Daniel Zimmerle, 970-581-9945, dan.zimmerle@colostate.edu
Clay Bell, clay.bell@colostate.edu

ENERGY INSTITUTE
COLORADO STATE UNIVERSITY

METEC Controlled Test Protocol:
Survey Emission Detection And Quantification

Revision 1.0

April 26, 2022

1 Purpose:

This testing will assess the performance of survey methods which perform leak detection and quantification (LDAQ) under single-blind controlled release testing over a range of environmental conditions and emission rates. Testing will evaluate system-level performance measures including Probability of Detection and Detection Time. Additional metrics including accuracy and precision of localization and quantification estimates will be evaluated if applicable.

2 Definitions

- Component – An individual part of a larger Equipment Unit. For example, a connector, flange, fitting, valve, pneumatic actuator, pneumatic controller, hatch, pressure relief valve or pressure vessel on an Equipment Unit.

- Controlled Release (CR) – A type of experiment where emissions are intentionally created for the purpose of evaluating emission detection and/or quantification systems. During a Controlled Release, the emission rate and location are known to the Test Center within well understood accuracy.

- Detection – An alert provided by an Emission Detection System to the Facility operator that an Emission is present. An elevated gas concentration measurement alone does not constitute a Detection, but instead must be accompanied by analytics to attribute the elevated concentration to an Emission within the Facility. This attribution must be established with a high enough confidence to warrant providing a detection alert to the Facility operator.

- Emission – a release of gas from a system to the ambient environment.

- Emission Detection System – A system including the sensor(s), deployment platform, auxiliary equipment, and analytics capable of detecting emissions and attributing them, at minimum, to a facility. Emission Detection Systems may include analytics to estimate emission rate and/or the location of the emission source.

- Emission Survey – An inspection performed at a facility using a handheld or otherwise mobile Emission Detection System. Emission surveys must localize emission sources detected during the survey. Emission surveys may include emission quantification estimates at the emission source- and/or facility-level.

- Equipment Group – A set of Equipment Units in proximity of one another.
- Equipment Unit – An individual unit of equipment such as a wellhead, separator, or liquid storage tank.
- Experimental Design (a test matrix) – A set of Experimental Design Points defined to investigate correlation between variation in a dependent variable and variation of one or more independent variables.
- Experimental Design Point (an experiment) – A single combination of settings for the independent variables of a controlled release experiment. Independent variables include both the emission rate of the Controlled Release(s) and environmental conditions.
- Facility – A set of Equipment Units and/or Equipment Groups with a common purpose and defined boundary which may be physical (such as a fenceline) or implied.
- False Negative (FN) – A Controlled Release which was not detected by a Performer. See Section 6.1 for Classification of Detections.
- False Negative Fraction (FNF) – The number of False Negative Controlled Releases relative to the total number of Controlled Releases. See section 6.2.3.
- False Positive (FP) – A Detection reported by a Performer that cannot be attributed to a Controlled Release. See section 6.1 for Classification of Detections.
- False Positive Fraction (FPF) – The number of False Positive Detections relative to the total number of Detections. See section 6.2.2.
- Final Report – A report issued by the Test Center after the conclusion of testing. See section 8.
- Localization Accuracy (LA)– A measure of the distance between the location of an emission estimated by a Performer and the location where a Controlled Release occurred. In this protocol location accuracy is 2D. Three localization accuracies may be calculated based on an Equipment Unit, a single latitude-longitude coordinate pair, or a pair of coordinates indicating a bounding box reported by the Performer (see sections 6.2.6, 6.3.5 and 6.3.6 respectively).
- Localization Precision (LP) – A measure of the area to which an emission source is attributed by a Performer. Two Localization Precisions may be calculated based on an Equipment Unit, or a pair of coordinates indicating a bounding box reported by the Performer (see sections 6.2.5 and 6.3.8 respectively).
- Performer – A single participant in the testing under this protocol. The Performer includes the personnel and an Emission Detection System.
- Probability of Detection (PD) – Fraction of Controlled Releases, over an extended test period, that the Performer reported as Detections. The Probability of Detection may vary across

independent variables such as the emission rate and/or the meteorological conditions, resulting in a Probability of Detection curve or surface. See section 6.2.1.

- Quantification Accuracy (QA) – A measure of the difference between the emission rate estimated by a Performer and the metered emission rate of a Controlled Release. Quantification Accuracy may be represented as an absolute difference, or as a percentage difference relative to the metered emission rate (see sections 6.3.1 and 6.3.2 respectively).

- Quantification Precision (QP) – A measure of the difference between the upper and lower confidence limits reported by a Performer for an emission rate estimate (see sections 6.3.3 and 6.3.4 respectively).

- Single-Blind – An experimental procedure in which the controlled testing facility knows the location and emission rate of all emissions, but operators of the systems being tested (i.e. the Performers) do not.

- Survey Solution – A survey solution includes the Emission Detection System, personnel, and methodology to perform an emission survey.

- Survey Time – The time required by a performer to complete an emission survey measured as the difference between the time of arrival at a facility and time of departure from the facility.

- Test Center – The location at which testing is performed under this protocol. The term 'Test Center' includes the physical facilities, the personnel performing the evaluation, and any supporting software or analysis.

- True Positive (TP) – A Detection reported by a Performer that can be attributed to a Controlled Release. See Section 6.1 for classification of Detections.

3 Variables and Subscripts

The variables listed in Table 1 are used in equations throughout the protocol:

Table 1: List of variables

Variable	Description
N	Total number across all experiments
n	Number during a single experiment or subset of all experiments
t	Time
FP	False Positive Detection(s)
FN	False Negative Detection(s)
PD	Probability of Detection
FPF	False Positive Fraction
FNF	False Negative Fraction
OF	Operational Factor
QA	Quantification Accuracy
QP	Quantification Precision

LA	Localization Accuracy
LP	Localization Precision
DT	Detection Time

The subscripts in Table 2 are used in equations throughout the protocol:

Table 2: List of subscripts

Subscript	Description
CR	Controlled Release(s)
RD	Reported Detection(s)
TP	True Positive Detection(s)
FP	False Positive Detection(s)
FN	False Negative Detection(s)
Unit	Equipment Unit Precision
Group	Equipment Group Precision
Facility	Facility Precision

4 System Types Covered by Testing

Survey solutions include many designs and configurations, but generally use a handheld or otherwise mobile Emission Detection System deployed periodically at a facility to detect and locate emission sources to the equipment unit-level or component-level.

Survey solutions typically differ from "screening solutions" by the precision to which attribution can be established. In general, screening solutions are performed remotely downwind or from aircraft with limited access to a facility and establish equipment group- or facility-level attribution of emission detections. A follow-up survey solution is generally required to pinpoint emission sources to the equipment unit or component level. Modifications to this protocol to test Screening Solutions are included in section 9. Section 10 provides adaptations to the protocol for testing remote sensing systems such as aircraft or satellite solutions.

5 Test Method

Testing consists of three activities – documentation of system under test, emission surveys, and reporting.

5.1 Documentation of System Under Test

The configuration of the survey solution under test shall be documented and reported. Documentation must be sufficient for a reviewer to fully identify the *as tested* revision and configuration of the survey solution.

5.1.1 Documentation Requirements

At a minimum, documentation shall include:

1. Detailed description of system configuration and primary components including the sensor and deployment platform. Additionally, the location (latitude, longitude, height) of auxiliary components such as meteorological station or any other equipment installed at or near the Test Center must be recorded.

2. Model number of each component in (1).

3. Revision number of software installed in each component in (1) that includes performer-specific software components, revisions, or customizations.

4. Revision number of any software analytics installed offsite.

5. Detailed description of the methodology used by the performer during emission detection surveys.

6. Confidence level at which emission detection data are reported.

7. Personnel are considered part of the deployed survey solution and should match typical deployment in field use. The number of people participating in surveys and their roles must be documented. Additional performer personnel may not interact with the survey team during the experiments, either onsite or via remote contact methods. Any remote personnel participating in the survey in any fashion should be documented as part of the survey team in this section.

Installation documentation should be considered public information, and Performers should not include proprietary information (e.g. algorithmic details, reasons for locating sensor in specific locations, performance data of sensors, etc.) as part of this documentation.

5.1.2 Testing Cautions

Performers should recognize results are applicable only to the system *as tested and documented.* Future reviewers of results will be interested in whether systems proposed for field deployment include the same quality of sensors, deployment platform, and methodology as were tested under the protocol. Deploying more sensors, higher cost-performance sensors, more extensive analytics, or more human intervention than would be typical in field deployments may render the results produced in these tests inapplicable to future field deployments, regulatory applications, or other uses of the test results.

5.2 Emission Surveys

Performer personnel will be present at the test center to perform emission detection. For each experiment, the following process will be followed:

1) A facility boundary will be defined by the test center for the experiment. The facility may include one or more equipment groups. The facility designation may change between consecutive experiments.

2) An experimental design point including a predetermined number of controlled releases on the designated facility will be established by the test center. An experimental design point

may include zero controlled releases. Controlled releases in an experimental design point may be steady, unsteady, or intermittent as described in section 7.2. For each Controlled Release, the Test Center will record the location, timing, gas composition, metered emission rate, and uncertainty (95% confidence limit) of the metered emission rate.

3) Performers will complete an emission survey of the facility. Emission surveys must be completed according to the performer methodology as documented in section 5.1. The performer will record survey data as described in section 5.3.1 and section 5.3.3.

4) Performers will submit survey data from (3) to the Test Center. The Test Center will record the date and time which Performer emission detection reports are received and store them for results analysis.

5.3 Reporting

Experiment and detection data must be reported as described in this section. Performers are encouraged to submit experiment and detection reports to the test center daily during the testing period. Results must be submitted by performers within 1 week of the end of the testing period. Reports received later than 1 week after the experiment will be discarded. The date emission reports are submitted will be recorded by the test center and used in result analysis.

5.3.1 Survey Summary

A survey summary will be reported for each experiment completed. Each survey summary will include the data fields listed in Table 3. Surveys which are missing fields will not be considered in the analysis. If a survey takes more than one day (e.g. on a very large test site), the Performer will provide one report for each day. The total survey time will be computed by the addition of the *SurveyTime* indicated in the reports.

Table 3: Survey summary data fields

Field	Description	Acceptable Values	Mandatory or Optional
ExperimentID	A unique ID assigned to the individual experiment. *This number should be **incremented** for every experiment performed.* Duplicate numbers will be assumed to be multiple transmissions of the same report; only one (arbitrarily chosen) report will be logged. The ExperimentID should be synchronized between the performer and the test facility.	Integer	Mandatory
FacilityID	Facility which this experiment was performed on	Defined by test center	Mandatory

StartDateTime	Time (UTC) which survey was started formatted as *yyyy/mm/dd_hh:mm*	Formatted DateTime	Mandatory
EndDateTime	Time (UTC) which survey was completed formatted as *yyyy/mm/dd_hh:mm*	Formatted DateTime	Mandatory
SurveyTime	Elapsed time during survey formatted as *hh:mm*	Formatted Time	Mandatory

5.3.2 Facility Quantification Data

Facility-level quantification estimates may be provided for each survey. The fields identified as mandatory in Table 4 must be included for each facility-level quantification estimate provided. A performer may provide multiple facility level quantification estimates for a single experiment to indicate the emission rate estimate of different gas species.

Table 4: Facility-level quantification estimate data fields

Field	**Description**	**Acceptable Values**	**Mandatory or Optional**
FacilityReportID	A unique ID assigned by the performer to the individual facility-level quantification report. *This number should be **incremented** for every report sent.* Duplicate numbers will be assumed to be multiple transmissions of the same report; only one (arbitrarily chosen) report will be logged. The increment amount between reports is arbitrary and need not be constant; report ID should never be decremented.	Positive Integer	Mandatory
ExperimentID	A unique ID assigned to the individual experiment. *This number should be **incremented** for every experiment performed.* Duplicate numbers will be assumed to be multiple transmissions of the same report; only one (arbitrarily chosen) report will be logged.	Integer	Mandatory

	The ExperimentID should be synchronized between the performer and the test facility.		
FacilityID	Facility which this experiment was performed on	Defined by test center	Mandatory
Gas	The gas the survey system measured to perform a detection.	THC Methane Ethane Propane Butane	Mandatory
FacilityEmissionRate	Estimate of total emission rate from the facility for the experiment. The units of this field should be **grams per hour** of the gas specified in *Gas.*	Decimal Number ≥0	Mandatory
FacilityEmissionRateUpper	Upper estimate of total emission rate of the facility. The units of this field should be **grams per hour** of the gas specified in *Gas.*	Decimal number >0	Optional
FacilityEmissionRateLower	Lower estimate of total emission rate of the facility. The units of this field should be **grams per hour** of the gas specified in *Gas.*	Decimal number ≥0	Optional

5.3.3 Detection Data

Each emission source detected during an experiment will be reported using a detection report. Each detection report should include, at minimum, the fields listed as mandatory in Table 5. Detection reports which are missing mandatory fields will not be considered in the analysis. Optional fields listed in Table 5 may be included if the performer is capable of reporting these additional data. Performers that are capable of reporting optional data fields are encouraged to do so in order to support the evaluation of additional metrics under the same series of experiments.

Table 5: Detection data fields

Field	**Description**	**Acceptable Values**	**Mandatory or Optional**
DetectionReportID	A unique ID assigned by the performer to the individual detection report. *This number should be **incremented** for every detection report sent.* Duplicate numbers will be assumed to be multiple transmissions of the same report; only	Positive Integer	Mandatory

	one (arbitrarily chosen) report will be logged. The increment amount between reports is arbitrary and need not be constant; report ID should never be decremented.		
ExperimentID	The experiment this emission source was detected during.	Integer	Mandatory
EmissionSourceID	A unique ID assigned by the performer to the individual emission source the detection report refers to. Multiple detection reports may refer to the same *EmissionSourceID* **only** to report emission rate estimates for two different gas species.	Positive Integer	Mandatory
EquipmentUnit	The Equipment Unit ID on which the emission was detected. An emission source attributed within the defined Facility but not attributed to an Equipment Unit should be reported as `OTHER`. An emission source detected by a survey system but not attributed to the Facility may be reported as `OFF_FACILITY`.	Defined by test center	Mandatory
Latitude1	If a bounding box is reported, the southern-most latitude of the bounding box in decimal degrees. Otherwise, the estimated latitude of the emission source location in decimal degrees.	Defined by test center	Optional
Latitude2	If a bounding box is reported, the northern-most latitude of the bounding box, in decimal degrees. Otherwise this field may be omitted or reported as `NULL`.	Defined by test center	Optional
Longitude1	If a bounding box is reported, the eastern-most longitude of the bounding box, in decimal degrees. Otherwise, the estimated longitude of the emission source in decimal degrees.	Defined by test center	Optional

Longitude2	If a bounding box is reported, the western-most longitude of the bounding box, in decimal degrees. Otherwise this field may be omitted or reported as `NULL`.	Defined by test center	Optional
Gas	The gas the survey system measured to perform a detection.	Defined by test center	Mandatory
EmissionRate	Estimated emission rate of the source. The units of this field should be **grams per hour** of the gas specified in *Gas.*	Decimal number >0	Optional
EmissionRateUpper	Upper estimate of emission rate of the source. The units of this field should be **grams per hour** of the gas specified in *Gas.* If *EmissionRateUpper* is reported, *EmissionRateLower* must also be provided.	Decimal number >0	Optional
EmissionRateLower	Lower estimate of emission rate of the source. The units of this field should be **grams per hour** of the gas specified in *Gas.* If *EmissionRateLower* is reported, *EmissionRateUpper* must also be provided.	Decimal number ≥0	Optional

Each detection report should refer to a single emission source identified by the performer using an *EmissionSourceID.* Two detection reports should refer to the same EmissionSourceID **only if** a performer would like to report quantification estimates for two different gases. In this case both detection reports should list the same location data (EquipmentUnitID, Latitude1, Latitude2, Longitude1, Longitude2). If two detection reports refer to the same EmissonSourceID but include different location data, only one (arbitrarily selected) set of location data will be used in the analysis of metrics.

6 Performance Metrics

To evaluate performance metrics, detection reports and Controlled Releases will first be classified as True Positive or False Positive Detections. Results will then be used to evaluate primary and secondary metrics. Primary metrics will be evaluated for all solutions under test; secondary metrics will be evaluated for systems that report optional data fields.

Caution: Performance metrics and the operational and environmental conditions during the experiment will be reported in the Final Report (see section 8). Performance metrics may only be applicable under

the conditions tested and caution should be exercised in extrapolating test results to operational or environmental conditions not encountered during the testing period.

6.1 Classification of Detections

Detection reports which refer to the same *EmissionSourceID* will be grouped together as one "Detection" during the classification process.

Prior to classification, Detections where the *EquipmentUnit* is `OFF_FACILITY` will be removed from the classification process. Detections removed in this step will not be classified as True Positives or False Positives.

By default, the classification methodology below will be performed where controlled releases and detection reports are sorted by emission rate. Performers may elect for the classification to be performed prioritized by localization accuracy instead of emission rate. This must be elected for all detections and controlled releases. If prioritization by localization is selected, performers must submit latitude and longitude for all reported detections.

The Test Center will perform the classification using the following process for each experiment:

1) The list of Controlled Releases performed within the Facility boundary during the experiment will be sorted by Equipment Unit, then by emission rate in descending order.

2) The list of all Detections where *EmissionStartDateTime* is between the start time and end time of the experiment will be sorted by *EquipmentUnit*, then by *EmissionRate* (if reported) in descending order. If a single detection includes *EmissionRate* reported for multiple gas species, the sorting will consider the sum of all gas species.

3) For each Controlled Release in (1), if a Detection in (2) is reported on the same Equipment Unit, the Detection and Controlled Release will be paired as a True Positive Detection and removed from further matching. True Positives matched in this step will be identified as correct Equipment Unit Detections (see section 6.2.5).

4) The list of Controlled Releases and list of Detections remaining after (3) will be resorted by Equipment Group, then by emission rate in descending order.

5) For each Controlled Release in (4), if a Detection in (4) is reported on the same Equipment Group, the Detection and Controlled Release will be paired as a True Positive Detection and removed from further matching. True Positives matched in this step will be identified as correct Equipment Group Detections (see section 6.2.5).

6) The list of Controlled Releases and list of Detections remaining after (5) will be resorted by Facility, then by emission rate in descending order.

7) For each Controlled Release in (6), if Detection in (6) is reported on the same Facility, the Detection and Controlled Release will be paired as a True Positive Detection and removed

from further matching. True Positives matched in this step will be identified as correct Facility Detections (see section 6.2.5).

8) Any Controlled Releases remaining after (7) will be identified as False Negative Detections.

9) Any Detections remaining after (7) will be identified as False Positive Detections.

This process will classify all Detections attributed to the Facility as either True Positive or False Positive, and all Controlled Releases occurring on the Facility as either True Positive or False Negative, and result in the three possible scenarios illustrated in Table 6 for each experiment. If the number on Controlled Releases, n_{CR}, is greater than the number of reported Detections, n_{RD}, then each reported Detection will be classified as True Positive and the remaining Controlled Releases will be classified as False Negative. If the number of Controlled Releases is equal to the number of reported Detections, then each reported Detection will be classified as True Positive and no Controlled Releases will be classified as False Negative. If the number of Controlled Releases is less than the number of reported Detections, then each Controlled Release will be classified as True Positive and the remaining Detections will be classified as False Positive.

Table 6: Detection classification outcomes for each experiment

Relationship between n_{CR} and n_{RD}	Number of True Positives, n_{TP}	Number of False Positives, n_{FP}	Number of False Negatives, n_{FN}
$n_{CR} > n_{RD}$	n_{RD}	0	$n_{CR} - n_{RD}$
$n_{CR} = n_{RD}$	n_{RD}	0	0
$n_{CR} < n_{RD}$	n_{CR}	$n_{RD} - n_{CR}$	0

6.2 Primary Metrics

The following performance metrics have been identified as primary metrics:

6.2.1 Probability of Detection

Probability of Detection (PD) will be calculated as a curve or surface. Detection data will be binned by conditions (environmental and controlled). For each set of conditions, the PD will be calculated as the number of True Positive Detections divided by the sum of the number of True Positive Detections and False Negative Detections in the relevant conditions:

$$PD|_x = \frac{n_{TP}}{n_{TP} + n_{FN}}\Bigg|_x$$

Where x is the combination of conditions at which the PD is evaluated at.

PD results will be calculated for the following three cases unless otherwise agreed by the Performer and Test Center:

1) PD vs emission rate

2) PD vs average wind speed

3) PD vs emission rate and average wind speed

The Performer may request PD be calculated against an independent variable other than wind speed, if they believe the performance of their solution is more impacted by another, recorded and available variable. The Performer may also request only (1) to be calculated with (2) and (3) omitted, producing only a PD curve instead of a surface or series of curves. While the Final Report will contain only the requested PD curve/surface, all data will be released, and other parties may compute other PD curves/surfaces.

6.2.2 False Positive Fraction

The False Positive Fraction will be calculated for the set of all experiments as the number of False Positive Detections divided by the total number of reported Detections.

$$FPF = \frac{N_{FP}}{N_{RD}} = \frac{N_{FP}}{N_{FP} + N_{TP}}$$

The False Positive Fraction *does not* represent the rate at which a Performer reported a Detection when there were no emissions at the Facility.

6.2.3 False Negative Fraction

The False Negative Fraction will be calculated for the set of all experiments as the number of False Negatives divided by the total number of Controlled Releases.

$$FNF = \frac{N_{FN}}{N_{CR}}$$

The False Negative Fraction does not represent the rate at which Controlled Releases were undetected by a Performer.

6.2.4 Survey Time

Survey time will be calculated as the time between the start of the emission survey and the end of the emission survey. The testing method will be designed to minimize the need to setup and breakdown equipment between consecutive surveys, allowing performers to complete a larger number of experiments in a single day of testing. For example, the time required to unpack and assemble a unmanned aerial vehicle (UAV) for an emission survey, or the time required to setup auxiliary equipment such as a meteorological measurement station or RTK GPS base station would not be included in the survey time. Therefore, survey time measured in this testing is likely less than or equal to the survey time required at a facility in a field deployment.

6.2.5 Localization Precision (Equipment Unit)

For primary metrics, localization uses only the *EquipmentUnit* provided in the detection report to determine the precision of each True Positive. Each True Positive Detection will be classified into one of three levels of precision, from most precise to least precise:

1) Correct unit: The *EquipmentUnit* was the Equipment Unit on which the Controlled Release occurred.

2) Correct group: The *EquipmentUnit* was in the Equipment Group where the Controlled Release occurred.

3) Correct Facility: The *EquipmentUnit* was within the facility boundary where the controlled release occurred.

6.2.6 Localization Accuracy (Equipment Unit)

Localization Accuracy will be calculated for the set of all experiments as the fraction of reported Detections at each level of precision.

1) Correct unit:

$$LA_{Unit} = \frac{N_{TP_{Unit}}}{N_{RD}} = \frac{N_{TP_{Unit}}}{N_{TP} + N_{FP}}$$

2) Correct group

$$LA_{Group} = \frac{N_{TP_{Group}} + N_{TP_{Unit}}}{N_{RD}} = \frac{N_{TP_{Group}} + N_{TP_{Unit}}}{N_{TP} + N_{FP}}$$

3) Correct Facility

$$LA_{Facility} = \frac{N_{TP_{Facility}} + N_{TP_{Group}} + N_{TP_{Unit}}}{N_{RD}} = \frac{N_{TP}}{N_{TP} + N_{FP}}$$

6.3 Secondary Metrics

Secondary metrics will only be evaluated when optional data fields necessary for their calculation are included in detection reports. The following performance metrics have been identified as secondary metrics:

6.3.1 Quantification Accuracy (Absolute)

Quantification Accuracy will be calculated for each True Positive Detection as the absolute difference (in g/hr) between the *EmissionRate* reported and the metered emission rate of the matched Controlled Release.

6.3.2 Quantification Accuracy (Relative)

Quantification Accuracy will also be calculated for each True Positive Detection as the relative difference (in 5) between the *EmissionRate* reported and the metered emission rate of the matched Controlled Release.

6.3.3 Quantification Precision (Absolute)

Quantification Precision will be calculated for each True Positive Detection as the absolute difference between *EmissionRateLower* and *EmissionRateUpper.*

6.3.4 Quantification Precision (Relative)

Quantification Precision will also be calculated for each True Positive Detection as the absolute difference between *EmissionRateLower* and *EmissionRateUpper normalized by the* metered emission rate of the matched Controlled Release.

6.3.5 Localization Accuracy (Single Coordinate)

Localization Accuracy will be calculated for each True Positive Detection with a single coordinate pair as the absolute difference (in meters) between the reported coordinate and the location where the *matched* Controlled Release occurred.

6.3.6 Localization Accuracy (Bounding Box)

Localization Accuracy will be calculated for each True Positive Detection with a bounding box coordinate set as the absolute difference (in meters) between the center of the reported bounding box and the location where the Controlled Release occurred. A true/false value will also be calculated for each True Positive Detection with a bounding box coordinate set to indicate if the Controlled Release was within the reported bounding box.

6.3.7 Bounding Box Accuracy

A true/false value will also be calculated for each True Positive Detection with a bounding box coordinate set to indicate if the Controlled Release was within the reported bounding box. The Bounding Box Accuracy will be calculated as the fraction of True Positive Detections with a bounding box reported where the Controlled Release was within the bounding box.

6.3.8 Localization Precision (Bounding Box)

Localization Precision will be calculated for each True Positive Detection with a bounding box coordinate set as the area (in square meters) of the bounding box.

7 Experimental Design

All testing will be performed “single-blind”. Performers *will not* be informed of the number, location(s), or emission rate(s) of controlled releases during the course of the experiment. Each survey will be performed following the steps in section 5.2.

7.1 Facility to be Monitored

The Test Center will define the Facility to be monitored during each experiment using a bounding box of coordinates. The bounding box may correspond to physical infrastructure, such as a fenceline, or an implied boundary such as a property line, right of way, or easement. Consecutive experiments may be performed on the same Facility or different Facilities.

7.2 Selection of Experimental Design Points

Each Experimental Design Point will be selected by the Test Center during the test period to cover a range of emission rates. Enough Experimental Design Points should be performed in each Controlled Release emission rate of interest to evaluate a Probability of Detection curve. The Test Center will keep

track of the number of Experimental Design Points at each emission rate in a design matrix similar to the matrix illustrated in Table 7.

Table 7: Example experimental design matrix for emission detection testing.

	Emission Rate			
	Zero	Low	Med	High
Number of experiments				

7.2.1 Gas Composition

Gas composition may vary between Experimental Design Points. The range of expected gas compositions will be provided by the Test Center to the Performer in advance of testing. The Test Center will select the gas composition considering the engineering design of the controlled release system, realism of the test, completion of the test matrix, and operational safety considerations. Gas composition may vary between emission locations included in an Experimental Design Point. Gas composition for each controlled release should not vary during one Experimental Design Point.

The actual gas composition of Controlled Releases will be recorded by the Test Center for inclusion in the analysis. Gas composition will be applied to the flowrate of Controlled Releases to calculate the mass flowrate of each gas species. Probability of Detection curves derived from test results will use the mass flowrate of the gas specified in the Performer detection reports (see *Gas* in Table 5).

7.2.2 Emission Rate

One of the primary objectives of this protocol is to evaluate the Probability of Detection curve across a range of emission rates. Therefore, emission rates will be selected by the Test Center for each Experimental Design Point to extend outside – above and below – the normal operating range of the Performer(s) participating at the time of testing.

Emission rates will be restricted to within the constraints of the Test Center controlled release system. The lower limit and upper limit of the Test Center will be provided by the Test Center to the Performer(s) in advance of testing. The Test Center has the final authority to select the emission rates considering the engineering design of the controlled release system and operational safety considerations.

7.2.3 Simultaneous Controlled Releases

Experimental Design Points may include multiple simultaneous Controlled Releases.

Other emission sources may occur near the Facility during testing. These emissions may be associated and controlled by the Test Center, or unassociated with the Test Center. If the Test Center performs Controlled Releases outside the Facility boundary during the test period, the releases shall be recorded as potentially interfering sources and included in the final report data.

7.2.4 Environmental Conditions

The environmental conditions during each experiment will be summarized using the maximum, minimum, mean and standard deviation of each parameter during the full duration of the test.

One test period may include a limited range of environmental conditions. Results from multiple test periods may be combined to cover a variety of environmental conditions, provided the survey solution as documented in section 5.1 remains the same between testing periods. Combining results from multiple test periods may be required to produce a probability of detection curve where one axis is an environmental condition such as wind speed.

8 Final Report

The Test Center will perform the classification of detections and calculation of metrics after all experiments are completed and either (a) detection reports have been provided by the Performer for all experiments or (b) one week has elapsed since the last experiment. The calculation of metrics will be performed across the full duration of the testing program.

The Test Center will provide a results report to the Performer. A copy of the original results report will be available to other parties from the Test Center, by request, with the Performers consent for release. The results report will include, at minimum, the information described in this section.

8.1 Experiment Summary

The experiment summary will include the date range in which experiments were performed, the total number of experiments and the total number of Controlled Releases. Experimental conditions will be summarized including the Controlled Release rates, Controlled Release durations, and environmental conditions included during the experiments.

8.2 Performance Metrics

Performance metrics will include all primary metrics as described in section 6.2. Secondary metrics will be reported if the Performer detection reports included the required data for their calculation. Metrics which are calculated individually for each True Positive Detection, for example Quantification Accuracy (section 6.3.1), will be included as histograms.

8.3 Documentation of Test Protocol

A copy of the test protocol utilized in the experiments will be included.

8.4 Documentation of System Under Test

Documentation of the system under test as reported by the Performer to the Test Center in section 5.1.1 will be included.

8.5 Controlled Release and Detection Data

All Controlled Release and Detection data will be included. Each True Positive, False Positive, and False Negative Detection will be included. Each Detection will include:

1) The Detection classification (True Positive, False Positive, False Negative)

2) Performer reported detection data, as received by the Test Center, including all data fields listed in Table 5 (applicable to True Positive and False Positive Detections only).

3) The Controlled Release data including timing, metered emission rate with upper and lower 95% confidence limits, Equipment Unit ID, latitude, longitude and height (applicable to True Positive and False Negative Detections only).

4) Meteorological conditions as measured by the Test Center for each Controlled Release (applicable to True Positive and False Negative Detections only).

5) Time to detect, Localization Accuracy, Localization Precision, Quantification Accuracy and Quantification Precision metrics calculated for the individual Detection (applicable to True Positive Detections only).

8.6 Flow Meter Calibrations

The Test Center will include calibration records for the flowmeters used in the experiments.

End of the protocol specification.

9 Modifications to Test Screening Solutions

This protocol may be adapted to evaluate screening solutions which cannot attribute emission detections to the equipment unit-level or better with the following changes:

1) Access to the facility may be limited to a similar amount available to the screening solution as deployed in the field;

2) The emission detection reports in section 5.3.3 may include an *EquipmentGroupID* in lieu of an *EquipmentUnitID.* The classification in section 6.1 may then begin at step 4, considering only matches at the Equipment Group- or Facility-level.

10 Aerial Survey Emission Detection And Quantification

The primary test protocol "Survey Detection and Quantification" is designed principally for close-range approaches such as optical gas imaging, where particular emphasis is given to source detection and identification amongst an array of possible sources. Emerging mobile systems with sensors placed on drones, aircraft, or satellites are being deployed to rapidly survey large number of sites. Due to the deployment methodologies, modifications to the protocol are necessary to efficiently evaluate detection limits and quantification accuracy of these aerial detection systems.

10.1 Test Method

10.1.1 Documentation of System Under Test

The configuration of the aerial survey solution under test shall be documented and reported. Documentation must be sufficient for a reviewer to fully identify the as tested revision and configuration of the survey solution. In addition to requirements from Section 5.1.1, Documentation must include the following:

- Definition of a survey measurement: For imaging or scanning solutions a survey measurement should be defined as a single pass over the facility. Some aerial solutions utilize multiple plume transects to determine emissions rates; in this case the typical number of transects required and the maximum and minimum flight altitude should be recorded.

  A technology deployed in pre-set surveys should have pre-set flight lines (e.g. racetrack or figure-8 formation) during testing in which each pass should be considered a unique measurement. A technology that conducts multiple passes for each emission, potentially at different angles of approach, should be allowed to operate as it would in the field.

- Uncertainty Type: A description of how uncertainty is reported. E.g. max/min estimate or 95% CI

10.1.2 Emissions Surveys

The following process will be followed:

1) Aerial technologies conduct repeated measurements of the release facility. A single CH4 source, or multiple sources may be utilized. Since the release facility location is generally "known" by the performer in order to enable repeat passes, the testing should be considered "partially single-blind" (location of emission source is generally known by the performer, but the emission rate is not).

2) Performers should acquire measurements using a protocol that mimics the field operation characteristics as closely as possible, following documentation given in Section 10.1.1. Each measurement should be reported independently.

3) Between experiments, the Test Center will adjust the gas release following the experimental design. The frequency and timing of these changes will not be communicated with the Performer, in accordance with the single-blind testing methodology. The test center may

allow successive measurements to be conducted at the same controlled release rate, or to adjust the release rate between each measurement.

10.1.3 Measurement Data

Performers being tested are required to report measurement data for each completed measurement. Each measurement must be included (including measurements where no gas was detected), and each entry will include the data fields listed in Table 8. In contrast to the standard protocol, for testing of aerial technologies, data collection is limited to a single workbook, the "Measurement Data" workbook.

Table 8 Aerial Survey Summary Data Fields

Field	Description	Acceptable Values	Mandatory or Optional
MeasurementID	A unique ID assigned by the performer to the individual measurement. This number should be incremented for every measurement.	Positive Integer	Mandatory
StartDatetime	The datetime of the start of the measurement.	Date Time MM/DD/YYYY HH:MM:SS	Mandatory
EndDatetime	The datetime of the end of the measurement.	Date Time MM/DD/YYYY HH:MM:SS	Mandatory
Gas	The gas the survey system measured to perform a detection	THC Methane Ethane Propane Butane	Mandatory
Altitude	The flight altitude of the measurement (Feet above ground level). If multiple transects were conducted as part of the measurement report the minimum and maximum	Integer	Mandatory

	flight altitudes as a pair: [min, max]		
EmissionRate	Estimate of total emission rate from the release location. The units of this field should be kilograms per hour of the gas specified in *Gas*	Decimal number > 0	Optional
EmissionRateUpper	Lower estimate of the emission rate from the release location. The units of this field should be kilograms per hour of the gas specified in Gas	Decimal number > 0	Optional
EmissionRateLower	Upper estimate of the emission rate from the release location. The units of this field should be kilograms per hour of the gas specified in Gas	Decimal number > 0	Optional
WindSpeed	If applicable, report here the wind speed estimate used in computing the total emission rate. The units of this field should be meters per second. If not applicable, report "N/A"	Decimal number > 0	Mandatory
WindDirection	Azimuthal wind direction recorded	Degrees from North	Optional

497

10.2 Selection of Experimental Design Points

The range of emission sizes and fraction of zeros should be selected based on the anticipated capabilities of the solution under test. This should be determined through discussion with the Performer about their assessment of their capabilities.

Experiment design may be developed with a focus on evaluating detection capabilities, or evaluating quantification capabilities, or evaluating both detection and quantification capablities. When evaluating detection capabilities, controlled release levels should target the Performer's expected detection limits, with buffer on the low and high end. Quantification testing should include the emission levels anticipated in the field and typically will be performed with all experiments above the anticipated detection level of the solution.

10.3 Classification of Measurements

Measurements will be matched to controlled releases by the test center using the StartDatetime and EndDatetime reported by the performer. Measurements where the emission plume was not established prior to the StartDatetime reported by the performer will be excluded from classification, and therefore not be interpreted as True Positive or False Negative, and will not be considered in quantification accuracy metrics. The test center shall establish and report the criteria used to determine if the plume was established.

10.4 Aerial Detection and Quantification Best Practices

10.4.1 Location

The test location (or set of locations) should be at least 1 km from potential confounding methane sources or environmental obstructions (e.g. standing water can interfere with hyperspectral methane imaging technologies). Other considerations about the environment which may impact the detection and quantification capabilities of these systems are the reflectivity of the ground surface, the type of ground cover, and the regional meteorology including humidity, typical wind speeds and cloud cover.

10.4.2 Gas Flow Metering

Appropriate metering equipment should be utilized for the anticipated range of release rates. Uncertainty in the metered flow rate should be reported by the test center for all data points considering the uncertainty in gas composition, meter accuracy, and flow stability during the measurement.

Expansion of gas from a pressurized source often results in significant change in temperature due to Joule-Thompson cooling. In extreme cases, when releasing at high release rates, this may impact the dispersion of the gas and/or the detection capabilities of the solutions. These effects can be managed by incorporating a heat exchanger in the release system design to allow the gas to warm back up prior to being released to atmosphere. Other methods to manage this challenge are by limiting the duration of releases or limiting the maximum pressure in the system to reduce the overall expansion ratio of the gas.

10.4.3 Wind Metering

The test center should measure the wind speed and direction near the release point for use in the analysis. It is often useful to evaluate detection limits in terms of mass flow normalized by wind speed in units of (kg/h)/(m/s). Since the wind speed varies with altitude, it is best practice to select a standardized altitude at which to measure the wind speed, and to use in the normalization (e.g. 10m agl).

10.5 Acknowledgement

We acknowledge the significant contribution to this chapter from Evan David Sherwin and Jeff Scott Rutherford of Stanford University.

11 Example Application: Testing at METEC

This section contains information specific to testing when METEC is the test center.

11.1 Data Reporting

Data will be reported in tabular format using a Microsoft Excel workbook. A template will be provided to performers by METEC. The template will include data validation to minimize data entry errors. Some summary fields will be calculated.

Acceptable values for all reporting fields set by the test center are listed in Table 9 for testing at METEC under this protocol.

Table 9: Accepted values for data fields during testing at METEC

Field	Acceptable Values
FacilityID	1 2 3 4 5
Gas	THC NMHC METHANE ETHANE PROPANE BUTANE
EquipmentUnit	1W-1 1S-2 1T-1 2W-1 2S-1 2T-1 3W-1 3W-2 3W-3 3S-1 3S-2 3T-1 3T-2 4W-1 4W-2 4W-3 4W-4 4W-5 4S-1 4S-2 4S-3 4S-4 4F-1 4F-2

	4T-1 4T-2 4T-3 5W-1 5W-2 5W-3 5S-1 5S-2 5S-3 6S-1 6D-1 6C-1 6F-1 OTHER OFF_FACILITY
Latitude1	min = 40.594800 max = 40.596550
Latitude2	min = 40.594800 max = 40.596550
Longitude1	min = -105.141480 max = -105.138400
Longitude2	min = -105.141480 max = -105.138400

11.2 Definition of Facilities to be Surveyed

Testing at METEC under this protocol will be performed on five facilities as defined in this section. Facilities designations have been selected to include a similar number of equipment units in each. The facilities' boundaries are defined by the maximum and minimum latitude and longitude listed in Table 10 which form the bounding boxes shown in Figure 1 . Controlled releases may occur anywhere within the designated facility boundary during an experiment under this protocol. Equipment unit IDs for use in detection reports are shown in Figure 2 through Figure 6. Performers will receive kml data including the facility boundaries and markers with equipment unit IDs prior to testing.

Table 10: Facilities' boundaries

Facility ID	Latitude		Longitude	
	Minimum	Maximum	Minimum	Maximum
1	40.595000°	40.595180°	-105.140300°	-105.139360°
2	40.595400°	40.596050°	-105.140960°	-105.140700°
3	40.595550°	40.596050°	-105.140550°	-105.140050°
4	40.595550°	40.596050°	-105.140000°	-105.139250°
5	40.594860°	40.595400°	-105.141180°	-105.140700°

All figures below are shown with North up.

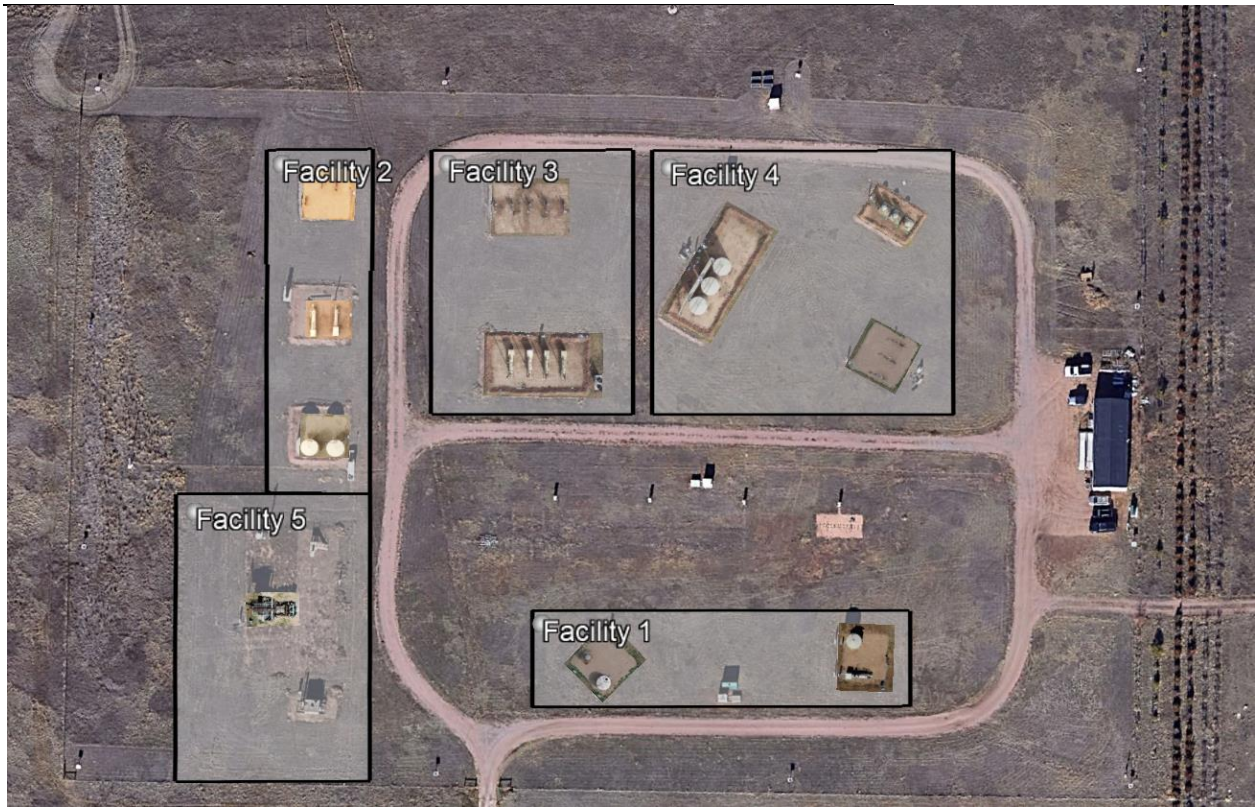
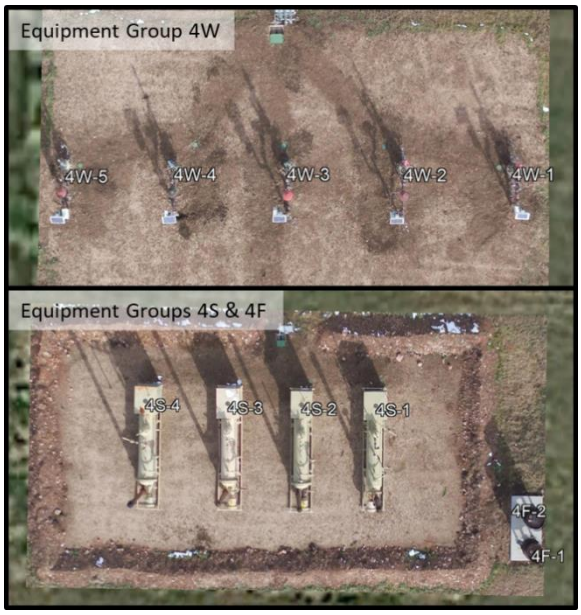


Figure 1: Facility boundaries

Figure 2: Facility 1 equipment group and equipment unit IDs

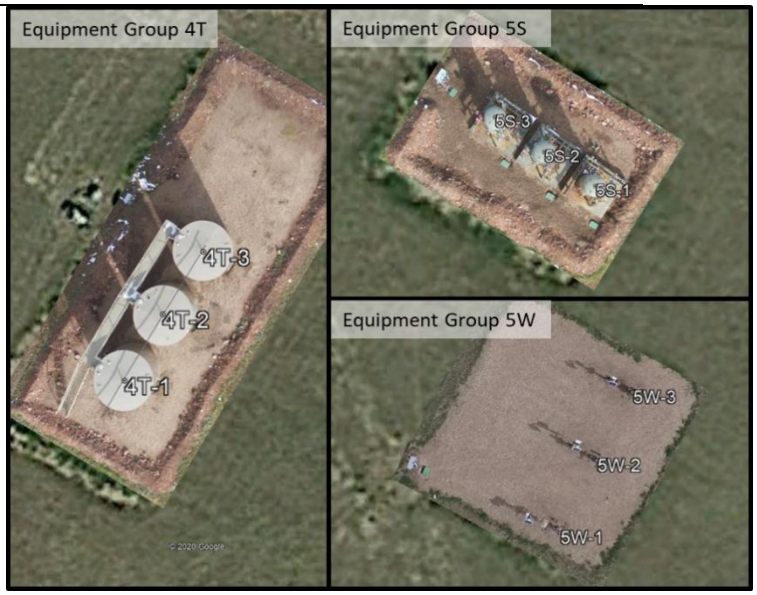
568

569 *Figure 3: Facility 2 equipment group and equipment unit IDs.*



570

571 *Figure 4: Facility 3 equipment group and equipment unit IDs.*



572

573 *Figure 5: Facility 4 equipment group and equipment unit IDs.*

574

575 *Figure 6: Facility 5 equipment group and equipment unit IDs.*

576 11.3 Maximum Survey Time and Testing Efficiency

577 METEC will support up to 3 performers testing simultaneously under this protocol. During testing,
578 multiple performers will not measure the same facility simultaneously. Facility assignments, and order

of progression between facilities, will change between consecutive experiments. To maximize testing efficiency, METEC will offer testing periods one week in length, each with a maximum survey time fixed for the full testing period. In order to maintain an efficient test program for all performers, METEC will strictly enforce the maximum survey time. Performers should test during a period with the shortest maximum survey time that is aligned with their expected survey speed in order to maximize the number of experiments performed in a week. Initially METEC will offer 2 periods with a maximum survey time of 20 minutes per survey and 40 minutes per survey, respectively.

11.4 Gas Composition

Controlled releases at METEC will largely use compressed natural gas (CNG). METEC measures the gas composition to allow emission rates to be reported as whole gas or individual species (e.g. methane). While CNG composition at METEC varies over time, to date METEC gas composition has been measured at 85%-90% methane, 8-12% ethane, and ≈1% propane. While the CNG is odorized using mercaptans, some testing may be performed without the release of odorized gas. METEC may include some experimental design points with higher ethane and propane content in the test matrix.

For the evaluation of methods using handheld methane detection solutions, METEC may elect to operate some or all controlled releases using non-odorized methane cylinders to reduce the influence of odorization on detection performance.

11.5 Emission Rates

The controlled release system at METEC supports emission rates between 0.375 slpm (using the Compressed Gas Association standard conditions of 70°F, 1 atm) to 375 slpm whole gas. This corresponds to a range of 15 g CH_4/hr to 15000 g CH_4/hr assuming gas composition is 100% CH_4.

11.6 Quality Control

METEC will perform some quality checking to make sure emissions are occurring as intended. METEC personnel will use a combination of audio/visual/olfactory (AVO), optical gas imaging (OGI), and portable gas monitors to validate the location of emission sources. Quality control (QC) issues will be documented including the experiment ID, date and time, and emission point affected. Detections associated with experiment IDs with QC issues will be addressed on a case-by-case basis and may be flagged for exclusion from the results analysis.

Leak surveys will be performed by the Test Center prior to testing to ensure no leaks occur from the controlled release system.