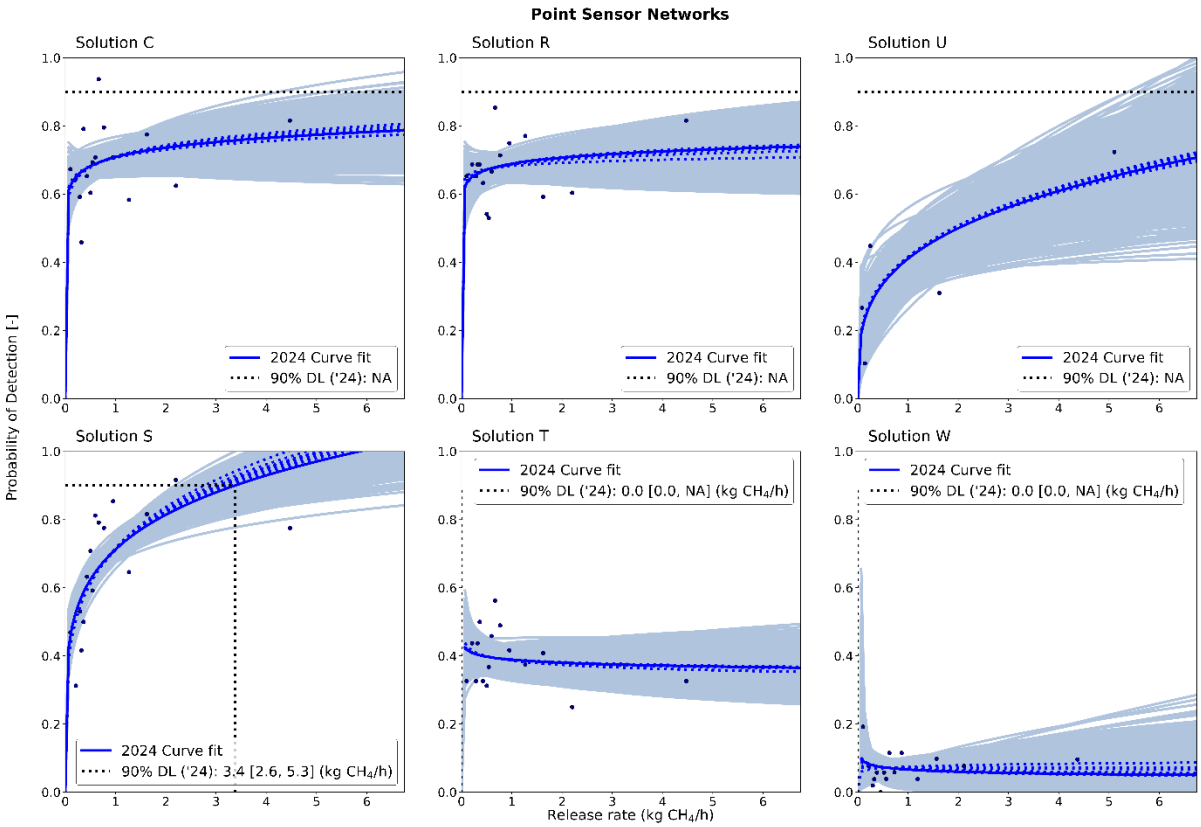
how fast a solution identifies an emission. Scanning/imaging solutions also deployed low sensor counts (1 sensor each for solutions I and V, and 7 sensors for solution M; Table 1). This means that they have a likelihood of missing emissions.

Figure 2.



The probability of detection against the controlled release rate for scanning/imaging solutions. The DL90 of one of three solutions (Solution M) is outside the range of tested release rates; therefore, this is presented as NA.

Figure 3.



The probability of detection against the controlled release rate for the newly tested PSNs. The DL90 of one solution is within the range of tested release rates. The DL90s for solutions C, R, T, U, and W are outside the tested emission rate range.

Inter-solution POD comparison: The DL90s for 4 PSNs were within the tested emission rate range, ranging from 0.5 to 5.5 kg CH_4 hr^{-1}, somewhat better than the scanning/imaging solutions. The DL90s for two scanning/imaging solutions (I and V) were 5.5 and 6.7 kg CH_4 h^{-1}, respectively. The DL90 for the other scanning/imaging solution (M) could not be estimated; therefore, this is presented as NA. The 4 PSNs deployed more sensors than scanning/imaging solutions, and had low FP and FN rates, less than 20% and 50%, respectively. Additionally, 3 of the 4 PSNs were previously tested by Ilonze *et al.*, (2024) hence, better performance than scanning/imaging solutions. The DL90s for the remaining 6 PSNs (5 newly tested) could not be estimated or did not produce data that reflected enough trends to estimate a DL90.

Finally, the curves for POD as a function of emission rate for all scanning/imaging solutions reflected a quasi-linear relationship. This suggests a stronger POD and emission rate relationship than the independent variables tested.

Table 2.
A summary of the solutions' DL90, overall true positive rate (TP), overall false negative rate (FN), overall no-controlled release, and excess reports- false positive rates (FP) (SI Section S8).

Solution ID (Type)	90 % DL (kg CH_4 hr^{-1})	DL Within Tested Rates	FP			TP (%)	FN (%)
			Overall (%)	No Controlled Release (%)	Excess Reports (%)		
F (PSN)	0.5 [0.3, 0.6]	Yes	14.0	7.8	92.2	92.0	8.0
P (PSN)	3.2 [2.2, 5.6]	Yes	8.0	6.1	93.9	72.3	27.7
C[a] (PSN)	NA	No	24.0	27.8	72.2	69	30.8
R[a] (PSN)	NA	No	34.0	3.0	97.0	67.7	32.3
D (PSN)	5.5 [3.3, 16.0]	Yes	7.8	15.9	84.1	67.1	32.9
S (PSN)	3.4 [2.5, 5.2]	Yes	18.9	5.0	95.0	65.9	34.1
N[a] (PSN)	NA	No	3.7	94.1	5.9	56.8	43.2
T[a] (PSN)	NA	No	40.5	33.2	66.8	39.5	60.5
U[a] (PSN)	NA	No	34.1	78.6	21.4	37.7	63.0
I (Imaging)	5.5 [4.3, NA]	Yes	22.8	35.1	64.9	18.7	81.3
M[a] (Imaging)	NA	No	5.1	20.0	80.0	14.2	85.8
V (Imaging)	6.7 [5.9, 8.0]	Yes	12.8	50.0	50.0	10.9	89.1
W[a] (PSN)	NA	No	5.1	100.0	0.0	7.2	92.8

The table is presented in decreasing order of the TP rate. Overall, FP in the table represents the fraction of the detection reports classified as false positives. The FP detection reports were further classified as a no-controlled release or an extra report. For the FP no-controlled release, this is a detection report sent in when the test center was not releasing gas. The FP extra report is when a solution sends multiple detection reports linked to an already TP-classified detection report.
[a]Solutions with no observable POD trend with emission rates or whose DL90 is outside the range of tested rates. These POD curves and DLs should be used with caution. The POD for solutions T and W did not exceed 60% and 20% in any bin, respectively (Figure 3).

Comparing the POD Performance in the Current Study to the Two Previous Studies

Logistic regression models were used to generate the POD curve and to calculate the Method Detection Limits in Bell *et al.*, (2023). This is referred to as the DL90 in this study and in Ilonze *et al.*, (2024). The controlled releases in the studies by Bell *et al.*, (2023) and Ilonze *et al.*, (2024)

ranged from 0.4 to 6.4 kg CH_4 h^{-1} and 0.006 to 7.1 kg CH_4 h^{-1}, respectively. The DL90 ranged from 2.7 to 30.1 kg CH_4 h^{-1} in Bell *et al.*, (2023) and from 3.9 to 18.2 kg CH_4 h^{-1} in Ilonze *et al.*, (2024). Six of 11 and four of 9 solutions had DL90s within the range of controlled release rates in Bell *et al.*, (2023) and Ilonze *et al.*, (2024), respectively.

Ilonze et al. relative to this study: The DL90s for the three retested solutions showed substantial improvements, though with increased variability. The DL90s for these solutions ranged from 3.9 to 14.1 kg CH_4 hr^{-1} in Ilonze *et al.* (2024). By the current study, these DL90s ranged from 0.5 to 5.5 kg CH_4 hr^{-1}, with notable gains for solutions F and P. Solution F's DL90 improved from 6.2 [3.6, 17.4] kg CH_4 hr^{-1} in Ilonze *et al.*, (2024) to 0.5 [0.3, 0.6] kg CH_4 hr^{-1} in the current study. Solution P's DL90 improved from 6.0 [4.1, 11.8] kg CH_4 hr^{-1} to 3.2 [2.2, 5.7] kg CH_4 hr^{-1} in the current study. The two solutions' improvement in performance can be attributed to continuous testing, improvement in the solution algorithm, and an increased number of sensors in solution P.

Solution F in the current study had a DL90 above EPA's detection threshold (0.4 kg h^{-1}) (US EPA, 2024) for continuous monitoring systems by 0.1 kg CH_4 h^{-1}. The EPA regulation requires detecting a facility's fugitive emissions of 0.4 kg CH_4 hr^{-1} in the presence of baseline emissions, i.e., routine emissions common to most O&G sites. However, the current testing protocol has no baseline emissions.

Bell et al. relative to this study: The DL90 for solutions tested in the current study, relative to Bell et al. (2023), partially improved. Three solutions, C, D, and F, were also tested by Bell *et al.* (2023). The DL90 for solution C could not be estimated in the current study, although the number of sensors doubled. The DL90 remained at 5.5 kg CH_4 h^{-1} for solution D and decreased from 3.4 [2.4, 5.4] kg CH_4 h^{-1} to 0.5 [0.3, 0.6] kg CH_4 h^{-1} for solution F. Although the DL90 for solution D remained constant, the upper and lower bounds had a wider range than during testing by Bell *et*

al. (2023). Additionally, solution D has deployed a constant number of sensors since the protocol's first implementation; therefore, the steady performance is attributed to improving the solution algorithm and continuous testing.

Localization Precision and Accuracy

The protocol required the performers to submit the equipment unit where the solution detected an emission. Localization used the reported equipment unit as a primary metric to establish the precision of each TP detection (SI S9.2). As for the localization accuracy at the equipment unit, this was calculated as a ratio of the number of TP detections at the equipment unit level to the total number of reported detections (nTP_{EqUnit} / (nTP + nFP)) for each solution (SI Section S9.2). We define sensor density as the ratio of the surface area (in hectares) of the test center, i.e., pads 4 and 5, to the number of sensors deployed by a performer. Scanning/imaging solutions had higher sensor densities (ha/sensor) as their sensors monitored larger areas than PSNs. The sensor densities for scanning/imaging solutions (I, V, and M) were 0.12, 0.85, and 0.85 ha/sensor. For PSNs, the sensor density ranged from 0.02 to 0.28 ha/sensor.

Overall, most solutions (7) presented high precision (> 50%) at the unit level. Although scanning/imaging solutions have higher sensor density, results show that they localize emission sources better than PSNs. All scanning/imaging solutions indicate high precision (> 50%) and accuracy (> 40%) at the unit level (Table 3; SI Figure S-21). The localization precision of scanning/imaging solutions (solutions I, M, and V) ranged from 56.0 % to 84.9%, and the localization accuracy ranged from 43.2% to 80.6% at the unit level. As shown in Table 3, for the PSNs at the unit level, 4 out of 10 solutions achieved precisions of 50% or greater and accuracies of 40% or greater.

Table 3.

Summary of the localization precision and accuracies at the unit, group, and facility levels.

Solution ID	Solution Type	Sensor Density (ha/sensor)	Localization (Equipment Unit)					
			Localization precision (%)			Localization Accuracy (%)		
			unit	group	facility	unit	group	facility
U	PSN	0.28	96.3	3.7	N/A	63.4	65.9	65.9
W	PSN	0.02	89.3	5.4	5.4	84.7	89.8	94.9
M	Imaging	0.12	84.9	10.8	4.3	80.6	90.8	94.9
V	Imaging	0.85	59.8	29.3	11.0	52.1	77.7	87.2
I	Imaging	0.85	56.0	34.4	9.6	43.2	69.8	77.2
N	PSN	0.02	55.7	41.2	1.1	55.5	95.2	96.3
F	PSN	0.09	50.4	48.4	1.3	43.3	84.9	86.0
S	PSN	0.07	29.2	49.9	20.9	23.7	64.1	81.1
P	PSN	0.06	25.5	65.0	9.5	23.5	83.3	92.0
D	PSN	0.11	24.6	64.2	11.2	22.6	81.8	92.2
R	PSN	0.08	23.2	33.0	43.8	15.3	37.1	66.0
T	PSN	0.12	22.9	35.0	42.2	13.6	34.4	59.5
C	PSN	0.07	22.0	54.5	23.5	16.7	58.2	76.0
Results from Ilonze et al. for the solutions that were retested in the current study								
D	PSN	0.11	27.2	68.8	4.0	25.3	89.3	93.1
F	PSN	0.08	40.8	53.9	5.3	36.5	84.7	89.4
N	PSN	0.05	51.6	41.8	6.6	42.2	76.2	81.6
P	PSN	0.14	27.0	56.9	16.1	23.4	72.8	86.8
Results from Bell et al. for the solutions that were retested in the current study								
C	PSN	0.14	1.0	0.5	98.5	0.8	1.2	79.3
D	PSN	0.11	0.0	52.8	47.2	0.0	47.3	89.6
F	PSN	0.11	24.8	50.2	25.0	19.2	58.1	77.5

The table is organized in decreasing order of the localization precision (for the current testing) at the equipment unit level. Precision at the unit level refers to a TP detection that identified the correct equipment unit, precision at the group level refers to a TP detection that identified the correct equipment group, and precision at the facility level refers to a TP detection that identified an emission from the test center.
Solution U localized emissions to the unit and group levels because its sensors monitored a specific equipment group, i.e., the tanks on pad 4.

Comparing the Current Study's Localization (Equipment Unit) to Bell et al. and Ilonze et al.

Generally, localization precision and accuracy for solutions tested in previous studies changed slightly relative to the current study (less than ±22%). The precision and accuracy increased for 2 of 4 solutions relative to Ilonze *et al.* (2024). For solutions C, D, and F, the precision and accuracy

increased at the unit level relative to Bell *et al.* (2023) (Table 3). Overall, there is an increase in precision and accuracy for the 2 of 4 and all three solutions relative to Ilonze *et al.* (2024) and Bell *et al.* (2023), respectively. This shows that, as a group, the current study's localization precision and accuracy partially improved compared to the two previous studies. Although the results may indicate partial improvement in precision and accuracy, these changes imply that the solutions' algorithms have improved.

Quantification Accuracy

Quantification accuracy for all TP detections from all solutions that reported the source emission rate is presented in two ways: as a percentage of single estimates within a factor of 3 and as the relative quantification error. In the current study, 11 solutions reported the source emission rate. The relative quantification error (in %) represents the relative difference between the reported and metered emission rates.

Overall, the percentage of single emission estimates within a factor of 3 for the 11 solutions ranged from 31 % to 92% (Table 4). The mean relative error from single estimates for controlled releases (CR) between 0.1 and 1 kg CH_4 h^{-1} and those greater than 1 kg CH_4 h^{-1} for 3 and 9 solutions were less than ±100%, respectively (Table 4). Small emission rates (between 0.1 and 1 kg CH_4 h^{-1}), often associated with component leaks, exhibit consistently high relative quantification error for most solutions (8). In contrast, larger emissions (> 1 kg CH_4 h^{-1}), often linked to process failures, show lower quantification error for most solutions (9 solutions); this indicates that most CM solutions could be well-positioned to mitigate larger emissions. Sections S8.4 and S10.3 of the SI provide additional information on quantification accuracies.

Table 4. The percentage of estimates within a factor of 3, and the relative quantification errors for the 11 solutions that reported the source emission rate.

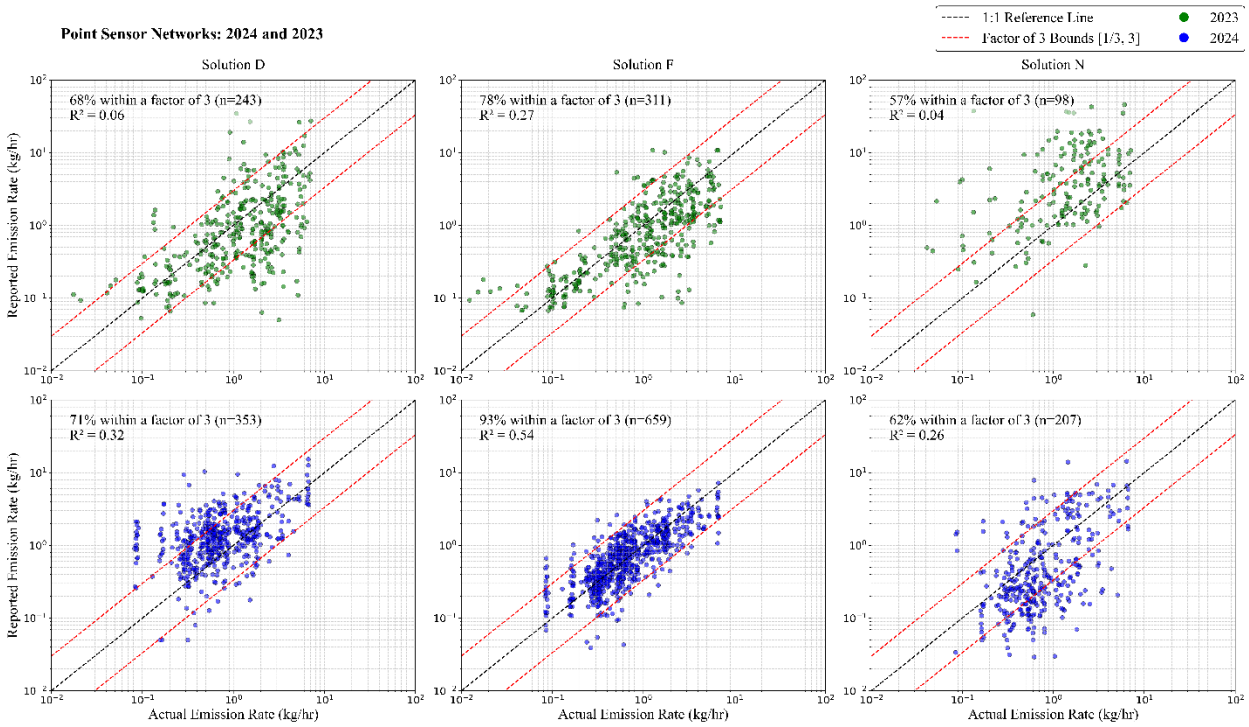
Solution	Solution type	Percent of Estimates within 3× (%)	Relative quantification error (%)					
			CR [0.1 to 1] (kg CH_4 h^{-1})			CR > 1 (kg CH_4 h^{-1})		
			Data Count	Mean Error [95% CI] (%)	Median Error (%)	Data Count	Mean Error [95% CI] (%)	Median Error (%)
F	PSN	93	510	35 [27, 43]	15	203	-12 [-18, -5]	-22
S	PSN	74	343	186 [153, 219]	90	168	27 [12, 41]	8
D	PSN	71	340	260 [213, 308]	124	158	33 [17, 49]	3
C	PSN	70	387	96 [63, 129]	-2	149	10 [-11, 30]	-40
M	Imaging	67	35	110 [56, 164]	98	57	-20 [-45, 6]	-52
N	PSN	62	282	33 [1, 66]	-41	135	3 [-20, 26]	-51
V	Imaging	57	17	367 [17, 716]	204	64	-29 [-46, -12]	-47
W	PSN	55	40	569 [319, 820]	394	16	4 [-34, 43]	-14
R	PSN	40	380	1326 [1003, 1648]	478	145	305 [205, 406]	111
T	PSN	39	233	962 [775, 1149]	383	73	3578 [-2832, 9988]	114
I	Imaging	31	54	422 [-279, 1122]	-61	71	-28 [-63, 8]	-79

The *percent of estimates 3× (%)* represents the percentage of single emission rate estimates within a factor of 3 [33% to 300%] of the CR rate. Also, the table shows the relative quantification errors' mean, median, and 95% CI of the mean for control release rates between 0.1 and 1 kg CH_4 h^{-1} and those > 1 kg CH_4 h^{-1}. The table is ranked based on the decreasing order of the percent estimates within a factor of 3.

Comparing the Quantification Accuracy for the Current Study to Ilonze et al.

Relative to Ilonze *et al.* (2024), 3 of 4 solutions reported the source emission rate (Figure 4). Between 57 and 78% of the single estimates were within a factor of 3 during testing by Ilonze *et al.* (2024). In the current study, between 62% and 93% of the single estimates were within a factor of 3 (Table 4 and Figure 4). The percentage of estimates within a factor of 3 has increased. Additionally, there is a stronger correlation between the reported and actual release rates.

Figure 4.



The reported against the actual emission rate for the three re-tested solutions. The plots with green points are from 2023 quantification data, while those with blue points are from the current study's quantification data. Each subplot shows the number of data points whose estimates are within a factor of 3. The black dotted line represents the 1:1 line; data points along this line illustrate that the reported equals the actual emission rate. The red dotted lines highlight the region where the single estimates are within a factor of 3 of the actual emission rates. The R^2 illustrates the correlation between the reported and actual emission rates. Generally, there is a stronger correlation between the reported and actual emission rates relative to Ilonze *et al.* (2024).

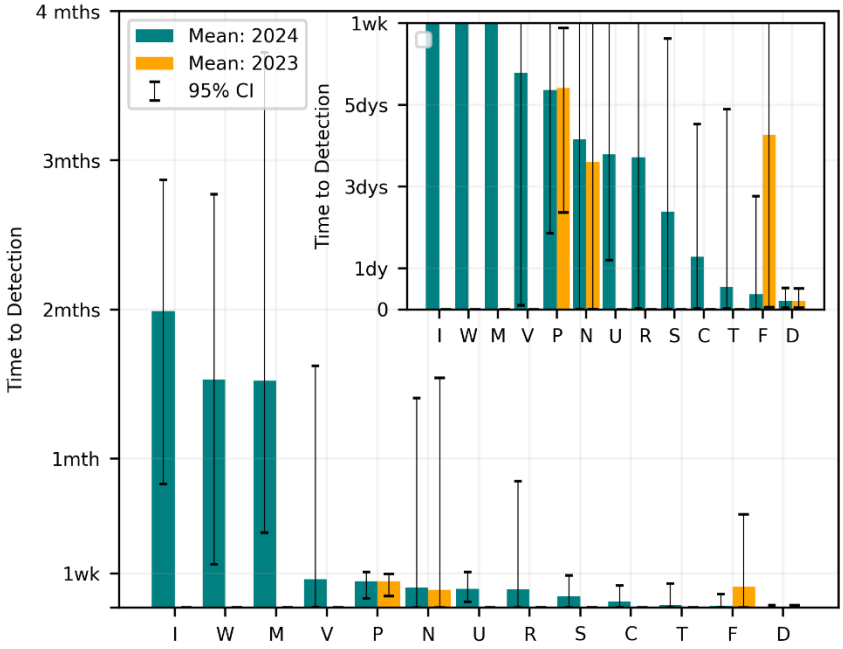
The overall emissions quantification has improved despite the consistently large quantification errors in single estimates. This improvement can be attributed to consistent testing, the improved emission quantification algorithm of solutions, and an increased number of deployed sensors for solution N. The improvement is illustrated by an increase in the percentage of single estimates within a factor of 3 in the current study, and Ilonze et al. (2024). Additionally, the solutions consistently tested in previous studies had the lowest relative quantification errors for small and large leaks (Table 4).

Detection Time

The detection time was calculated as the time between the start time of a controlled release and the time our email-based reporting system received the solution's detection report. In cases where we received multiple detection report IDs, the time for the first detection report was used for calculation. Generally, the mean detection time of solutions ranged from 5 hrs to 5 days (Figure 5). All solutions except solution D took longer (≥ 9 hrs) than the longest duration (8 hrs) of controlled releases in the study. Overall, detection time for these solutions varies widely due to the performers' manual evaluation of the detection report before submission. Additionally, the automated reporting system may have failed for solutions with long detection times, resulting in delayed detection report processing.

The mean detection time of solutions in the current study was larger than the mean detection time in Bell *et al.* (2023) and Ilonze *et al.* (2024). For the current study, the least mean detection time for 1 of 13 solutions was 5 hrs. For Bell *et al.* (2023) and Ilonze *et al.* (2024), 3 of 11 and 4 of 9 solutions had a mean detection time of ≤ 5 hrs, respectively. The mean detection time increased for 3 of the four solutions tested by Ilonze *et al.* (2024).

Figure 5.



Bar plot showing the mean time to detection for all the solutions. The plot shows the mean time to detection for solutions tested in the current study (green bars) and those tested in Ilonze et al. (gold bars); the bars are organized in descending order of the mean time to detection for solutions tested in 2024. The whiskers represent the 2.5th and 97.5th percentiles of each solution's detection time distribution. The inserted plot represents the outer plot's trimmed y-axis (at 1 week). For solutions I, W, and M, the mean time to detect is over 1 month.

Operation Downtime

Operation downtime is when a solution is not operational during testing, as a performer reports. The operation downtime for 12 of 13 solutions in this study was ≤ 6.4%, meaning that the 12 solutions met EPA's stipulated maximum operation downtime for continuous monitoring systems (US EPA, 2024). Although the EPA considers operational downtime on a 12-month rolling average, the solutions assessed in the current study demonstrate the ability to operate for more than three months without malfunctioning.

Implications

This study represents the third implementation of the ADED continuous monitoring protocol. The major findings for the current study are (1) solutions that tested in the two previous studies have improved or held steady in various performance metrics, including lower DL90, lower FP and FN rates, and higher localization accuracies at equipment group and facility levels, (2) scanning/imaging solutions offer higher localization precision and accuracy than most PSNs, and (3) solutions have higher quantification accuracies for larger emissions (> 1 kg CH_4 hr^{-1}) than smaller emissions (< 1 kg CH_4 hr^{-1}).

These results highlight the importance of controlled testing in understanding the potential performance of solutions in the field. Although the newly tested solutions performed well on localization precision and accuracy, these solutions struggle with the probability of detection, have large FP and false negative FN rates, and generally have long detection times. These results suggest that repeat testing, with concomitant attention to solution evaluation and improvement, positively impacts solution performance. Therefore, providing extended testing windows at low cost will likely benefit the development of next-generation LDAQ solutions.

Additionally, different solutions perform well on different metrics, indicating that they could or may already be tailored towards desirable functionalities or applicability to certain facility types or leak detection and repair programs. Some solutions have DL90 that qualify for current voluntary and regulatory leak detection programs. Also, these solutions have sufficiently low FP rates, meaning they should be well-positioned for deployment in those programs.

Finally, as in prior test programs, release rates did not span the range necessary for many solutions to achieve DL90 within the tested range. Comparing the three release programs to recently approved EPA regulations shows that there is a need for future programs to test above the

survey mode requirements listed in EPA's subpart OOOOb (US EPA, 2024) of 15 kg/h. Also, field testing indicates a need for future testing to include non-fugitive emission sources during controlled release testing (Day et al. 2024). Additionally, we have learnt that, with repeated testing, we needed data representative of field performance (Day et al. 2024).

To address these concerns, revisions to the controlled testing protocol are underway. Therefore, we propose the following revisions to the controlled testing protocol: 1. Including non-fugitive emissions to fugitive releases; 2. Including more test conditions, such as during summer, to widen the applicability of the test results; and 3. Encouraging performers to define detection thresholds relevant to their solutions. Non-fugitive emission sources are due to normal operational conditions at a facility. This creates non-steady emission rates and introduces temporal emission variation like the field's emission profile. This allows a representative assessment of a solution's performance.

Data accessibility statement

All data from different solutions that participated in the current testing are included as supplementary material.

Supporting Materials

All the solutions' reports are included in one zip folder. Additional information about the test facility and supplementary results are in PDF form. The final reports on each solution's performance can also be found on METEC's ADED Results website.

Acknowledgment

We recognize the effort and commitment of the solution vendors who paid to participate in the study and the test program. The authors acknowledge the contributions of the protocol development committee to the test protocol's development. Finally, the authors wish to thank Ryan Buenger and Wendy Hartzell for their assistance during testing.

Funding:

1. The Department of Energy (DOE: DE-FE0031873) supported the test protocol development and subsidized the cost of testing for participating solutions.
2. Matching funds from industry partners and associations for the DOE contract.

Competing Interest

After the testing was concluded and before manuscript submission, Ethan Emerson began working for Sonoma Technology, Inc., headquartered in Petaluma, California. Sonoma Technology, Inc. did not participate financially or otherwise in drafting this manuscript; therefore, the views and opinions expressed do not represent those of Sonoma Technology, Inc.

Author Contributions

Contributed to conception and design: FC, EE, CI, RD, DF, RB, DZ.

Contributed to data acquisition: FC, EE, DF, RB, DZ.

Contributed to analysis and interpretation of data: FC, EE, CI, EL, DZ.

Drafted and/or revised this article: FC, EE, CI, RD, EL, DF, RB, DZ.

Approved the submitted version for publication: FC, EE, CI, RD, EL, DF, RB, DZ.

References

"5-CCR-1001-9_eff-021424.Pdf." n.d. Google Docs. Accessed July 21, 2024. https://www.sos.state.co.us/CCR/GenerateRulePdf.do?ruleVersionId=11327&fileName=5%20CCR%201001-9.

Bell, Clay, Chiemezie Ilonze, Aidan Duggan, and Daniel Zimmerle. 2023. "Performance of Continuous Emission Monitoring Solutions under a Single-Blind Controlled Testing Protocol." *Environmental Science & Technology* 57 (14): 5794–5805. https://doi.org/10.1021/acs.est.2c09235.

Bell, Clay, Timothy Vaughn, and Daniel Zimmerle. 2020. "Evaluation of next Generation Emission Measurement Technologies under Repeatable Test Protocols." Edited by Detlev

Helmig and Brian Lamb. *Elementa: Science of the Anthropocene* 8 (January):32. https://doi.org/10.1525/elementa.426.

Bell, Clay, and Daniel Zimmerle. 2022. "METEC Controlled Test Protocol: Continuous Monitoring Emission Detection and Quantification." Testing,reports. Mountain Scholar. https://doi.org/10.25675/10217/235364.

Brandt, A. R., G. A. Heath, E. A. Kort, F. O'Sullivan, G. Pétron, S. M. Jordaan, P. Tans, et al. 2014. "Methane Leaks from North American Natural Gas Systems." *Science* 343 (6172): 733–35. https://doi.org/10.1126/science.1247045.

CDPHE. 2024. "Oil and Gas Greenhouse Gas Intensity Program | Department of Public Health & Environment." 2024. https://cdphe.colorado.gov/oil-and-gas-greenhouse-gas-intensity-program.

Chen, Zhenlin, Sahar H. El Abbadi, Evan D. Sherwin, Philippine M. Burdeau, Jeffrey S. Rutherford, Yuanlei Chen, Zhan Zhang, and Adam R. Brandt. 2024. "Comparing Continuous Methane Monitoring Technologies for High-Volume Emissions: A Single-Blind Controlled Release Study." *ACS ES&T Air*, June. https://doi.org/10.1021/acsestair.4c00015.

Collins, William, Raymond Orbach, Michelle Bailey, Sebastien Biraud, Ian Coddington, David DiCarlo, Jeff Peischl, Anuradha Radhakrishnan, and David Schimel. 2022. "Monitoring Methane Emissions from Oil and Gas Operations ‡." *Optics Express* 30 (14): 24326. https://doi.org/10.1364/OE.464421.

Day, Rachel Elizabeth, Ethan Emerson, Clay Bell, and Daniel Zimmerle. 2024. "Point Sensor Networks Struggle to Detect and Quantify Short Controlled Releases at Oil and Gas Sites." *Sensors* 24 (8): 2419. https://doi.org/10.3390/s24082419.

EIA. 2024. "U.S. Natural Gas Production Grew by 4% in 2023, Similar to 2022 - U.S. Energy Information Administration (EIA)." 2024. https://www.eia.gov/todayinenergy/detail.php?id=61646.

EPA. 2024. "40 CFR Part 60 Subpart OOOOb -- Standards of Performance for Crude Oil and Natural Gas Facilities for Which Construction, Modification or Reconstruction Commenced After December 6, 2022." 2024. https://www.ecfr.gov/current/title-40/part-60/subpart-OOOOb.

"Gd-56-Alternative-Test-Methods-in-the-Advanced-Methane-Detection-Program.Pdf." n.d. Accessed January 26, 2025. https://www.epa.gov/system/files/documents/2024-05/gd-56-alternative-test-methods-in-the-advanced-methane-detection-program.pdf.

Huang, Ling, Shannon Stokes, Qining Chen, and David T. Allen. 2024. "Uncertainties in the Estimated Methane Emissions in Oil and Gas Production Regions Based on Aircraft Mass Balance Flights." *ACS Sustainable Chemistry & Engineering* 12 (29): 11024–32. https://doi.org/10.1021/acssuschemeng.4c03945.

IEA, International Energy Agency. 2024. "Global Methane Tracker 2024 – Analysis." IEA. 2024. https://www.iea.org/reports/global-methane-tracker-2024.

Ilonze, Chiemezie, Ethan Emerson, Aidan Duggan, and Daniel Zimmerle. 2024a. "Assessing the Progress of the Performance of Continuous Monitoring Solutions under a Single-Blind Controlled Testing Protocol." *Environmental Science & Technology* 58 (25): 10941–55. https://doi.org/10.1021/acs.est.3c08511.

———. 2024b. "Assessing the Progress of the Performance of Continuous Monitoring Solutions under a Single-Blind Controlled Testing Protocol." *Environmental Science & Technology*, June, acs.est.3c08511. https://doi.org/10.1021/acs.est.3c08511.

IPCC. 2021. *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. 1st ed. Cambridge University Press. https://doi.org/10.1017/9781009157896.

Lavoie, Tegan N., Paul B. Shepson, Maria O. L. Cambaliza, Brian H. Stirm, Stephen Conley, Shobhit Mehrotra, Ian C. Faloona, and David Lyon. 2017. "Spatiotemporal Variability of Methane Emissions at Oil and Natural Gas Operations in the Eagle Ford Basin." *Environmental Science & Technology* 51 (14): 8001–9. https://doi.org/10.1021/acs.est.7b00814.

Liu, Yunsong, Jean-Daniel Paris, Gregoire Broquet, Violeta Bescós Roy, Tania Meixus Fernandez, Rasmus Andersen, Andrés Russu Berlanga, et al. 2024. "Assessment of Current Methane Emission Quantification Techniques for Natural Gas Midstream Applications." *Atmospheric Measurement Techniques* 17 (6): 1633–49. https://doi.org/10.5194/amt-17-1633-2024.

"Methane | Vital Signs." n.d. Climate Change: Vital Signs of the Planet. Accessed November 14, 2024. https://climate.nasa.gov/vital-signs/methane?intent=121.

MONITOR| Arpa-e. 2014. "MONITOR | Arpa-e.Energy.Gov." April 29, 2014. http://arpa-e.energy.gov/technologies/programs/monitor.

OGMP. 2024. "A Solution to the Methane Challenge – OGMP 2.0." 2024. https://www.ogmpartnership.org/partnership.

Parliament, European. n.d. "Reducing Methane Emissions in the Energy Sector | Legislative Train Schedule." European Parliament. Accessed April 24, 2024. https://www.europarl.europa.eu/legislative-train/package-fit-for-55/file-reducing-methane-emissions-in-the-energy-sector.

Ravikumar, Arvind P., Sindhu Sreedhara, Jingfan Wang, Jacob Englander, Daniel Roda-Stuart, Clay Bell, Daniel Zimmerle, et al. 2019. "Single-Blind Inter-Comparison of Methane Detection Technologies – Results from the Stanford/EDF Mobile Monitoring Challenge." Edited by Detlev Helmig and Brian Lamb. *Elementa: Science of the Anthropocene* 7 (January):37. https://doi.org/10.1525/elementa.373.

Schuit, Berend J., Joannes D. Maasakkers, Pieter Bijl, Gourav Mahapatra, Anne-Wil Van Den Berg, Sudhanshu Pandey, Alba Lorente, et al. 2023. "Automated Detection and Monitoring of Methane Super-Emitters Using Satellite Data." *Atmospheric Chemistry and Physics* 23 (16): 9071–98. https://doi.org/10.5194/acp-23-9071-2023.

Siebenaler, Shane P., Adam M. Janka, David Lyon, John P. Edlebeck, and Aileen E. Nowlan. 2016. "Methane Detectors Challenge: Low-Cost Continuous Emissions Monitoring." In *Volume 3: Operations, Monitoring and Maintenance; Materials and Joining*, V003T04A013. Calgary, Alberta, Canada: American Society of Mechanical Engineers. https://doi.org/10.1115/IPC2016-64670.

Titchener, James, Doug Millington-Smith, Chris Goldsack, George Harrison, Alexander Dunning, Xiao Ai, and Murray Reed. 2022. "Single Photon Lidar Gas Imagers for Practical and Widespread Continuous Methane Monitoring." *Applied Energy* 306 (January):118086. https://doi.org/10.1016/j.apenergy.2021.118086.

US EPA, OA. 2023. "Methane Emissions Reduction Program." Collections and Lists. January 20, 2023. https://www.epa.gov/inflation-reduction-act/methane-emissions-reduction-program.

———. 2024. "Biden-Harris Administration Announces Final Rule to Cut Methane Emissions, Strengthen and Update Greenhouse Gas Emissions Reporting for the Oil and Gas Sector."

News Release. May 6, 2024. https://www.epa.gov/newsreleases/biden-harris-administration-announces-final-rule-cut-methane-emissions-strengthen-and.

US EPA, OAR. 2016. "Understanding Global Warming Potentials." Overviews and Factsheets. January 12, 2016. https://www.epa.gov/ghgemissions/understanding-global-warming-potentials.

———. 2024. "Overview of Greenhouse Gases." Overviews and Factsheets. 2024. https://www.epa.gov/ghgemissions/overview-greenhouse-gases.

US EPA, ORD. 2017. "Greenhouse Gases." Reports and Assessments. November 2, 2017. https://www.epa.gov/report-environment/greenhouse-gases.

Vaughn, Timothy L., Clay S. Bell, Cody K. Pickering, Stefan Schwietzke, Garvin A. Heath, Gabrielle Pétron, Daniel J. Zimmerle, Russell C. Schnell, and Dag Nummedal. 2018. "Temporal Variability Largely Explains Top-down/Bottom-up Difference in Methane Emission Estimates from a Natural Gas Production Region." *Proceedings of the National Academy of Sciences* 115 (46): 11712–17. https://doi.org/10.1073/pnas.1805687115.

Zavala-Araiza, Daniel, David Lyon, Ramón A. Alvarez, Virginia Palacios, Robert Harriss, Xin Lan, Robert Talbot, and Steven P. Hamburg. 2015. "Toward a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites." *Environmental Science & Technology* 49 (13): 8167–74. https://doi.org/10.1021/acs.est.5b00133.

Zimmerle, Daniel, Sonu Dileep, and Casey Quinn. 2024. "Unaddressed Uncertainties When Scaling Regional Aircraft Emission Surveys to Basin Emission Estimates." *Environmental Science & Technology* 58 (15): 6575–85. https://doi.org/10.1021/acs.est.3c08972.