# Continuous Monitoring Final Report

Performer: Solution S
Report Generated: September 6, 2024

# 1 Experiment Summary

## 1.1 Experiment Demographics

Testing included a total of 351 experiments performed between 2024/02/06 00:00 and 2024/04/29 00:00. A total of 783 controlled releases were performed during these experiments. Individual experiments included between 1 and 5 controlled releases, with an average of 2.23 controlled releases per experiment (Figure 1). Figure 2 shows the emission rate of controlled releases during experiments. Figure 3 shows a histogram of the duration of controlled releases performed. Emission rates ranged from 2.3 to 189.0 slpm whole gas. The variation of controlled release rates (kg $CH_4$/h) with experiment duration (h) is captured in Figure 4. Each controlled release is classified as either a True positive (TP) or False negative (FN). Figure 5 summarizes the location of controlled releases by equipment unit. The range of controlled release rates (kg $CH_4$/h) and detection classifications across the equipment groups involved in the experiments is illustrated in Figure 6.

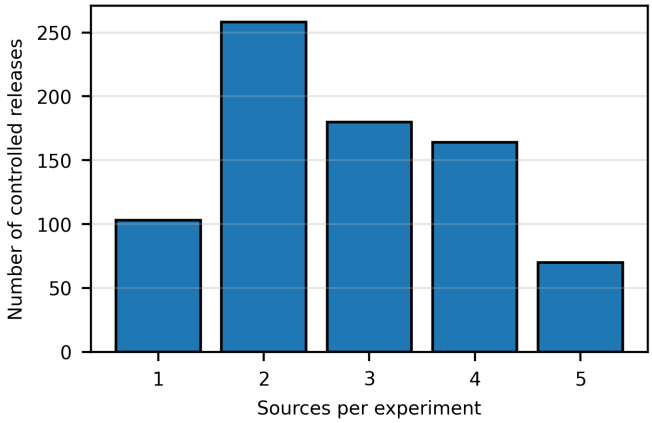


Figure 1: Histogram of number of controlled releases included in each experiment

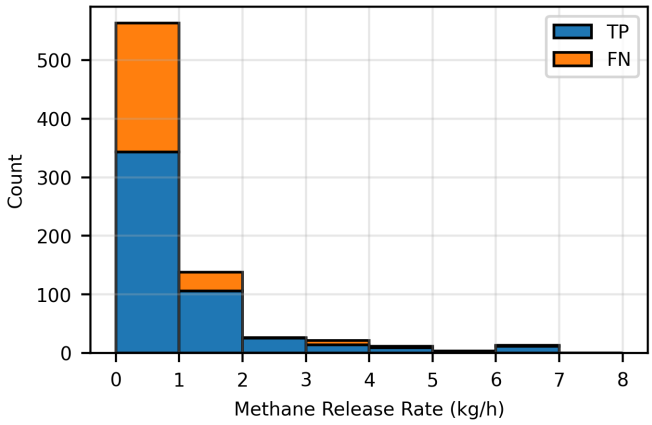


Figure 2: Histogram of metered emission rate (kg $CH_4$/h) for all controlled releases performed during experiments. Each controlled release is classified as true positive (TP) or false negative (FN).

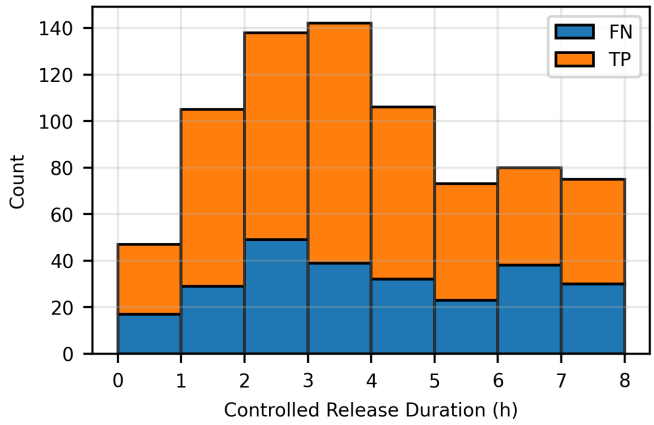


Figure 3: Histogram of controlled release duration (h) for all controlled releases performed during experiments. Each controlled release is classified as true positive (TP) or false negative (FN).

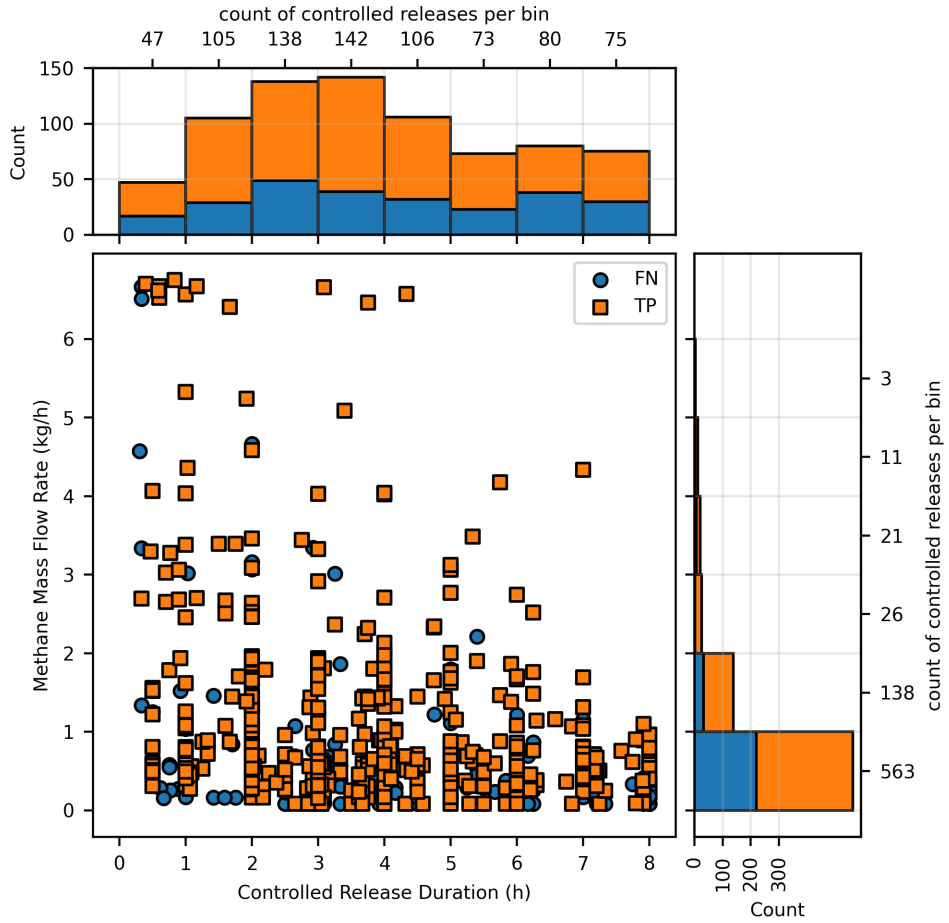


Figure 4: Controlled release rate (kg $CH_4$/h) versus controlled release duration (h). Markers indicate True Positive (TP) and False Negative (FN) detection classification. Data are summarized by histogram shown parallel to the respective axes.

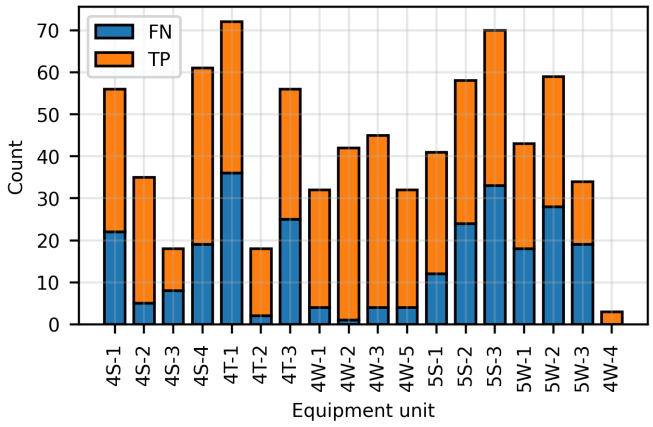


Figure 5: Distribution of controlled releases across equipment units. Each controlled release is classified as true positive (TP) or false negative (FN).

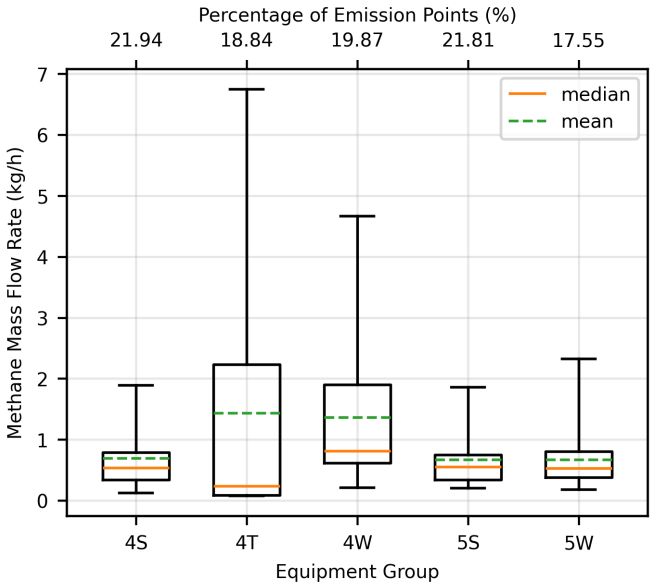


Figure 6: Controlled release rate (kg $CH_4$/h) versus equipment group. The whiskers of each boxplot represent the range of all controlled releases metered in each equipment group. While the boxes represents 50% of data points about the median controlled release rate

## 1.2 Meteorology

Meteorological data during testing is summarized in the following figures. Figure 7 shows the average ambient temperature during each experiment. Temperature ranged from -11.0 $^o$C to 23.4 $^o$C, with a mean of 5.71 $^o$C. Figure 8 shows the average wind speed of each experiment. Average wind speed ranged from 0.516 m/s to 12.9 m/s, with a mean of 2.9 m/s. Figure 9 shows the average direction during each experiment. Figure 10 shows the average wind speed and direction of each experiment as a wind rose.

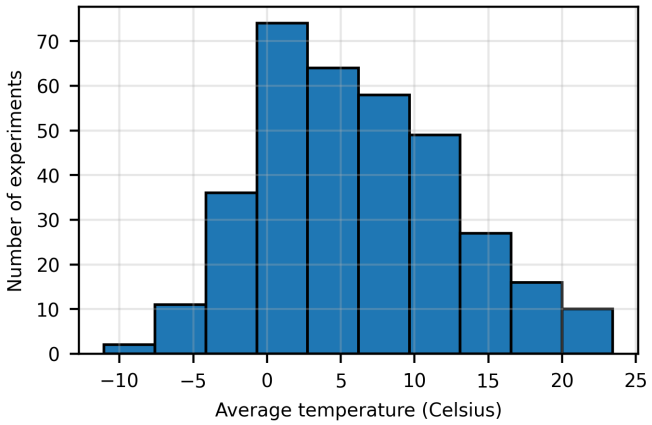


Figure 7: Histogram of average temperature ($^o$C) during each experiment

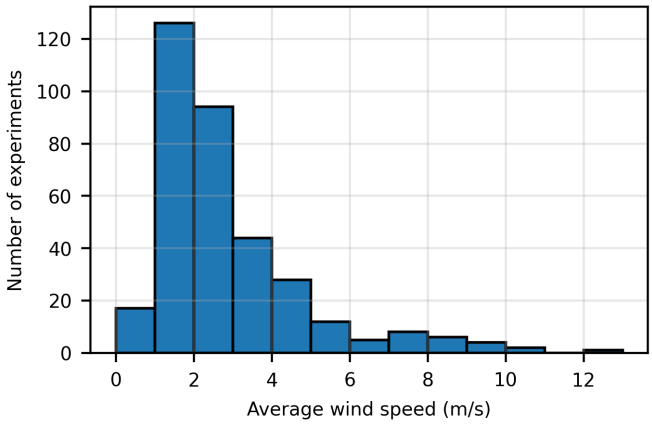


Figure 8: Histogram of average wind speed (m/s) during each experiment

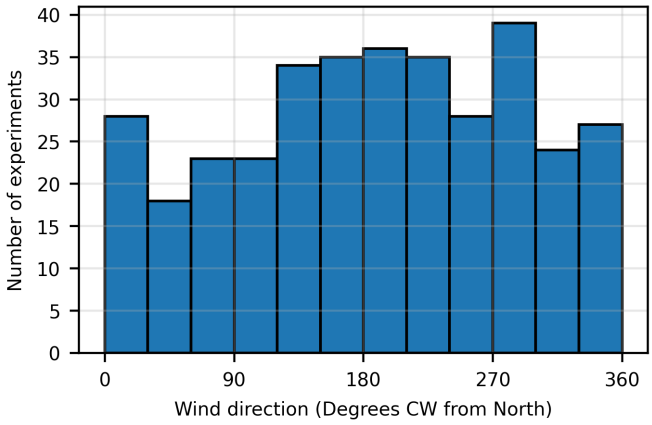


Figure 9: Histogram of average wind direction during each experiment

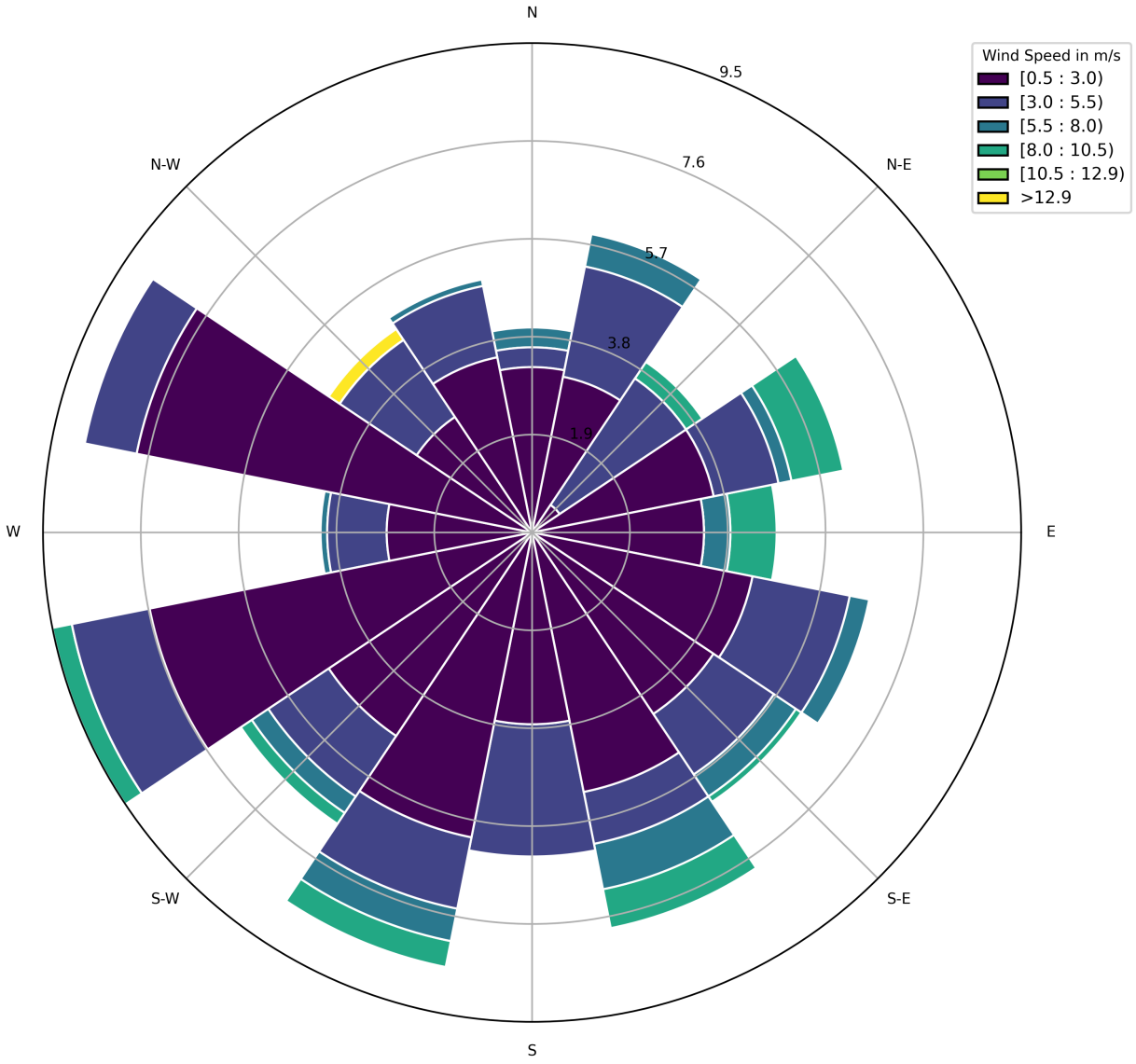


Figure 10: Experiment wind rose. Wedges show percent of experiments by average wind speed and direction. $0^o$ represents the true North of the test center.

# 2 Performance Metrics

Metrics as described in section 6 in the protocol are reported in this section. Primary metrics are reported for all performers. Secondary metrics are reported only if the performer detection reports included the required data for their calculation.

## 2.1 Primary Metrics

### 2.1.1 Classification of Detections

Solution S reported a total of 1936 detection reports including 735 unique EmissionSourceIDs. 105 reported sources were excluded from classification as they may indicate an emission from the test center during a maintenance period or another test conducted at the test center that was not following the continuous monitoring protocol. Remaining detection reports were matched to controlled releases following the method outlined in the test protocol to identify true positives, false positives, and false negatives. The classification of controlled releases and detections is shown in Table 1. The classification of individual detections and controlled releases can be found in the accompanying data files.

Table 1: Classification of controlled releases and detection reports.

| Level | True Positive | False Negative | False Positive | Excluded from Analysis | Total |
|---|---|---|---|---|---|
| Controlled Releases | 511 | 264 | - | 8 | 783 |
| Detection Reports | 511 | - | 119 | 105 | 735 |

### 2.1.2 Probability of Detection

The probability of detection (POD) curves derived from the classified detection and controlled release records is illustrated in Figure 11. The upper subplot of the figure shows a logistic regression performed against the true positive and false negative results. The lower subplot shows the POD calculated using a variable bin width each containing approximately equal number of data points. The bins are obtained through quantile-based discretization limiting number of points per bin to between 30 to 50. Power function with intercept set to zero is used to fit calculated POD data points. In this approach, the statistical significance of each bin, is approximately equal. Figure 12 shows the probability of detection curves for controlled release rates normalized by wind speed.

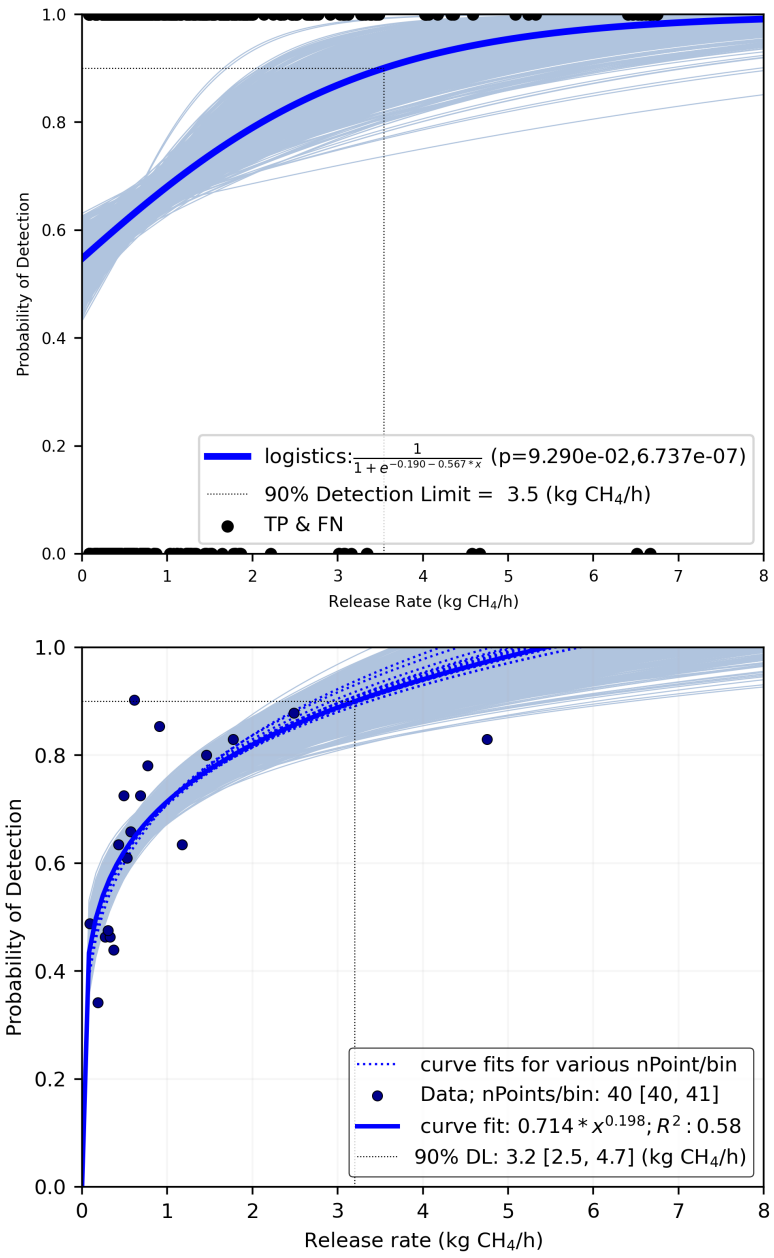


Figure 11: Probability of detection versus emission rate. Upper subplot is the probability of detection vs emission rate (kg $CH_4$/h) assessed with logistic regression including all true positives (equipment unit-, equipment group-, and facility-levels). True positive and false negative detections are shown with markers at y = 1 and y = 0 respectively. The regression is performed on bootstrapped results to show a cloud of curves to illustrate uncertainty in the result. The lower subplot shows the probability of detection versus emission rate (kg $CH_4$/h) fitted using power functions. The x-axis is divided into equal-sized bins with each marker (pod) as the fraction of controlled releases in a bin classified as TPs. The emission rate at which the POD reaches 90% is shown. Each pod data point is bootstrapped to produce a cloud of curves illustrating associated uncertainty. When the bootstrapping could not evaluate the lower and upper empirical confidence limit on the emission rate of 90% POD, they are given as 0 and NA respectively. Curve fits (dotted colored lines) obtained using other (5-16) quantile-based discretizations are shown for comparison. The mean count of points per bin along with the minimum and maximum counts across all bins is also shown in the figure.

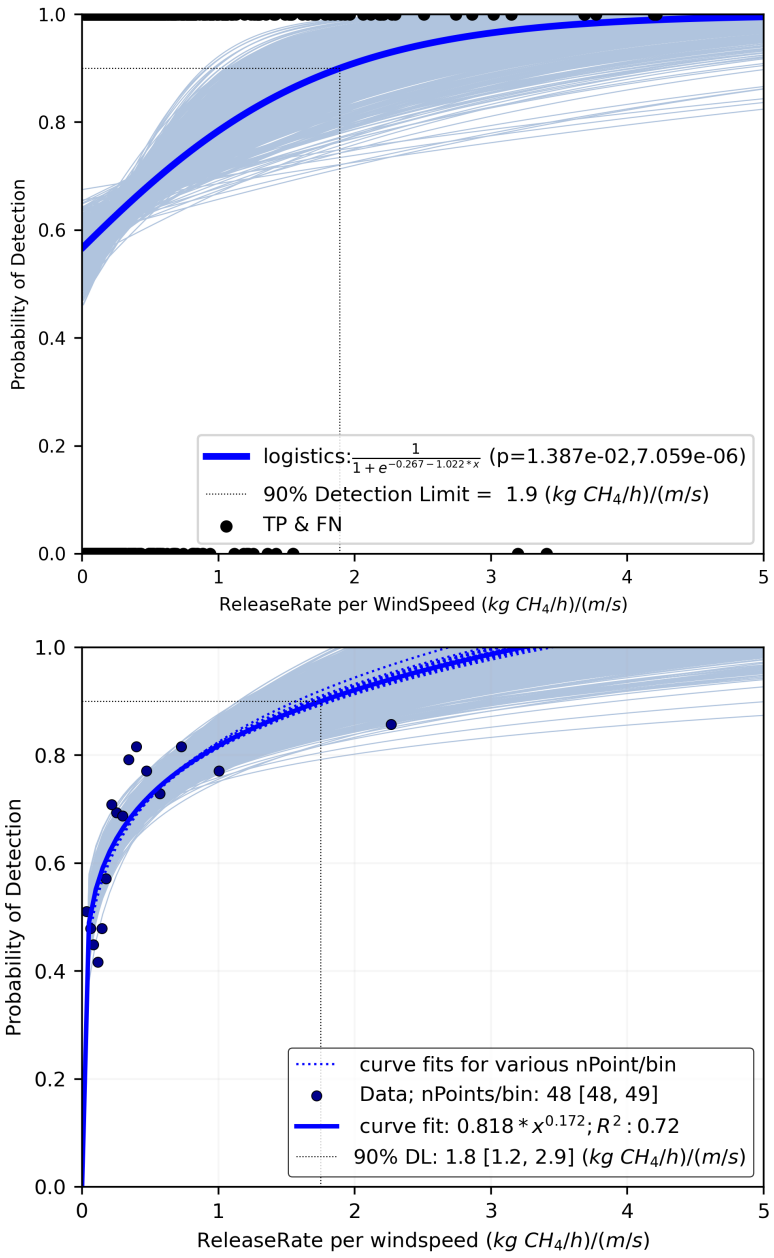


Figure 12: Probability of detection versus wind normalized emission rate. Upper subplot is the probability of detection vs wind normalized emission rate $(kg\ CH_4/h)/(m/s)$ assessed with logistic regression including all true positives (equipment unit-, equipment group-, and facility-levels). True positive and false negative detections are shown with markers at y = 1 and y = 0 respectively. The lower subplot shows the probability of detection versus wind normalized emission rate (kg $CH_4$/h) fitted using power functions. The x-axis is divided into equal-sized bins with each marker (pod) as the fraction of controlled releases in a bin classified as TPs. The wind normalized emission rate at which the POD reaches 90% is shown. Each pod data point is bootstrapped to produce a cloud of curves illustrating associated uncertainty. When the bootstrapping could not evaluate the lower and upper empirical confidence limit on the wind normalized emission rate of 90% POD, they are given as 0 and NA respectively. Curve fits (dotted colored lines) obtained using other (5-16) quantile-based discretizations are shown for comparison. The mean count of points per bin along with the minimum and maximum counts across all bins is also shown in the figure.

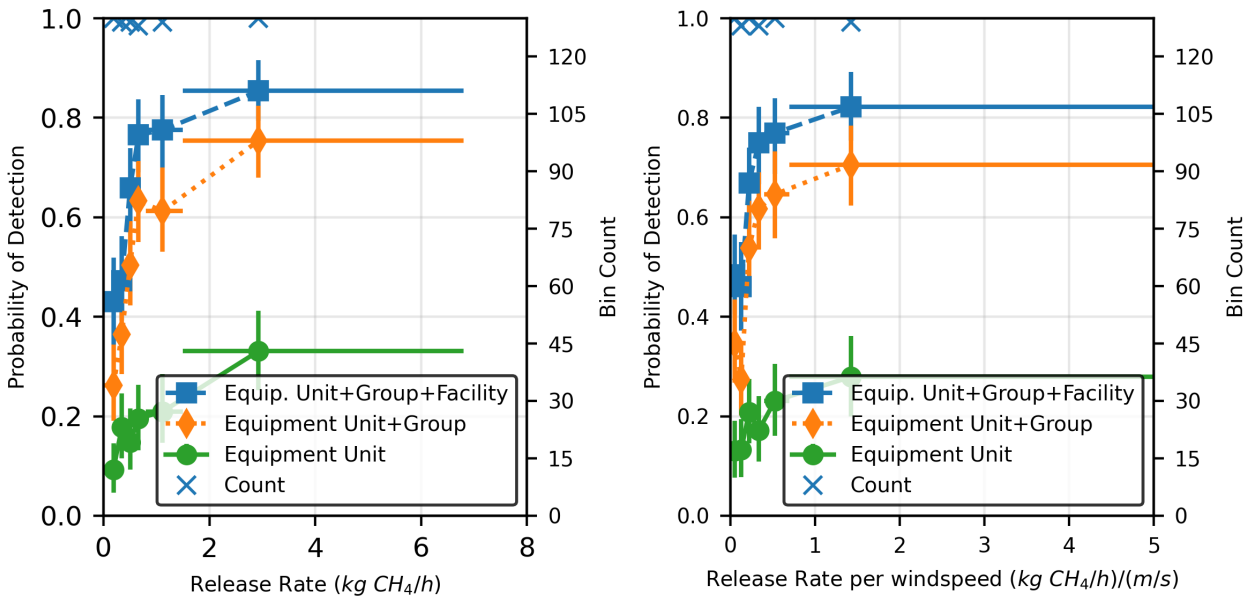


Figure 13: The left subplot shows the probability of detection vs emission rate (kg $CH_4$/h) calculated by binning data along the x-axis. Separate curves are illustrated for equipment unit, equipment group, and facility levels. Markers represent mean emission rate and observed probability of detection within each bin. X whiskers indicate maximum and minimum emission rate in each bin. Y whiskers indicate maximum and minimum probability of detection when empirical data is bootstrapped. The number of data points within each bin is plotted on right hand axis. The right subplot shows the probability of detection vs wind normalized emission rate ($kg\ CH_4/h)/(m/s$) calculated by binning data along the x-axis. Separate curves are illustrated for equipment unit, equipment group, and facility levels. Markers represent mean emission rate and observed probability of detection within each bin. X whiskers indicate maximum and minimum wind normalized emission rate in each bin. Y whiskers indicate maximum and minimum probability of detection when empirical data is bootstrapped. The number of data points within each bin is plotted on right hand axis.

### 2.1.3 False Positive Fraction

The false positive fraction ($N_{FP}/(N_{FP} + N_{TP})$) derived from all detections was 0.189.

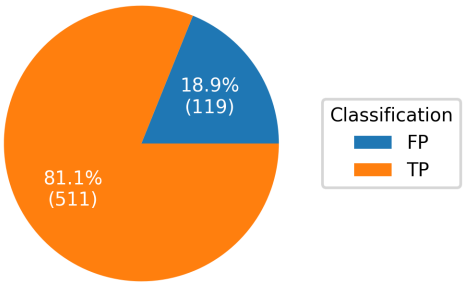


Figure 14: Classification of performer detection alerts. True Positive (TP) alerts were paired with a controlled release at the test center. False Positive (FP) alerts were unpaired.

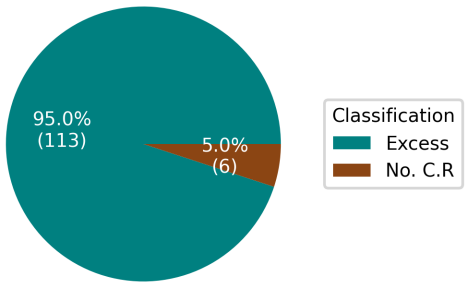


Figure 15: Categorization of false positives. The pie chart shows the fraction of FPs due to no controlled release and excess detection alerts sent during experiments.

#### 2.1.4 False Negative Fraction

The false negative fraction ($N_{FN}/(N_{FN} + N_{TP})$) derived from all controlled releases was 0.341.

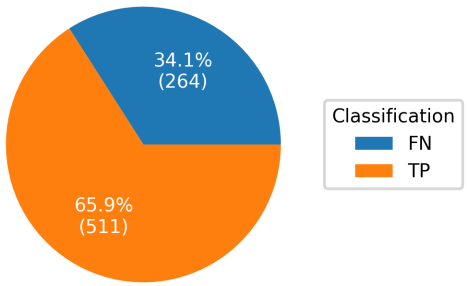


Figure 16: Classification of controlled releases. True Positive (TP) releases were paired with a detection reported by the performer. False Negative (FN) releases were unpaired.

#### 2.1.5 Detection Time

The detection time of individual detections is illustrated in figure 17. The detection time ranged from 0.071 to 166.166 hours. The 2.5% and 97.5% percentiles of the detection time were 0.278 and 158.802 hours respectively. The mean detection time was 57.495 hours. Note, the detection time may only be calculated for true positive detections.

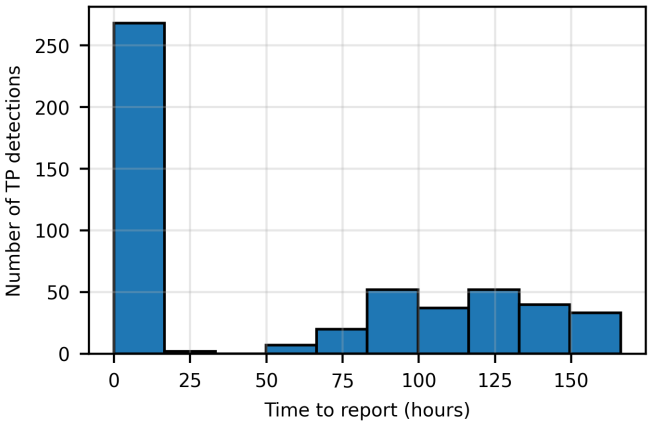


Figure 17: Detection time was calculated as the time difference between the start time of a controlled release and the time when the matched detection was received by the test center.

## 2.2 Secondary Metrics

### 2.2.1 Localization Precision (Equipment Unit)

Table 2 lists the number of true positives at the equipment unit, equipment group, and facility-level.

Table 2: Localization Precision (Equipment Unit)

| Level | True Positive Count |
|---|---|
| Equipment Unit | 149 |
| Equipment Group | 255 |
| Facility | 107 |

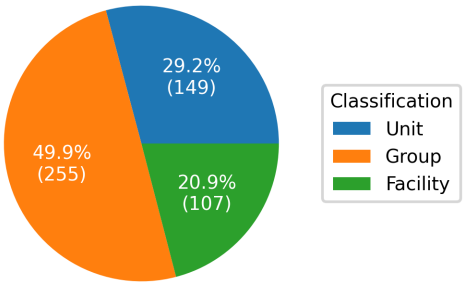


Figure 18: Categorizing True Positive detections by localization precision levels. The localization precision levels include the Equipment unit level, Equipment group level, and the Facility level.

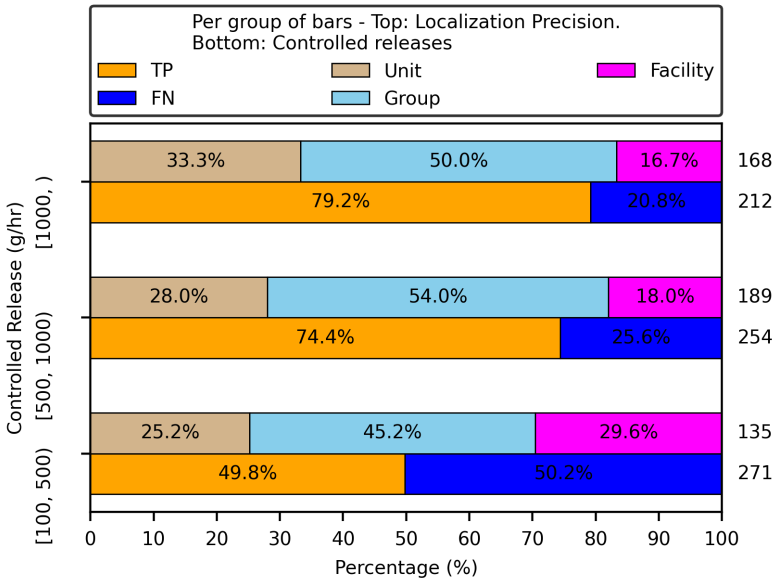


Figure 19: Categorizing the classification of controlled releases and equipment localization precision based on the range of release rate. Controlled releases are classified as True Positives (TP) and False Negatives (FN). The localization precision levels include the Equipment unit level, Equipment group level, and the Facility level.

### 2.2.2 Localization Accuracy (Equipment Unit)

Table 3 lists the localization accuracy (the fraction of reports identified as true positive) at each level of precision. Precision levels are the equipment unit, the equipment group or better (i.e. equipment unit + equipment group), and facility or better (i.e. equipment unit + equipment group + facility). The fraction of reports identified as false positive is also included in the table.

Table 3: Localization Accuracy (Equipment Unit)

| Level | Localization Accuracy |
|---|---|
| Equipment Unit | 0.237 |
| Equipment Group | 0.641 |
| Facility | 0.811 |
| False Positive Fraction | 0.189 |

### 2.2.3 Operational Factor

The operational factor (1 - $\sum t_{offline}$ / $t_{total}$) derived from all offline reports received during testing was 1.0.

### 2.2.4 Quantification Accuracy (Absolute)

The quantification accuracy (absolute) of individual detections is illustrated in figure 20. The upper panel shows the reported emission rate versus metered emission rate for true positive detections. Markers are colored by controlled release duration. The center panel shows the absolute error in emission rate estimates provided by the performer, relative to the metered emission rate. The lower panel shows a box plot summarizing the quantification error for data within each order of magnitude of release rate. The correlation line is derived using a linear regression model yielding a slope of 0.956 and a coefficient of determination ($R^2$) of -0.05. The regression bias for this model is -4.4%. The quantification accuracy ranged from -6000.0 g/h to

7220.0 g/h. The mean quantification accuracy (absolute) was 491.0 g/h. 375 quantification estimates were higher the metered flow rate, and 136 quantification estimates were lower than the metered flow rate.

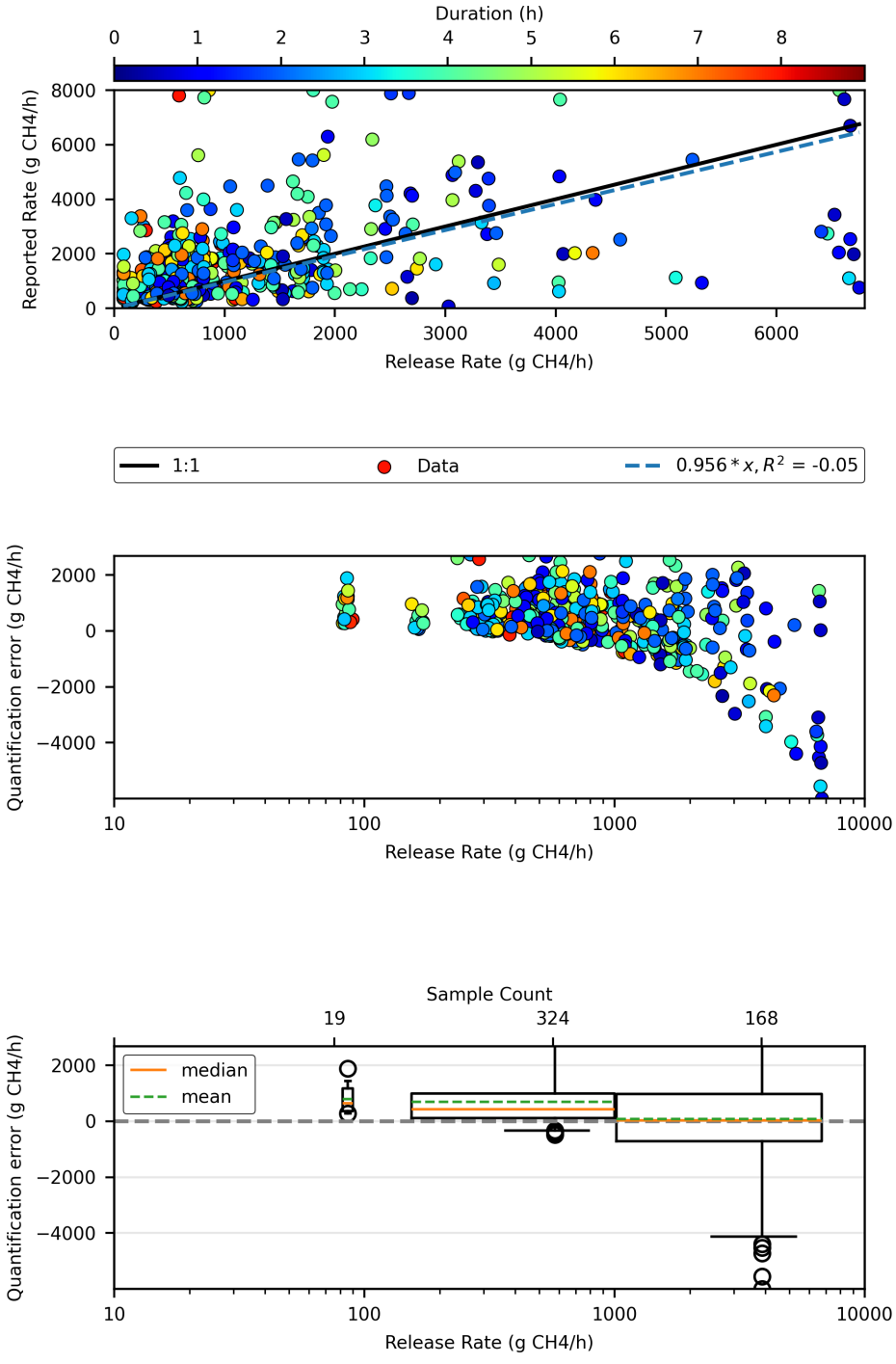


Figure 20: Quantification accuracy (absolute) across experiments

### 2.2.5 Quantification Accuracy (Relative)

The quantification accuracy (relative) of individual detections is illustrated in figure 21. The upper panel shows the reported emission rate versus metered emission rate for true positive detections. Markers are colored by controlled release duration. The center panel shows the relative error in emission rate estimates provided by the performer, normalized by the metered emission rate. The lower panel shows a box plot summarizing the quantification error for data within each order of magnitude of release rate. The quantification accuracy (relative) ranged from -0.978 to 22.1 times the metered flow rate. The mean quantification accuracy (relative) was 1.34 times the metered flow rate. Figure 22 suggests that the 54%, 74%, and 91% of measurements underestimated/overestimated the true rates by factors within 2, 3, and 5 respectively. The correlation line is derived using a linear regression model yielding a slope of 0.956 and a coefficient of determination ($R^2$) of -0.05. The regression bias for this model is -4.4%. The quantification accuracy (relative) ranged from -0.978 to 22.1 times the metered flow rate. The mean quantification accuracy (relative) was 1.34 times the metered flow rate. The y-axis upper limit is cut off at 0.95 quantile.
Quantification Accuracy Relative Hist

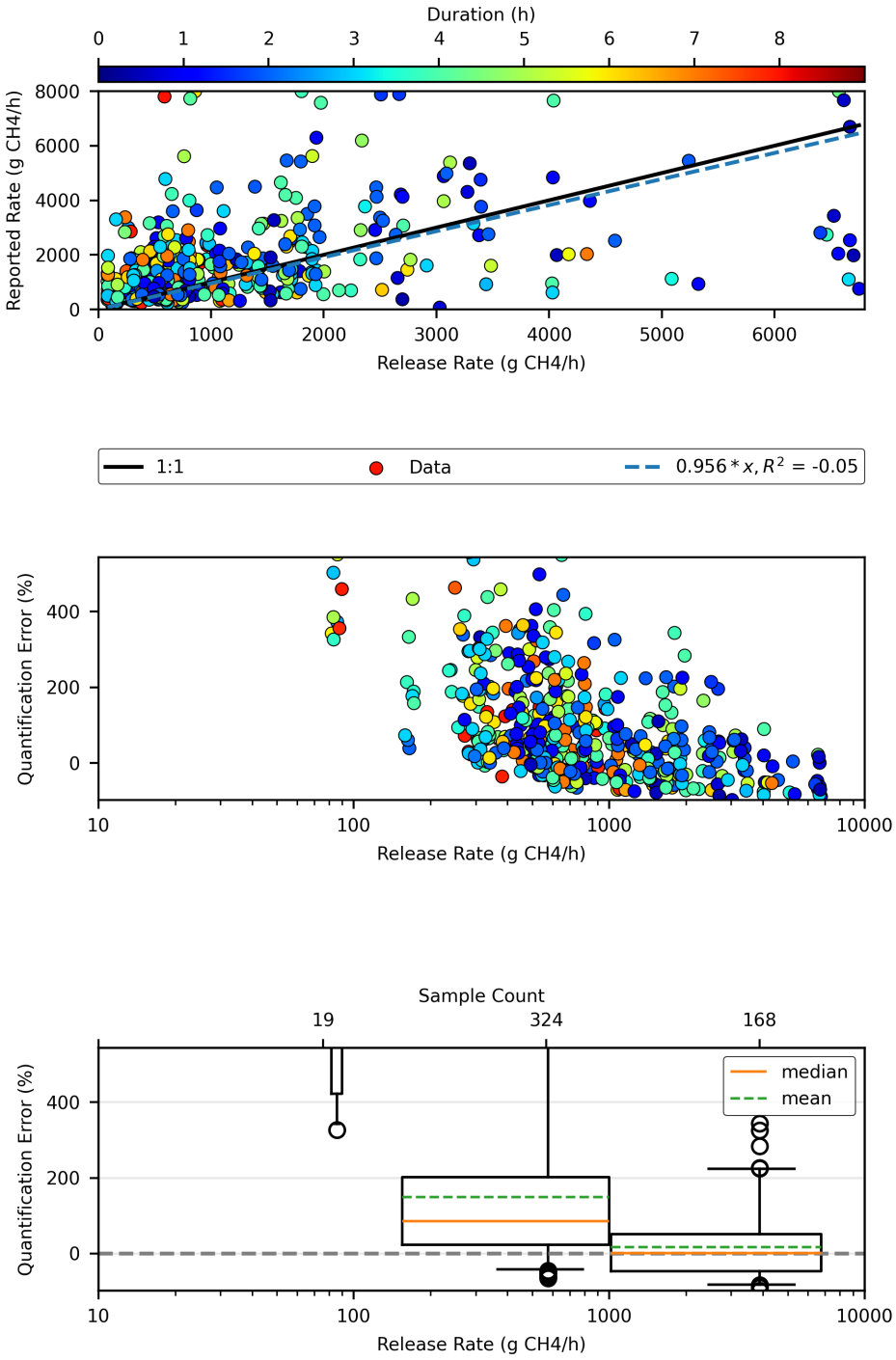


Figure 21: Quantification accuracy (relative) across experiments

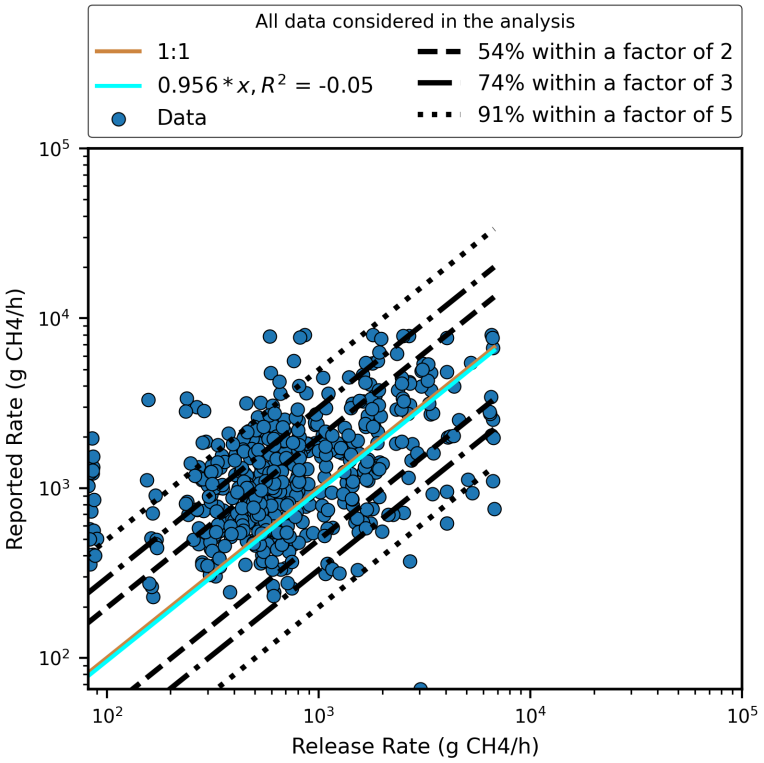


Figure 22: Quantification accuracy (relative) across experiments: The figure shows the relationship between reported rates and the controlled release rates. Regression and one to one lines to illustrate the correlation between reported and controlled release rates. The percentage of measurements within factors of quantification of 2, 3, and 5 are also shown.

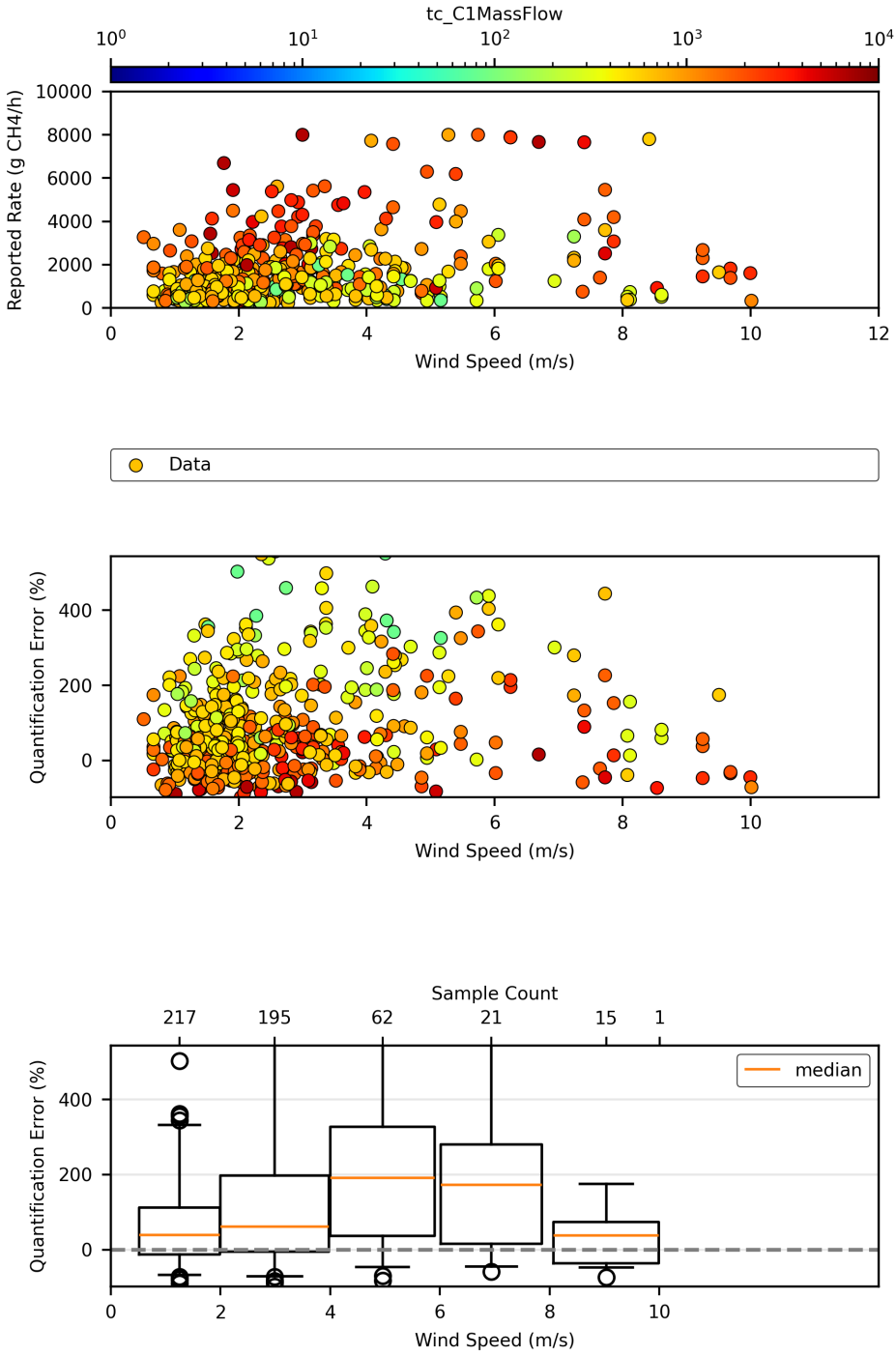


Figure 23: Quantification accuracy (relative) across experiments: The top panel shows the relationship between reported rates and the test center's mean wind speed. Markers are colored by controlled release rates (g $CH_4$/h). The center panel shows how the relative error in emission rate estimates varies mean windspeed. The lower panel shows a box plot summarizing the quantification error for data within arbitrary bin size of the test center's mean wind speed. The y-axis upper limits of each of the panel is trimmed at 0.95 quantile of the whole data.

### 2.2.6 Quantification Precision (Absolute)

Solution S did not report data required to compute this metric.

### 2.2.7 Quantification Precision (Relative)

Solution S did not report data required to compute this metric.

### 2.2.8 Localization Accuracy (Single Coordinate)

The localization accuracy (single coordinate) of individual detections is illustrated in figure 24. The localization accuracy (single coordinate) ranged from 0.802 to 92.0 meters. The mean localization accuracy (single coordinate) was 15.4 meters.

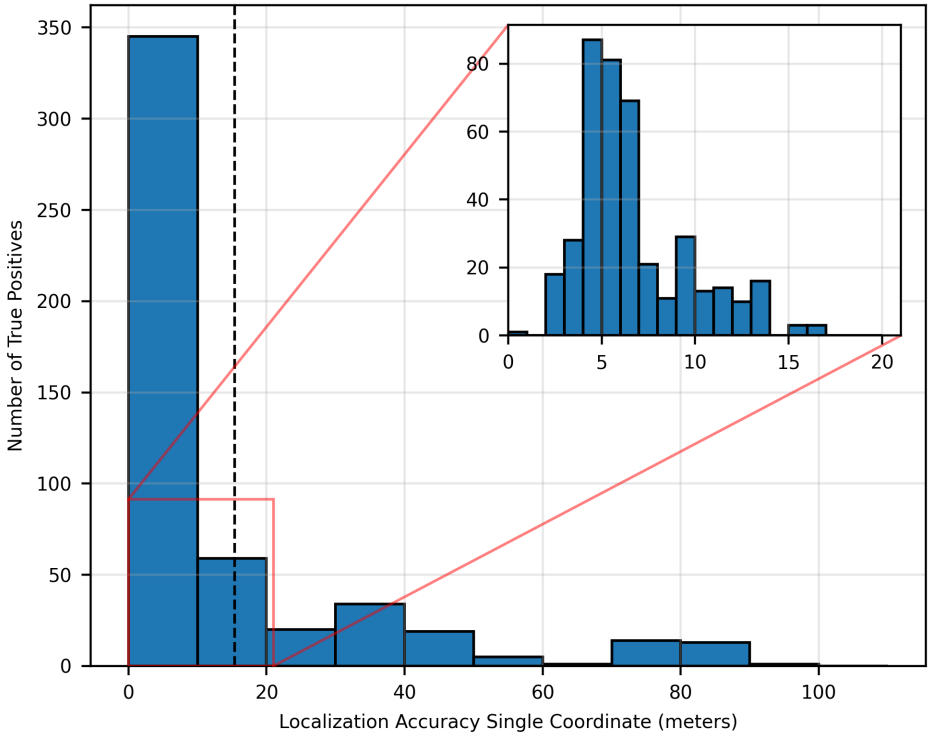


Figure 24: Localization accuracy (single coordinate) across experiments. The main histogram shows the localization accuracy of a true positive (TP) detections calculated as the distance from reported coordinates to the controlled release sources. The insert shows 1 meters resolution for distances between 0 meters to 20 meters.

### 2.2.9 Localization Accuracy (Bounding Box)

The localization accuracy (bounding box) of individual detections is illustrated in figure 25. The localization accuracy (bounding box) ranged from 0.275 to 91.5 meters. The mean localization accuracy (bounding box) was 13.4 meters.

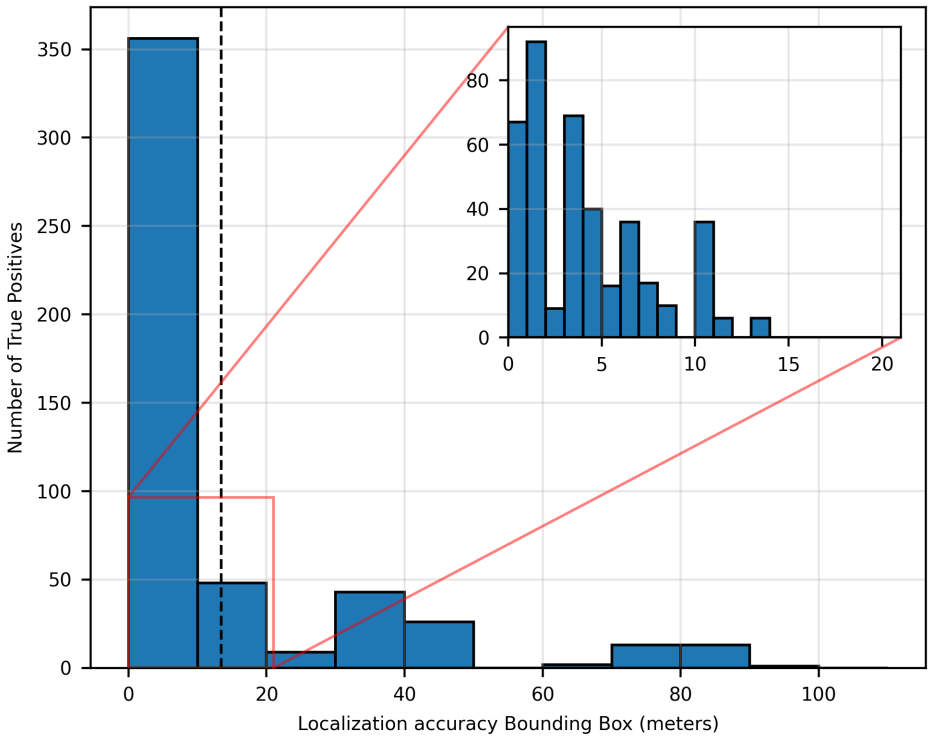


Figure 25: Localization accuracy (bounding box) across experiments. The main histogram shows the localization accuracy of a true positive (TP) detections calculated as the distance from the center of the bounding box of the controlled release source. The insert shows 1 meters resolution for distances between 0 meters to 20 meters.

### 2.2.10 Bounding Box Accuracy

The number of emissions points within the reported bounding box was 197 (38.6% of true positive detections). The number of emissions points outside the reported bounding box was 314 (61.4% of true positive detections)

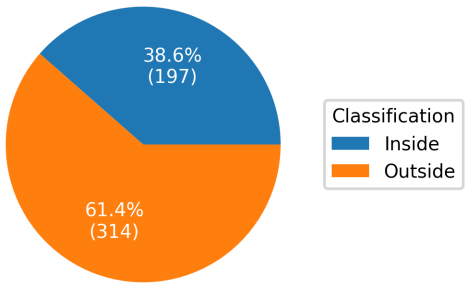


Figure 26: Classification of True Positive (TP) detections based on leak source location with respect to bounding box reported by performer. Each true positive detection is either inside or outside of the reported bounding box.

### 2.2.11 Localization Precision (Bounding Box)

The localization precision (bounding box) of individual detections is illustrated in figure 27. The localization precision (bounding box) ranged from 24.2 to 69.7 $m^2$. The mean localization precision (bounding box) was 41.1 $m^2$.

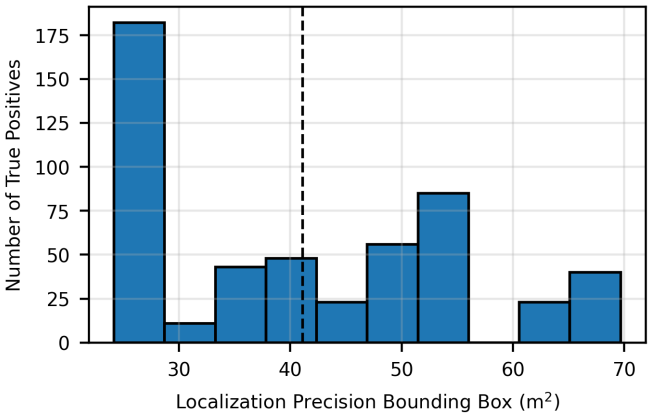


Figure 27: Localization precision (bounding box) across experiments

# 3 Documentation of Test Protocol

A copy of the test protocol is provided in Continuous Monitoring Protocol R1.0.pdf in the zip folder with this report.

# 4 Documentation of System Under Test

This solution deployed 12 sensors at a sensor density of 704.0 $m^2$ per sensor. The documentation of system under test has been redacted to preserve confidentiality of the performers in this publication.

# 5 Controlled Release and Detection Data

Controlled release and classification data are provided in classifiedReports.csv in the zip folder with this report. The raw data reported by the performer is provided in detectionReports.xlsx in the zip folder with this report.

# 6 Offline Reports

Offline reports are provided in offlineReports.csv in the zip folder with this report.

# 7 Flow Meter Calibrations

Meter calibrations are available from the test center by request.